

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ESTUDOS DA LINGUAGEM

PAULA RIBEIRO E SOUZA

O EFEITO RETROATIVO DO EPLIS NAS PERCEPÇÕES, ATITUDES E AÇÕES DE PROFESSORES E ALUNOS DE UM CURSO DE FORMAÇÃO EM CONTROLE DE TRÁFEGO AÉREO

CAMPINAS
2018

PAULA RIBEIRO E SOUZA

O EFEITO RETROATIVO DO EPLIS NAS PERCEPÇÕES, ATITUDES E AÇÕES DE PROFESSORES E ALUNOS DE UM CURSO DE FORMAÇÃO EM CONTROLE DE TRÁFEGO AÉREO

Tese de doutorado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para obtenção do título de Doutora em Linguística Aplicada na área de Linguagem e Educação.

Orientadora: Profa. Dra. MATILDE VIRGINIA RICARDI SCARAMUCCI

Este exemplar corresponde à versão final da Tese defendida pela aluna Paula Ribeiro e Souza e orientada pela Profa. Dra. Matilde Virginia Ricardi Scaramucci.

CAMPINAS
2018

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.
ORCID: https://orcid.org/0000-0001-7202-531

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Estudos da Linguagem
Dionary Crispim de Araújo - CRB 8/7171

So89e

Souza, Paula Ribeiro e, 1982-
O efeito retroativo do EPLIS nas percepções, atitudes e ações de professores e alunos de um curso de formação em controle de tráfego aéreo / Paula Ribeiro e Souza. – Campinas, SP : [s.n.], 2018.

Orientador: Matilde Virginia Ricardi Scaramucci.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Efeito retroativo. 2. Avaliação de proficiência. 3. Lingua inglesa - Exames. 4. Pesquisa - Metodologia. 5. Brasil. Força Aérea Brasileira. I. Scaramucci, Matilde Virginia Ricardi. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: The washback of EPLIS on teachers' and learners' perceptions, attitudes and actions in an air traffic initial training course
Palavras-chave em inglês:
Washback effect
Proficiency assessment
English language - Examinations
Research - Methodology
Brazil. Força Aérea Brasileira
Área de concentração: Linguagem e Educação
Titulação: Doutora em Linguística Aplicada
Banca examinadora:
Matilde Virginia Ricardi Scaramucci [Orientador]
Sandra Regina Buttros Gattolin de Paula
Silvia Matravolgyi Damião
Gladys Plens de Quevedo Pereira de Camargo
Ana Cecilia Cossi Bizon
Data de defesa: 31-08-2018
Programa de Pós-Graduação: Linguística Aplicada

BANCA EXAMINADORA:

Matilde Virginia Ricardi Scaramucci

Sandra Regina Buttros Gattolin de Paula

Silvia Matravolgyi Damião

Gladys Plens de Quevedo Pereira de Camargo

Ana Cecilia Cossi Bizon

IEL/UNICAMP
2018

Ata da defesa, com as respectivas assinaturas dos membros da banca, encontra-se no SIGA - Sistema de Gestão Acadêmica.

E, tudo quanto fizerdes, fazei-o de todo o coração, como ao Senhor e não aos homens. Colossenses, 3:23.

AGRADECIMENTOS

A Deus, pela força, perseverança, discernimento e sabedoria com que me presenteou ao longo desta jornada.

À profa. Dra. Matilde Virginia Ricardi Scaramucci, por me guiar cuidadosamente na realização deste trabalho. Seu apreço e zelo pela avaliação em contexto de línguas me encantam e me motivam a buscar sempre mais conhecimentos nessa área.

Às profas. Dras. Gladys Quevedo Camargo, Sandra Regina Gattolin, Silvia Matravolgyi Damião, Ana Cecília Bizon e Maria de Fátima Amarante, cujos comentários e sugestões, nas qualificações e na defesa da tese, enriqueceram enormemente esta pesquisa.

Aos professores e instrutores de inglês da EEAR, por se colocarem dispostos e interessados em participar da pesquisa. Meu agradecimento especial ao tenente Roberto, ao suboficial Olivar e aos professores Carlos, Daniela, Daudt, Élida, Helena, Márcia Rita, Patrícia e Renata Tito, que fizeram com que este trabalho se concretizasse.

Ao diretor e ex-diretores do DECEA, comandante e ex-comandantes da EEAR, chefe e ex-chefes do SDAD, diretor e ex-diretores do ICEA, chefe e ex-chefes da Divisão de Ensino da EEAR e do ICEA, por darem todo o apoio e incentivo a esta pesquisa, e, em especial, ao Coronel Evandro José Alves, meu chefe imediato, por ter sempre argumentado a favor do meu estudo e de sua importância para o ICEA.

Às amigas, colegas de orientação e de trabalho, Adriana, Beatriz, Gabriela, Natália Andrade, Natália Guerreiro, Lígia, Luciana, Mônica, Patrícia e Raquel, pelas valiosas contribuições a esta pesquisa. Agradeço sempre a oportunidade ímpar de conviver com pessoas tão queridas e competentes como vocês.

Aos funcionários da Biblioteca e da Secretaria da Pós-Graduação do IEL, pela cordialidade e profissionalismo com que sempre me receberam e, prontamente, atenderam aos meus inúmeros pedidos durante esses quase cinco anos de estudo.

À querida amiga Deborah que, muito gentilmente e atenciosamente, revisou este trabalho, mesmo estando a chegada da pequena Luísa tão próxima.

À minha família, por ser meu porto seguro, por apoiar as minhas decisões e por se alegrar com as minhas vitórias. À minha mãe, Silvana, que me ensinou a pedir a inspiração do Espírito Santo para tudo o que eu faça. Ao meu pai, Paulo, pelo carinho com que cuidou de mim, de minha mãe, de minhas irmãs e de meu sobrinho. A saudade é imensa. Às minhas irmãs, por desejarem sempre o meu bem, e à Luiza e ao Henrique, por trazerem tanta alegria à minha vida.

Ao meu querido companheiro de vida, Thiago, por ser meu grande incentivador e admirador. Por estar ao meu lado em todos os momentos, comemorando minhas realizações, pequenas ou grandes, e me confortando nas horas de desânimo e insegurança.

RESUMO

SOUZA, P. R. **O efeito retroativo do EPLIS nas percepções, atitudes e ações de professores e alunos de um curso de formação em controle de tráfego aéreo**. 2018. 384f. Tese (Doutorado em Linguística Aplicada) − Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2018.

Esta pesquisa teve como foco de investigação o efeito retroativo do EPLIS (Exame de Proficiência em Inglês Aeronáutico do Sistema de Controle de Tráfego Aéreo Brasileiro) nas percepções, atitudes e ações de professores e alunos de um curso de formação em controle de tráfego aéreo. O EPLIS é um exame de proficiência, aplicado, anualmente, a controladores de tráfego aéreo e operadores de estação aeronáutica, desde 2007, em cumprimento às normas e regras da OACI (Organização da Aviação Civil Internacional) quanto à avaliação da proficiência dos profissionais envolvidos nas operações de voo internacionais. Em 2014, esse exame foi estendido a alunos do último semestre de um curso de dois anos de formação em controle de tráfego aéreo. Adotando um desenho metodológico misto, este estudo é composto por duas fases consecutivas. A primeira fase, denominada quantitativa, objetivou analisar padrões e tendências nas percepções e atitudes dos participantes em relação à influência do exame em diversos aspectos do ensino e aprendizagem de inglês. Para isso, foi elaborado um questionário, e as informações coletadas foram analisadas com o emprego de técnicas estatísticas. A segunda fase, de natureza qualitativa, pretendeu explicar, expandir e corroborar os resultados iniciais, gerando dados por meio de grupo focal, entrevistas individuais e observações de aulas. Os resultados mostraram que o exame influenciava as habilidades desenvolvidas, os tipos de tarefas realizadas e seus formatos, as avaliações em sala de aula e a abordagem de ensino do professor. Fatores como a proximidade do EPLIS, a proficiência dos alunos, o nível de conhecimentos de alunos e professores sobre o exame e a alta-relevância das provas de rendimento realizadas em sala de aula interferiram nesse processo. Assim, o efeito retroativo era mais intenso na última série do curso, quando aspectos do teste eram mais mencionados e discutidos, e as habilidades orais eram mais exploradas em sala de aula. Contudo, deficiências na capacitação dos professores em relação ao exame pareceram contribuir para que algumas características do teste fossem mais enfatizadas em detrimento de outras. Somado a isso, a alta-relevância das avaliações escritas fez com que o desenvolvimento da produção oral fosse diminuído em sala de aula, principalmente, nas 2ª e 3ª séries. O nível de proficiência percebido dos alunos também influenciou a intensidade do efeito. Assim, alunos mais proficientes demonstraram maior confiança, motivação e atitudes positivas, atribuindo níveis mais elevados de relevância ao exame, bem como à frequência com que buscavam informações se engajando em atividades preparatórias. Nas turmas mais avançadas, também foi constatado que os professores desenvolviam mais a produção oral, promoviam mais oportunidades de interação entre os alunos e exploravam mais atividades semelhantes às do exame. Por fim, limitações nas funções comunicativas avaliadas no exame se refletiram negativamente na sala de aula, o que fez com que tarefas com foco no gerenciamento do diálogo piloto-controlador e no uso de estratégias de negociação de sentido tendiam ser negligenciadas.

Palavras-chave: Efeito retroativo. Exames de proficiência. Inglês aeronáutico. Pesquisa de métodos mistos.

ABSTRACT

SOUZA, P. R. **The washback of EPLIS on teachers' and learners' perceptions, attitudes and actions in an air traffic control initial training course**. 2018. 384f. Tese (Doutorado em Linguística Aplicada) − Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2018.

This research aimed at investigating the washback effect of EPLIS (Aviation English Language Proficiency Exam for the Brazilian Airspace Control System) on teachers' and students' perceptions, attitudes and actions in an air traffic control initial training program. EPLIS is a proficiency test, which has been applied to Brazilian in-service air traffic controllers and aeronautical station operators every year since 2007, in compliance with ICAO's (International Civil Aviation Organization) guidelines pertaining to the proficiency assessment of professionals involved in international flight operations. In 2014, pre-service air traffic controllers in Brazil started sitting EPLIS in the last semester of their two-year training program. Adopting a mixed methods design, this study comprised two consecutive phases. In the first phase, the focus was to identify patterns and trends of perception and attitude among participants in relation to the influence of EPLIS on various aspects of English teaching and learning. To this end, a questionnaire was designed and statistical analyses were conducted on the data. In the second phase, qualitative data were collected through a focus group, individual interviews and class observations, in order to explain, expand on and corroborate the initial results. The results showed that the exam influenced the skills developed in the classroom, the types of tasks and their formats, the school achievement tests, and the teachers' teaching approach. Factors such as the proximity to the test date, the learners' proficiency level, the teachers' and learners' knowledge about the exam, as well as the high-stakes classroom tests played a role in this process. Thus, washback was more intense in the last semester of the course, when aspects of the exam were mentioned and discussed more often and oral skills were further explored in the classroom. However, gaps in teachers' training in the test seemed to contribute to some characteristics being overemphasized, to the detriment of other exam aspects. Moreover, the high-stakes school paper and pencil tests led to a decrease of oral production in the classroom, especially in the 2nd and 3rd grades. Learners' perceived level of English also influenced washback intensity. Learners with a higher proficiency demonstrated greater confidence, motivation and positive attitudes towards EPLIS, rating test relevance higher, as well as the frequency with which they searched for more information and engaged in preparation activities. In higher level classes, it was found that teachers developed the oral production skills further, promoted more opportunities for interaction among learners and explored more activities similar to those in the exam. Lastly, limitations in the communicative functions assessed in the exam had a negative impact in the classroom, therefore tasks that focused on the management of pilot-controller dialogue and on the use of meaning negotiation strategies tended to be overlooked.

Key-words: Washback effect. Proficiency exams. Aviation English. Mixed methods research.

LISTA DE QUADROS

Quadro 1 − Hipóteses de efeito retroativo 42
Quadro 2 − Modelos de pesquisa de métodos mistos 89
Quadro 3 − Panorama das etapas da pesquisa 92
Quadro 4 − O ensino de inglês no curso BCT 99
Quadro 5 − A estrutura e as características do EPLIS 101
Quadro 6 − A estrutura do questionário do professor 112
Quadro 7 − A estrutura do questionário do aluno 114
Quadro 8 − Valores de referência para TDE 125
Quadro 9 − Perfil dos participantes das observações 131
Quadro 10 − Cronograma de observação de aulas 132
Quadro 11 − Cronograma de entrevistas 134
Quadro 12 − Participantes do Grupo Focal 136
Quadro 13 − Efeito retroativo do EPLIS nas percepções dos professores 142
Quadro 14 − Estrutura fatorial do construto 1 177
Quadro 15 − Estrutura fatorial do construto 2 182
Quadro 16 − Estrutura fatorial do construto 3 187
Quadro 17 – Aulas observadas na 2ª série turma G 271
Quadro 18 – Aulas observadas na 2ª série turma A – parte 1 274
Quadro 19 − Aulas observadas na 2ª série turma A – parte 2 276
Quadro 20 − Aulas observadas na 2ª série turma B 278
Quadro 21 − Aulas observadas na 3ª série turma G – parte 1 280
Quadro 22 − Aulas observadas na 3ª série turma G – parte 2 282
Quadro 23 − Aulas observadas na 3ª série turma B – parte 2 283
Quadro 24 − Aulas observadas na 3ª série turma B – parte 1 285
Quadro 25 − Aulas observadas na 2ª série turma D 287
Quadro 26 − Aulas observadas na 2ª série turma C 289
Quadro 27 − Aulas observadas na 2ª série turma C 291
Quadro 28 − Aulas observadas na 3ª série turma F 293
Quadro 29 − Aulas observadas na 3ª série turma D 295
Quadro 30 – Aulas observadas na 3ª série turma E 296
Quadro 31 − Aulas observadas na 2ª série turma G 299
Quadro 32 − Aulas observadas na 2ª série turma E 301
Quadro 33 − Aulas observadas na 3ª série turma G 304
Quadro 34 − Aulas observadas na 3ª série turma E 307
Quadro 35 − Aulas observadas na 3ª série turma F 308
Quadro 36 − Aulas observadas na 4ª série turma E 311
Quadro 37 − Aulas observadas na 4ª série turma C 314
Quadro 38 − Aulas observadas na 4ª série turma F 316
Quadro 39 − Aulas observadas na 4ª série turma A 318

LISTA DE FIGURAS

Figura 1 − Modelo básico de efeito retroativo 44
Figura 2 − Modelo de significância atribuída a um teste 45
Figura 3 − Modelo de efeito retroativo 47
Figura 4 − Modelo de inovação curricular 55
Figura 5 − Matriz progressiva de validade 74
Figura 6 – Experiência no ensino de inglês aeronáutico e conhecimento sobre o exame 145
Figura 7 − Experiência no ensino de inglês aeronáutico e a prática pedagógica 146
Figura 8 − Experiência no ensino de inglês aeronáutico e as atividades em sala de aula 146
Figura 9 − Experiência no ensino de inglês aeronáutico e as avaliações em sala de aula 148
Figura 10 − Diagrama da estrutura fatorial do questionário do aluno 189
Figura 11 – Diagrama do efeito retroativo do EPLIS no curso BCT 332

LISTA DE GRÁFICOS

Gráfico 1 − Pressão exercida pelo EPLIS nos professores 143
Gráfico 2 − Experiência no ensino de inglês aeronáutico 150
Gráfico 3 − Dispersão das respostas dos subgrupos 1 e 2 em relação às tarefas da fase I do EPLIS .. 152
Gráfico 4 − Dispersão das respostas dos subgrupos 1 e 2 em relação às tarefas da fase II do EPLIS . 153
Gráfico 5 − Dispersão das respostas dos subgrupos 1 e 2 em relação às dúvidas dos alunos 161
Gráfico 6 − Nível de proficiência linguística autodeclarado 173
Gráfico 7 − Proficiência autodeclarada por habilidade linguística 174
Gráfico 8 − Extração dos componentes no construto 1 175
Gráfico 9 − Extração dos componentes no construto 2 179
Gráfico 10 − Extração dos componentes no construto 3 185

LISTA DE ABREVIATURAS E SIGLAS

AC	Análise de Conteúdo
ACC	Centro de controle de área
ACP	Análise de Componentes Principais
APP	Centro de controle de aproximação
ANAC	Agência Nacional de Aviação Civil
ATC	*Air traffic control*
BCT	Básico em Controle de Tráfego Aéreo
COLT	*Communicative Orientation of Language Teaching*
CHT	Certificado de Habilitação Técnica
CTA	Controle de Tráfego Aéreo
DECEA	Departamento de Controle de Tráfego Aéreo
DEPENS	Departamento de Ensino da Aeronáutica
DOC 9835	Manual de Implementação dos Requisitos de Proficiência Linguística
EEAR	Escola de Especialistas de Aeronáutica
EPLIS	Exame de Proficiência em Inglês Aeronáutico do SISCEAB
ESP	*English for Specific Purposes*
GEIA	Grupo de Estudos em Inglês Aeronáutico
Infraero	Empresa Brasileira de Infraestrutura Aeroportuária
ICEA	Instituto de Controle do Espaço Aéreo
IELTS	*International English Language Testing System*
KMO	Kaiser-Meyer-Olkin
LPRs	Requisitos de Proficiência Linguística
MCA	Manual do Comando da Aeronáutica
OACI	Organização da Aviação Civil Internacional
PCA	Plano do Comando da Aeronáutica
QA	Questionário do aluno
QP	Questionário do professor
RT	Radiotelefonia
SISCEAB	Sistema de Controle do Espaço Aéreo Brasileiro
TCLE	Termo de consentimento livre e esclarecido
TWR	Torre de controle
TDE	Tamanho de efeito
TEFL	*Teaching English as a Foreign Language*
Unicamp	Universidade Estadual de Campinas

LISTA DE SÍMBOLOS[1]

[[	Falas simultâneas
[	Sobreposição de vozes
[ ]	Sobreposição de vozes localizada – trecho curto de sobreposição
(+)	Pausas de 0.5 segundo e silêncios
(incompreensível)	Dúvidas e suposições – quando não se entende parte da fala
/	Truncamento brusco – quando o falante corta uma unidade
MAIÚSCULA	Quando uma palavra ou sílaba é produzida com ênfase ou acento forte
: ou ::	Alongamento de vogal
- - - -	Quando uma palavra é produzida silabadamente
”	Entonação para subida rápida (ponto de interrogação)
’	Entonação para subida leve (vírgula ou ponto e vírgula)
,	Entonação para descida leve ou brusca (pequena pausa)
Reduplicação	Repetição da letra ou sílaba
eh, ah, oh, ih::, mhm, ahã	Pausa preenchida, hesitação ou sinais de atenção
/.../ ou [...]	Indicação de transcrição parcial ou de eliminação
Iniciais maiúsculas	Para nomes próprios ou siglas

[1] Adaptado de Marcuschi (2008).

LISTA DE TABELAS

Tabela 1 − Retorno de resposta do QP....115
Tabela 2 − Retorno de Resposta do QA....116
Tabela 3 − Diretrizes para determinação de significância fatorial 121
Tabela 4 − Nível de conhecimento dos professores e alunos sobre o EPLIS....142
Tabela 5 − Conhecimentos dos professores sobre o EPLIS....151
Tabela 6 − Comparação entre professores e instrutores em relação aos conhecimentos sobre o EPLIS 151
Tabela 7 − Comparação entre experiência no ensino de inglês aeronáutico e conhecimentos sobre o EPLIS....154
Tabela 8 − Oportunidades de capacitação....154
Tabela 9 − Oportunidades de capacitação em porcentagem 155
Tabela 10 − Comparação entre professores e instrutores em relação às oportunidades de capacitação 155
Tabela 11 − Comparação entre experiência no ensino de inglês aeronáutico e acesso ao site do exame 155
Tabela 12 − Influência do EPLIS na sala de aula em porcentagem....157
Tabela 13 − Comparação da influência do EPLIS na sala de aula em 2014 e 2016....157
Tabela 14 − Comparação entre experiência no ensino de inglês aeronáutico e influência do EPLIS em sala de aula....158
Tabela 15 − Abordagem do exame em sala de aula 159
Tabela 16 − Abordagem do exame em sala de aula em porcentagem 159
Tabela 17 − Comparação entre professores e instrutores em relação à abordagem do exame em sala de aula 160
Tabela 18 − Comparação entre experiência no ensino de inglês aeronáutico e abordagem do exame em sala de aula....160
Tabela 19 − Influência do EPLIS no conteúdo 161
Tabela 20 − Influência do EPLIS no conteúdo em porcentagem 162
Tabela 21 − Atividades de compreensão e produção orais 163
Tabela 22 − Comparação entre professores e instrutores em relação às atividades orais em sala de aula 164
Tabela 23 − Comparação entre experiência no ensino de inglês aeronáutico e atividades orais em sala de aula....164
Tabela 24 − Atividades de leitura e escrita em sala de aula....165
Tabela 25 − Atividades de leitura e escrita em sala de aula em porcentagem 165
Tabela 26 − Preparação para o exame em porcentagem....166
Tabela 27 − Comparação entre professores e instrutores em relação à preparação para o exame 166
Tabela 28 − Comparação entre experiência no ensino de inglês aeronáutico e preparação para o exame 166
Tabela 29 – Visões e percepções sobre o EPLIS....167
Tabela 30 – Visões e percepções sobre o EPLIS em porcentagem 168
Tabela 31 − Comparação entre professores e instrutores em relação a visões e percepções sobre o EPLIS....168
Tabela 32 − Comparação entre experiência no ensino de inglês aeronáutico e visões e percepções sobre o EPLIS....169
Tabela 33 − Impacto negativo do exame 170

Tabela 34 − Impacto negativo do exame em porcentagem 170
Tabela 35 − Comparação entre experiência no ensino de inglês aeronáutico e impacto negativo do exame 171
Tabela 36 − EPLIS como exame de saída na perspectiva dos professores e instrutores 171
Tabela 37 − KMO e Teste de Bartlett no construto 1 175
Tabela 38 − Análise paralela para retenção de componentes no construto 1 175
Tabela 39 − Itens e cargas fatoriais no construto 1 com rotação varimax 176
Tabela 40 − Nível de conhecimentos específicos sobre o EPLIS 177
Tabela 41 − Nível de conhecimento dos alunos sobre o formato do exame 178
Tabela 42 − KMO e Teste de Bartlett no construto 2 179
Tabela 43 − Análise paralela para retenção de componentes no construto 2 180
Tabela 44 − Itens e cargas fatoriais no construto 2 com rotação varimax 181
Tabela 45 − Atividades de leitura, compreensão oral, produção oral e escrita 183
Tabela 46 − Atividades de produção oral relacionadas ao exame 184
Tabela 47 − KMO e Teste de Bartlett no construto 3 185
Tabela 48 − Análise paralela para retenção de componentes no construto 3 186
Tabela 49 − Itens e cargas fatoriais no construto 3 com rotação varimax 186
Tabela 50 − Média e desvio padrão das respostas dos subgrupos por proficiência 191
Tabela 51 − Resultados do teste t para o nível de proficiência 193
Tabela 52 − Tamanhos de efeito para a relação entre proficiência e componentes principais 194
Tabela 53 − Média e desvio padrão das respostas dos subgrupos por nível de conhecimento sobre o exame 195
Tabela 54 − Resultados do teste t para o nível de conhecimento sobre o exame 196
Tabela 55 − Tamanhos de efeito para a relação entre conhecimento sobre o exame e componentes principais 197
Tabela 56 − Média e desvio padrão das respostas dos subgrupos por série 198
Tabela 57 − Diferença entre séries no componente 1 do construto 1 199
Tabela 58 − Diferença entre séries no componente 2 do construto 1 200
Tabela 59 − Diferença entre séries no componente 1 do construto 2 200
Tabela 60 − Diferença entre séries no componente 2 do construto 2 201
Tabela 61 − Diferença entre séries no componente 3 do construto 2 202
Tabela 62 − Diferença entre séries no componente 4 do construto 2 202
Tabela 63 − Diferença entre séries no componente 5 do construto 2 203
Tabela 64 − Diferença entre séries no componente 1 do construto 3 203
Tabela 65 − Diferença entre séries no componente 2 do construto 3 204
Tabela 66 − Diferença entre séries no componente 3 do construto 3 205
Tabela 67 − Nível de conhecimento de alunos e professores sobre o EPLIS 206
Tabela 68 − Frequência com que o professor menciona o EPLIS em sala de aula 206
Tabela 69 − Frequência com que o professor discute aspectos do EPLIS em sala de aula 206
Tabela 70 − Frequência com que o professor trabalha questões semelhantes às do EPLIS em sala de aula 206
Tabela 71 − Frequência com que questões semelhantes às do EPLIS são trabalhadas em sala de aula 207
Tabela 72 − Proporção de atividades realizadas por série relacionadas ao exame 208
Tabela 73 − Adequação da preparação para o exame oferecida aos alunos 209
Tabela 74 − Comparativo das visões de professores/instrutores e alunos sobre o EPLIS 210

SUMÁRIO

CAPÍTULO I

A TRAJETÓRIA DA PESQUISA 20

1.1 CONSIDERAÇÕES INICIAIS 20
1.2 A COMUNICAÇÃO AERONÁUTICA, EXAMES DE PROFICIÊNCIA E EFEITOS PRETENDIDOS 23
1.3 O PROBLEMA DE PESQUISA E A JUSTIFICATIVA 27
1.4 AS PERGUNTAS DE PESQUISA 31
1.5 AS CONTRIBUIÇÕES DO ESTUDO 31
1.6 A ESTRUTURA DA TESE 33

CAPÍTULO II

OS ESTUDOS SOBRE EFEITO RETROATIVO 35

2.1 CONSIDERAÇÕES INICIAIS 35
2.2 O EFEITO RETROATIVO 35
2.3 A NATUREZA DO EFEITO RETROATIVO E OS EXAMES DE DESEMPENHO 38
2.4 O ESCOPO E A VARIABILIDADE DO EFEITO RETROATIVO 41
2.5 ESTUDOS SOBRE EFEITO RETROATIVO NOS PROGRAMAS 49
2.6 ESTUDOS SOBRE EFEITO RETROATIVO NAS PERCEPÇÕES E ATITUDES DE ALUNOS 58
2.7 PREPARAÇÃO PARA O EXAME 64
2.8 A RELAÇÃO ENTRE VALIDADE E EFEITO RETROATIVO 71
2.9 RESUMO DO CAPÍTULO 79

CAPÍTULO III

A METODOLOGIA DA INVESTIGAÇÃO 80

3.1 CONSIDERAÇÕES INICIAIS 80
3.2 OS PARADIGMAS METODOLÓGICOS NOS ESTUDOS SOBRE EFEITO RETROATIVO 80
3.3 O SURGIMENTO DA PESQUISA DE MÉTODOS MISTOS 83
3.3.1 A definição da PMM e seus propósitos 86
3.3.2 Modelos de PMM e suas principais características 87
3.4 O DESENHO METODOLÓGICO DA PESQUISA 90
3.5 O CENÁRIO DA PESQUISA 92
3.5.1 O Curso Básico em Controle de Tráfego Aéreo 93
3.5.2 Os participantes da pesquisa 94
3.5.3 As disciplinas de língua inglesa no curso BCT 94
3.5.3.1 Inglês I 96
3.5.3.2 Inglês II e III 97
3.5.3.3 Inglês ATC 98
3.5.4 O uso do EPLIS no contexto escolar 99
3.5.4.1 As especificações do EPLIS 100
3.5.4.2 A fase I do EPLIS 103
3.5.4.3 A fase II do EPLIS 104
3.6 O ESTUDO PRELIMINAR 107
3.7 A FASE QUANTITATIVA 110
3.7.1 Questionários 110
3.7.1.1 Questionário do professor (QP) 112
3.7.1.2 Questionário do aluno (QA) 113
3.7.1.3 Validação e pilotagem dos questionários 114
3.7.2 Procedimentos de coleta 115
3.7.3 As análises estatísticas 117
3.7.3.1 A análise fatorial e a análise de componentes principais 117

3.7.3.2 Os testes de hipóteses 123
3.8 A FASE QUALITATIVA 126
3.8.1 Observações *126*
3.8.1.1 Ficha de observação de aula 127
3.8.1.2 Procedimentos de coleta 131
3.8.2 Entrevistas individuais *133*
3.8.3 Grupo focal *135*
3.8.4 A análise de conteúdo *137*
3.9 RESUMO DO CAPÍTULO 139

CAPÍTULO IV

AS PERCEPÇÕES E ATITUDES DE PROFESSORES, INSTRUTORES E ALUNOS: UM ESTUDO QUANTITATIVO 140

4.1 CONSIDERAÇÕES INICIAIS 140
4.2 O EFEITO RETROATIVO NOS PROFESSORES ANTES DA IMPLEMENTAÇÃO DO EPLIS 140
4.2.1 O perfil dos professores em 2014 ..... *141*
4.2.2 A intensidade do efeito retroativo *141*
4.2.3 Padrões de resposta *144*
4.2.3.1 Conhecimento sobre o EPLIS 144
4.2.3.2 Influência do exame em sala de aula 145
4.2.3.3 Influência do exame nas avaliações de rendimento 147
4.3 O EFEITO RETROATIVO NOS PROFESSORES APÓS A IMPLEMENTAÇÃO DO EPLIS 148
4.3.1 O perfil dos professores e instrutores em 2016 ..... *149*
4.3.2 Construto 1 – Conhecimentos sobre o EPLIS *150*
4.3.3 Construto 2 – Impacto do EPLIS no ensino *156*
4.3.4 Construto 3 – Preparação dos alunos para o EPLIS *162*
4.3.5 Construto 4 – Visões e percepções sobre o EPLIS *166*
4.4 O EFEITO RETROATIVO NOS ALUNOS 172
4.4.1 O perfil dos alunos *173*
4.4.2 Construto 1 − Conhecimentos sobre o EPLIS *174*
4.4.2.1 Interpretação dos componentes no construto 1 176
4.4.3 Construto 2 – Preparação para o EPLIS *178*
4.4.3.1 Interpretação dos componentes no construto 2 181
4.4.4 Construto 3 – Visões e percepções sobre o EPLIS *185*
4.4.4.1 Interpretação dos componentes no construto 3 187
4.4.5 Diagrama da estrutura fatorial do QA *189*
4.4.6 Padrões de resposta entre os alunos *189*
4.4.6.1 Proficiência linguística autodeclarada 190
4.4.6.2 Nível de conhecimento sobre o EPLIS 194
4.4.6.3 Proximidade do exame 197
4.5 PERCEPÇÕES E ATITUDES DE PROFESSORES E ALUNOS COMPARADAS 205
4.6 RESUMO DO CAPÍTULO 211

CAPÍTULO V

AS PERCEPÇÕES, ATITUDES E AÇÕES DE PROFESSORES, INSTRUTORES E ALUNOS: UM ESTUDO QUALITATIVO 213

5.1 CONSIDERAÇÕES INICIAIS 213
5.2 AS PERCEPÇÕES E AS ATITUDES DOS PARTICIPANTES 213
5.2.1 Conhecimentos sobre o EPLIS *214*
5.2.2 A significância do exame *225*
5.2.3 As situações de ensino e aprendizagem *233*
5.2.4 As avaliações de rendimento na EEAR *253*
6.2.5 A preparação para o exame e sua natureza *265*
5.3 AS AÇÕES EM SALA DE AULA 268

5.3.1 A professora e coordenadora Pilar *269*
5.3.1.1 Aulas na 2ª série do curso BCT 270
5.3.1.2 Aulas na 3ª série do curso BCT 279
5.3.2 O professor Eduardo *286*
5.2.3.1 As aulas na 2ª série do curso BCT 286
5.2.3.2 As aulas na 3ª série do curso BCT 292
5.3.3 O instrutor João *298*
5.3.3.1 As aulas na 2ª série do curso BCT 299
5.3.3.2 As aulas na 3ª série do curso BCT 303
5.3.4 A professora Ana *309*
5.3.4.1 As aulas na 4ª série do curso BCT 310
5.3.5 Diferenças na prática pedagógica dos professores e instrutor *320*
5.4 RESUMO DO CAPÍTULO 322

CAPÍTULO VI

CONCLUSÕES E CONTRIBUIÇÕES DO ESTUDO 324

6.1 CONSIDERAÇÕES INICIAIS 324
6.2 PERGUNTAS DE PESQUISA REVISITADAS 324
6.2.1 Pergunta de pesquisa 1 ..... *324*
6.2.2 Pergunta de pesquisa 2 ..... *327*
6.2.3 Pergunta de pesquisa 3 ..... *330*
6.3 DIAGRAMA DO EFEITO RETROATIVO DO EPLIS NO CENÁRIO ESCOLAR 331
6.3 AS CONTRIBUIÇÕES DO ESTUDO 333
6.3.1 Teóricas *333*
6.3.1 Metodológicas *335*
6.3.3 Pedagógicas *336*
6.3.4 Elaboração e implementação de políticas linguísticas *336*
6.4 AS DIFICULDADES E LIMITAÇÕES DO ESTUDO 338
6.5 SUGESTÕES PARA ESTUDOS FUTUROS 340

REFERÊNCIAS 342

APÊNDICE A – Questionário inicial do professor 354

APÊNDICE B – Questionário inicial do aluno 356

APÊNDICE C – Questionário principal do professor 358

APÊNDICE D – Questionário principal do aluno 362

APÊNDICE E – Ficha de observação de aula 367

APÊNDICE F – Direcionamento para grupo focal 368

APÊNDICE G – TCLE do aluno 369

APÊNDICE H – TCLE do professor 373

ANEXO A – Parecer do CEP 377

ANEXO B – Escala de níveis de proficiência da OACI 384

CAPÍTULO I

A TRAJETÓRIA DA PESQUISA

1.1 Considerações iniciais

O uso da língua inglesa nas comunicações radiotelefônicas para a prestação do serviço de controle de tráfego aéreo internacional e de informações aeronáuticas está normatizado, desde 1944, no tratado de aviação civil internacional, conhecido por Convenção de Chicago,[2] mais especificamente, nos Anexos 1 e 10 a esse documento. Esses anexos estabelecem que as comunicações aeronáuticas entre aeronave e estação no solo devem ser conduzidas na língua utilizada pela estação, no caso de voos domésticos, ou na língua inglesa para servir aeroportos e rotas internacionais.

No entanto, em 2003, a Organização da Aviação Civil Internacional[3] (OACI) adotou uma nova política linguística em âmbito internacional, denominada Requisitos de Proficiência Linguística (LPRs), com o objetivo de assegurar que pilotos, controladores de tráfego aéreo e operadores de estação aeronáutica, envolvidos na aviação civil internacional, apresentem um nível mínimo de proficiência na língua inglesa para lidar com situações de tráfego aéreo inesperadas, incomuns ou emergenciais.

Foi, então, divulgada pela OACI uma escala contendo seis níveis de proficiência: Pré-elementar (Nível 1), Elementar (Nível 2), Pré-operacional (Nível 3), Operacional (Nível 4), Pós-operacional (Nível 5) e Experto (Nível 6) e seis aspectos linguísticos a serem avaliados: pronúncia, vocabulário, estrutura, compreensão, fluência e interação. O nível 4 foi considerado o nível mínimo para que esses profissionais pudessem operar na aviação civil internacional. Também ficou estabelecido que os profissionais que demonstrassem nível abaixo do Nível 6 deveriam ser formalmente avaliados em intervalos de tempo, conforme o nível de proficiência obtido: a cada três anos para aqueles que alcançassem o nível de proficiência 4, e a cada seis anos para o nível 5.

---

[2] A Convenção de Aviação Civil Internacional, conhecida por Convenção de Chicago, ocorreu em 1944, na cidade de Chicago, e reuniu representantes de 54 países. Nessa convenção, foi assinado o Tratado de Aviação Civil Internacional, composto por 19 anexos que estabelecem padrões e práticas recomendadas para o setor.

[3] A OACI é uma agência regulatória da Organização das Nações Unidas (ONU), criada durante a Convenção de Chicago, em 1944, com o objetivo de estabelecer os padrões e regulamentos necessários para a segurança, regularidade e eficiência da aviação civil mundial.

Somado a isso, a OACI publicou, em 2004, o Manual de Implementação dos Requisitos de Proficiência Linguística – Doc 9835 AN/453 (OACI, 2004, 2010), para auxiliar seus países membros[4] nos aspectos relativos à capacitação e avaliação do uso do inglês na comunicação aeronáutica (TOSQUI-LUCKS et al, 2016; SOUZA, 2017) e determinou que caberia a cada Estado a elaboração e aplicação de seu próprio exame ou a contratação de exames de instituições externas. Esse documento passou por atualizações e foi reeditado em 2010.

A necessidade de regular o nível dos profissionais envolvidos na aviação civil internacional é fruto da constatação de que o uso inadequado da língua inglesa e a baixa proficiência podem motivar falhas de comunicação e, consequentemente, colocar em risco a segurança de um voo, especialmente em situações de urgência e emergência, em que o nível de *stress* costuma estar elevado. Essa conclusão baseou-se em diversos relatórios da OACI sobre acidentes e incidentes aeronáuticos, ocorridos entre as décadas de 1970 e 1990, nos quais a deficiência no idioma foi apontada como fator contribuinte ou determinante.

No Brasil, o Instituto de Controle do Espaço Aéreo (ICEA) foi designado pelo Departamento de Controle do Espaço Aéreo (DECEA) como a instituição responsável pelo planejamento, desenvolvimento e aplicação dos instrumentos de capacitação e avaliação relativos à proficiência em língua inglesa dos controladores de tráfego aéreo (controladores, doravante) e operadores de estação aeronáutica brasileiros.[5] Em cumprimento à política linguística da OACI, o ICEA desenvolveu o Exame de Proficiência em Inglês Aeronáutico do SISCEAB[6] (EPLIS, doravante) em 2007 e, desde então, o tem aplicado anualmente a esses profissionais.

Exames no contexto aeronáutico, como é o caso do EPLIS, são considerados de alta-relevância, uma vez que decisões importantes são tomadas com base em seus resultados, acarretando sérias consequências para seus candidatos, instituições envolvidas e sociedade. A proficiência linguística se tornou um dos itens avaliados nas auditorias de vigilância da segurança operacional, realizadas pela OACI, nos seus 191 países membros, cujas não conformidades podem acarretar prejuízos políticos e financeiros (ARAGÃO, 2018). No Doc 9835 AN/453, a OACI alerta sobre a possibilidade de afastamento das operações internacionais ou transferência para aeroportos domésticos dos profissionais que não alcançarem o nível de

[4] Países que aderiram à Convenção de Chicago sobre Aviação Civil Internacional, sejam eles membros ou não da ONU.

[5] A avaliação da proficiência dos pilotos civis brasileiros é de responsabilidade da Agência Nacional de Aviação Civil (ANAC). Para isso, a ANAC desenvolveu o *Santos Dumont English Assessment* e é responsável pela manutenção, aplicação e certificação desse exame.

[6] SISCEAB é a sigla utilizada para designar o Sistema de Controle do Espaço Aéreo Brasileiro.

proficiência mínimo exigido. Para a sociedade, uma maior proficiência linguística por parte desses indivíduos significa a redução de problemas de comunicação relacionados ao uso inadequado do idioma, o que eleva, ainda mais, a segurança de voo. Em relação aos controladores e operadores de estação aeronáutica, o DECEA, por enquanto, adota medidas de incentivo para aqueles que atingem o nível de proficiência desejado, tais como possibilidade de trocas de serviço e redução de carga horária. Porém, existe a previsão de que o nível de proficiência se torne fator restritivo na emissão e revalidação do Certificado de Habilitação Técnica (CHT) desses profissionais. A CHT é uma licença que atesta a habilidade e a aptidão técnica do profissional para desempenhar a atividade operacional no órgão para o qual foi designado e deve ser revalidada a cada dois anos. Para garantir a excelência na prestação do serviço de controle de tráfego aéreo internacional, o DECEA estabelece ações mitigadoras na área de gestão operacional como, por exemplo, a distribuição do efetivo de controladores que obtiveram nível de proficiência 4 ou acima no EPLIS de forma igualitária nas equipes de serviço e o monitoramento das trocas de serviço para que esse equilíbrio seja mantido.

Shohamy (1993, p. 1) argumenta que "poucos instrumentos são tão poderosos, ou capazes de ditar tantas decisões, quanto os testes",[7] principalmente quando se atribui a eles um alto grau de relevância, e acrescenta que a adoção de determinado exame pelo Estado é um ato consciente de que ele é capaz de provocar e acelerar mudanças nos processos de ensino e aprendizagem que o precedem. A política linguística da OACI em relação aos requisitos de proficiência de pilotos, controladores e operadores de estação aeronáutica não foge à regra. Nesse sentido, os exames no contexto aeronáutico não são utilizados apenas como indicadores de proficiência, mas também como instrumentos "capazes de impor o que será estudado, aprendido e ensinado por professores e alunos"[8] em sala de aula (SHOHAMY, 2006, p. 103), visando à elevação do nível de proficiência na língua inglesa desses profissionais. A relação existente entre avaliação e ensino e aprendizagem é denominada efeito retroativo.

Nas últimas duas décadas, diversos pesquisadores têm se dedicado ao estudo do efeito retroativo de exames de alta-relevância no ensino e aprendizagem, assim como nas percepções e atitudes daqueles envolvidos no processo − professores, alunos, diretores, elaboradores de material didático e pais (NOBLE; SMITH, 1994; ALDERSON; HAMP-LYONS, 1996; CHENG, 1999; SCARAMUCCI, 1999; STONEMAN, 2006; TSAGARI, 2009; CHENG et al, 2011). Esses estudos têm constatado que mudanças atreladas à introdução de certo exame nem sempre são alcançadas conforme esperado, devido à natureza complexa,

[7] No original: "Few devices are as powerful, or are capable of dictating as many decisions, as tests".
[8] No original: "Tests are capable of dictating what will be studied, learned and taught by students and teachers".

multifacetada e variável do efeito retroativo. Em outras palavras, diversos fatores, além do teste em si, interferem no seu impacto, tais como as crenças, os valores e o conhecimento de alunos e professores, as exigências das escolas em que esses exames estão inseridos, o nível de relevância do teste, sua proximidade, entre outros.

Este estudo tem como objetivo investigar o efeito retroativo do EPLIS e os fatores que interagem na sua determinação nas percepções, atitudes e ações de professores e alunos da Escola de Especialistas de Aeronáutica (EEAR), onde são formados a maior parte dos controladores brasileiros. Nesta seção, apresentei a política linguística da OACI e a criação do EPLIS. Em seguida, discorro sobre a natureza da comunicação aeronáutica, as principais características dos exames de proficiência nesse contexto e os efeitos pretendidos, segundo a OACI. Passo, então, a tratar do problema de pesquisa a ser investigado e da justificativa pela escolha do tema. Por fim, apresento as perguntas de pesquisa e as contribuições do estudo para a área de Linguística Aplicada.

1.2 A comunicação aeronáutica, exames de proficiência e efeitos pretendidos

A comunicação aeronáutica via radiotelefonia é definida pela OACI como as interações trocadas por dois profissionais da aviação, controladores e pilotos, caracterizada pelo uso da fraseologia padrão e da língua comum (*plain language*).[9] A fraseologia padrão é considerada uma sublíngua, formada por frases e expressões pré-estabelecidas que permitem a troca de informações de forma clara, concisa e segura em situações normais de voo (OACI, 2010, p. 3-2). Apresenta vocabulário bastante reduzido (cerca de 400 palavras), geralmente exclusivo do contexto da aviação, e palavras funcionais, tais como pronomes, verbos de ligação e auxiliares, e algumas preposições tendem a ser excluídas ou evitadas. As frases, frequentemente, contêm nominalizações (verbos são transformados em substantivos) e cerca de 50% delas estão no imperativo ou na voz passiva (OACI, 2010, p. 3-4). As frases abaixo são exemplos de fraseologia de tráfego aéreo, retirados do MCA 100-16 (BRASIL, 2016, p. 22-23), manual que sanciona o uso da fraseologia em português e em inglês no espaço aéreo brasileiro.

a) PT AMS, unable RVSM due turbulence.

[9] As expressões *plain language* e *plain English* foram traduzidas para o português brasileiro como "língua comum" e "inglês comum" para comunicação aeronáutica, respectivamente, pelo Grupo de Estudos em Inglês Aeronáutico (GEIA).

> b) *PP EHJ, IFR flight cancelled at 13:42 Z, no traffic reported, frequency change approved.*
>
> c) *PUA 646, cleared visual approach runway 35, maintain own separation from preceding B 747, caution wake turbulence.*
>
> d) *PT OLL, descend at your discretion, no reported traffic, report leaving FL145.*

De acordo com o MCA 100-16 (BRASIL, 2016, p. 10), o uso da fraseologia no espaço aéreo é obrigatório para assegurar a padronização das comunicações, reduzir o tempo de transmissão de mensagens e, principalmente, proporcionar informações claras e precisas. Embora a fraseologia seja suficiente para lidar com a maioria das situações encontradas na prestação diária do serviço de controle de tráfego aéreo, essa linguagem pode se mostrar limitada em situações inesperadas, incomuns ou emergenciais (que envolvem desde problemas médicos com passageiros a panes na aeronave), ou até mesmo para atender algumas solicitações em situações normais de voo. Sobre isso, o próprio manual permite que frases adicionais sejam usadas, mas adverte que os profissionais devem se afastar o mínimo possível da fraseologia.

Essas "frases adicionais", mencionadas no MCA 16-100 (BRASIL, 2016), a OACI nomeia de *plain language*, língua comum para comunicação aeronáutica, e a define como sendo:

> o uso espontâneo, criativo e não codificado de determinada língua natural [no caso, o inglês], embora limitado por funções e tópicos demandados pela comunicação aeronáutica, bem como pelos requisitos específicos críticos de segurança em relação à inteligibilidade, objetividade, precisão, não ambiguidade e concisão[10] (OACI, 2010, p. 3-5).

Para se referir ao uso da língua inglesa em situações que não estão contempladas na fraseologia, a OACI utiliza o termo "inglês comum para comunicação aeronáutica" (*plain English*). O excerto abaixo, retirado de Emery (2014, p. 200) exemplifica o uso do inglês comum nesse contexto:

> *Pilot: Control. Redline 253. We have a passenger on board with suspected heart attack, we'd like to divert to the nearest available airfield. Request full medical assistance on arrival.*

[10] No original: "the spontaneous, creative and noncoded use of a given natural language, although constrained by the functions and topics (aviation and non-aviation) that are required by aeronautical radiotelephony communications, as well as by specific safety-critical requirements for intelligibility, directness, appropriacy, non-ambiguity and concision".

É importante salientar que o inglês comum para comunicação aeronáutica não deve ser confundido com o inglês geral, pelo fato de o primeiro estar restrito às especificidades do contexto em que as comunicações aeronáuticas ocorrem. Diversos estudos têm se dedicado ao uso da língua inglesa por pilotos e controladores na aviação civil internacional (ALDERSON, 2011; ESTIVAL; FARRIS, 2016; EMERY, 2014). No Brasil, o Grupo de Estudos em Inglês Aeronáutico (GEIA) conduz pesquisas sobre aspectos relacionados à descrição da linguagem, ao ensino e aprendizagem e à avaliação no âmbito do inglês aeronáutico (SCARAMUCCI, 2011b; CHINI, 2014; BONIFÁCIO, 2015; SILVA, 2016; PRADO; TOSQUI-LUCKS, 2017; SCARAMUCCI et al, 2018). O termo em português "inglês aeronáutico" é utilizado por esse grupo para se referir à linguagem utilizada nas comunicações aeronáuticas entre pilotos, controladores e operadores de estação aeronáutica, composta pelo uso da fraseologia e do inglês comum, fazendo contraponto com o termo "inglês para aviação" (*Aviation English*). O "inglês para aviação" apresenta escopo mais amplo, caraterizado pelo uso da língua por diversos profissionais no campo da aviação, tais como engenheiros aeronáuticos, mecânicos de aeronaves, técnicos, entre outros (OACI, 2010; SILVA, 2016). O inglês aeronáutico é, portanto, considerado uma subcategoria do inglês para aviação (OACI, 2010, p. 3-2).

Os requisitos de proficiência linguística da OACI se restringem ao uso do inglês comum para comunicação aeronáutica, em situações inesperadas, incomuns ou emergenciais, para as quais não há fraseologia prevista. O emprego da fraseologia para a OACI é considerado uma habilidade operacional, que deve ser avaliada separadamente do inglês comum, por pessoal qualificado operacionalmente. Isso, porém, não impede que a fraseologia seja utilizada nos exames para estabelecer um cenário ou para dar início a uma tarefa.

Para atender aos requisitos, o Doc 9835 AN/453 cita alguns parâmetros que devem ser seguidos pelos exames de proficiência no contexto aeronáutico: a) avaliar as habilidades de compreensão e produção orais; b) promover o uso do inglês comum para comunicação aeronáutica em situações inesperadas, incomuns e emergenciais; c) priorizar a inteligibilidade em um ambiente intercultural; d) abordar conteúdo pertinente às comunicações e operações aeronáuticas; e e) propor tarefas que se assemelhem às atividades profissionais. Somado a isso, o documento pontua que exames de desempenho seriam o método mais apropriado de avaliar essa proficiência.

Em relação aos efeitos exercidos por exames dessa natureza, a OACI explicita, no Doc 9835 AN/453, que os elaboradores devem demonstrar que seus instrumentos de avaliação promovem práticas voltadas para o desenvolvimento da competência comunicativa em detrimento do ensino de técnicas de memorização e preparação. A organização também enfatiza

ser imprescindível o desenvolvimento de programas de capacitação em inglês aeronáutico bem fundamentados e alinhados ao construto da avaliação na comunicação aeronáutica para que pilotos, controladores e operadores alcancem o nível de proficiência necessário na condução das operações de voo internacionais.

Bachman e Palmer (2010) argumentam que a intenção dos elaboradores de um teste, bem como dos usuários de seus resultados, em geral, é contribuir para que consequências benéficas sejam geradas tanto para os indivíduos e os programas educacionais quanto para a sociedade. No entanto, não podemos desconsiderar que o efeito retroativo é um fenômeno intermediado por diversos fatores, o que faz com que efeitos não pretendidos ou indesejáveis também possam ocorrer. Nesse sentido, estudos tornam-se imprescindíveis para investigar se o teste está produzindo efeitos positivos conforme esperado, da mesma forma que os aspectos negativos necessitam ser identificados, para que possam ser controlados e minimizados.

Conhecer as características que, idealmente, estarão presentes em sala de aula se os exames exercerem os efeitos a que se propõem é o primeiro passo em uma investigação sobre efeito retroativo (WALL; HORAK, 2007). O levantamento das intenções dos elaboradores do exame e sua comparação com as características presentes no teste e as consequências reais nos contextos de ensino e aprendizagem têm sido uma abordagem bastante utilizada em estudos sobre efeito retroativo, principalmente, naqueles que não possuem dados de base (*baseline data*) (SCARAMUCCI, 2004; QI, 2004; TSAGARI, 2009; SCARAMUCCI, 2011a).

Embora, o ICEA não especifique quais seriam os efeitos pretendidos de seu exame, pelo fato de ter sido elaborado seguindo as recomendações da OACI, podemos assumir que são os mesmos daqueles encontrados no Doc 9835 AN/453, especialmente no Capítulo 7 desse manual, dedicado a orientar os estados membros em questões relacionadas à capacitação de seus profissionais, e na Circular 323 AN/185 (OACI, 2009), publicada em 2009, na qual estão estabelecidos os princípios de boa prática e orientações para planejamento e desenvolvimento de programas de capacitação. Com base na análise desses documentos, foi possível depreender alguns efeitos pretendidos que devem ser almejados pelos elaboradores de exames de inglês aeronáutico na visão da OACI:

a) criação de programas de ensino de inglês focado nas necessidades dos profissionais envolvidos na comunicação aeronáutica;
b) adoção da abordagem de ensino comunicativa, com ênfase no desenvolvimento das habilidades orais (*listening* e *speaking*) de forma integrada;

c) foco no uso do inglês comum para comunicação aeronáutica no gerenciamento de situações inesperadas, incomuns e emergenciais, passíveis de ocorrer nas comunicações aeronáuticas;
d) ênfase nas funções linguísticas necessárias para o gerenciamento e solução dessas situações;
e) tópicos e situações semelhantes aos encontrados no ambiente profissional;
f) elaboração de atividades que promovam a interação em sala de aula, criando oportunidades para o desenvolvimento da produção oral; e
g) desenvolvimento das estratégias comunicativas para lidar com possíveis mal-entendidos.

Podemos dizer que existem vários pontos de convergência entre as características do EPLIS e os efeitos descritos acima, o que pode colaborar para que haja efeitos positivos nos contextos de ensino e aprendizagem. Ou seja, o exame se propõe avaliar as habilidades de compreensão e produção orais na solução de situações incomuns, inesperadas e/ou emergenciais. Os temas, tópicos e funções abordados no teste são próprios do contexto aeronáutico e se baseiam no inventário encontrado no Apêndice B do Doc 9835 AN/453. As tarefas propostas, por sua vez, buscam se assemelhar àquelas encontradas no ambiente profissional. E, como salienta Scaramucci et al (2010), pode ser considerado um exame de desempenho, com propósitos específicos, uma vez que o conhecimento técnico, necessário para o desempenho da profissão de controlador de tráfego aéreo, interage com o conhecimento linguístico exigido em uma situação comunicacional. A estrutura do EPLIS, seu formato, tipos de atividades e exemplos de tarefas serão apresentados no Capítulo III.

1.3 O problema de pesquisa e a justificativa

Nos anos de 2008 e 2009, atuei como professora de inglês na Escola de Especialistas de Aeronáutica (EEAR), na cidade de Guaratinguetá. A EEAR é uma organização militar do Comando da Aeronáutica, responsável pela formação e o aperfeiçoamento dos sargentos especialistas de aeronáutica. São oferecidos mais de 23 cursos, dentre os quais está o Curso de Formação de Sargentos em Controle de Tráfego Aéreo. Fui designada para a EEAR para lecionar inglês a esse grupo específico de alunos – controladores de tráfego aéreo em formação.

Logo que cheguei à escola, deparei-me com uma reestruturação do currículo das disciplinas de língua inglesa do curso de Controle de Tráfego Aéreo. As mudanças abrangeram diversos aspectos, tais como implementação da abordagem de ensino de inglês para fins específicos (ESP) em substituição ao ensino de inglês geral, adoção de um material didático específico para pilotos e controladores, ênfase no desenvolvimento das habilidades de compreensão e produção orais, busca dos professores para compreender melhor a atividade profissional que seus alunos iriam desempenhar; entre outros.

Intrigada para saber o que estava impulsionando tantas mudanças, conversei com os professores mais antigos que me apresentaram um documento, no qual estavam descritas as medidas de capacitação a serem tomadas visando à elevação do nível de inglês dos controladores de tráfego aéreo brasileiros. O documento do qual eles falavam era o *Plano de Implementação dos Requisitos de Proficiência Linguística em Inglês* – PCA 37-9, referente ao período 2008-2010, publicado pelo DECEA pela primeira vez em 2008 e reeditado em 2011 e 2014 (BRASIL, 2008, 2011, 2014).

A existência desse plano e a cobrança das autoridades para que as medidas fossem implementadas provocaram, imediatamente, sem sombra de dúvida, uma comoção interna entre os professores e coordenadores de língua inglesa na EEAR, que se viam pressionados por uma política linguística sobre a qual não puderam opinar. Houve, no início, bastante resistência e questionamento por parte de alguns, principalmente, em relação à abordagem de ensino de línguas para fins específicos. Esse grupo argumentava que não se sentia preparado e seguro para ministrar aulas voltadas às necessidades linguísticas do controlador de tráfego aéreo, pois não tinha conhecimento suficiente sobre a área de atuação desse profissional.

Além da descrição das ações de capacitação, esse plano versava também sobre a criação do EPLIS, como parte das exigências feitas pela OACI em relação à necessidade de testagem periódica da proficiência linguística de controladores e operadores de estação aeronáutica. No entanto, até aquele momento, não havia nenhuma menção a respeito da obrigatoriedade de se aplicar o EPLIS aos alunos dos cursos de formação. Devido a esse distanciamento, somado ao fato de ser um exame ainda recente, acredito que o efeito retroativo do EPLIS na EEAR era bastante limitado nessa época. Em minhas aulas, por exemplo, por diversas vezes, chamei a atenção dos alunos para a importância da língua inglesa no controle de tráfego aéreo e mencionei que eles teriam que realizar um exame que atestaria o seu nível de inglês sem me aprofundar sobre essas questões. Na verdade, meu conhecimento sobre o teste era relativamente pouco e não havia muita informação disponível a nós, professores.

No final de 2009, deixei a EEAR para fazer parte da, então, Subdivisão de Língua Inglesa do ICEA, responsável pelo planejamento, desenvolvimento e aplicação do EPLIS. Nesse mesmo ano, o EPLIS foi submetido a um processo de validação, liderado pela professora Dra. Matilde Virginia Ricardi Scaramucci, da Universidade Estadual de Campinas, com o objetivo de coletar e analisar evidências sobre sua validade. Uma das deficiências apontadas por Scaramucci et al (2010) foi, exatamente, a falta de maior divulgação do EPLIS tanto para a comunidade aeronáutica como também para a sociedade. Existiam informações sobre a estruturação do exame e modelos de questões disponíveis em um endereço eletrônico, porém o acesso era restrito aos candidatos. Essa dificuldade em se obter informações mais específicas sobre o teste coincide com a minha experiência como professora na EEAR. Ademais, diversos estudos sobre impactos reconhecem que a carência de informações pode afetar o potencial de um exame em exercer efeitos positivos (HUGHES, 1989, 1993; ALDERSON; HAMP-LYONS, 1996; GREEN, 2007a).

Todavia, mesmo com pouco tempo de implementação e com uma divulgação deficiente, Scaramucci et al (2010) e Azevedo (2009) perceberam algumas evidências de efeito retroativo do EPLIS em entrevistas com professores, controladores, elaboradores do exame e candidatos: a) alunos ingressantes no curso de formação de controladores estavam mais conscientes sobre a necessidade de uma maior proficiência em inglês; b) os cursos de formação passaram a dar maior ênfase às habilidades orais e ao conteúdo aeronáutico; c) houve aumento de materiais didáticos voltados para o inglês aeronáutico e d) tornou-se obrigatório frequentar cursos de inglês, mesmo contrário à vontade de alguns controladores de tráfego aéreo em serviço.

Em 2011, para solucionar o problema da divulgação, a equipe de elaboradores do EPLIS criou uma página na web dedicada exclusivamente ao exame e aberta ao público, elaborou um manual ao candidato com informações sobre o teste, suas características, os critérios de atribuição de notas e exemplos de tarefas, e disponibilizou vídeos ilustrativos da fase II do exame, juntamente com um exemplar completo da prova.

Outra medida implementada foi a introdução do EPLIS versão *ab-initio* nos cursos de formação ocorrida no final de 2014. Essa versão do exame foi elaborada exclusivamente para avaliar a proficiência de alunos dos cursos de Controle de Tráfego Aéreo e Operador de Estação Aeronáutica e tem sido aplicada, desde então, aos alunos da última série da EEAR. Ressalta-se que, antes de 2014, esses alunos só se submetiam ao EPLIS quando já estavam atuando nos órgãos de controle para os quais foram designados, ou seja, seis a doze meses após a conclusão do curso, com raríssimas exceções. Embora não tenha sido divulgada oficialmente

a razão que levou as autoridades a implementarem o EPLIS na EEAR, acredita-se que essa decisão tenha sido motivada pelo interesse em aumentar o grau de relevância do EPLIS na escola, na expectativa de que houvesse aumento no nível de proficiência desses alunos. Frente a essa mudança, foi oferecida, em 2015, uma oficina sobre o EPLIS, com dois dias de duração, aos professores da EEAR, na qual foram discutidos diversos aspectos referentes ao exame, tais como, sua estrutura, objetivo e construto, a tabela de níveis de proficiência do OACI, os critérios avaliados e o perfil de desempenho desejado.

Apesar dos impactos potenciais que os exames de alta-relevância exercem, como é o caso do EPLIS, no contexto de ensino e aprendizagem, as pesquisas sobre o assunto ainda são escassas no cenário nacional e carecem de mais evidências empíricas, tornando necessária a coleta de dados sobre efeito retroativo, principalmente, nas percepções e atitudes de alunos e professores, considerados os principais envolvidos nesse processo (SCARAMUCCI et al, 2010; SCARAMUCCI, 2011b).

Sabemos que o efeito retroativo é um fenômeno complexo mediado por fatores que vão além do teste em si e sua natureza. Forças diversas interagem com as características do exame para determinarem o seu efeito, tais como as crenças e experiências dos professores e alunos, as exigências das instituições em que estão inseridos, além de questões políticas, sociais e econômicas. Scaramucci et al (2010, p. 209) explica que "as características vistas como positivas e inovadoras [de um exame] podem ter impactos negativos dependendo de como as pessoas as interpretam". Isso quer dizer que um teste ruim pode trazer efeitos positivos, assim como um bom teste pode estar associado a efeitos negativos (MESSICK, 1996).

A minha experiência como professora de inglês no contexto investigado, bem como as informações levantadas por Scaramucci et al (2010) durante o processo de validação do EPLIS, apontam para a existência de efeito retroativo do exame em diferentes aspectos do ensino e aprendizagem de inglês, que pode ter sido potencializado com a entrada do exame na EEAR e com as melhorias implementadas em sua divulgação. Uma investigação sobre o assunto faz-se urgente e necessária e, como parte da equipe de desenvolvimento do EPLIS, tenho um compromisso profissional de averiguar se o exame está exercendo os impactos desejáveis, se existem efeitos não pretendidos ou até mesmo negativos. Porém, devido à minha relação com a escola, meu compromisso torna-se também pessoal e afetivo para com a formação desses alunos. Por essas razões, decidi analisar o impacto do exame nas percepções e atitudes de professores e alunos da EEAR, a fim de contribuir para a melhoria do exame e para a qualidade do ensino e aprendizagem de inglês nesse contexto.

1.4 As perguntas de pesquisa

Compartilho da visão de que exames, especialmente aqueles considerados de alta-relevância, são capazes de induzir mudanças no sistema educacional e de que esse impacto pode sofrer influência de outros fatores como, por exemplo, as interpretações que o professor e os alunos fazem do teste, baseadas em suas crenças, valores, conhecimentos e experiência. Assim sendo, esta pesquisa pretende investigar o efeito retroativo do EPLIS no ensino e na aprendizagem de inglês no Curso de Formação de Controladores de Tráfego Aéreo, mais especificamente, nas percepções, atitudes e ações de professores e alunos em sala de aula. Neste estudo, as "percepções" são entendidas como as interpretações que os participantes desenvolvem sobre o EPLIS. A "atitude", por sua vez, é compreendida como a disposição para responder favoravelmente ou não ao exame, ou seja, é uma disposição para uma possível ação e, de certa forma, dependente da percepção. A ação é, por sua vez, caracterizada pelos eventos que ocorrem em sala de aula envolvendo professores e alunos (CHENG, 2005).

Esta pesquisa pretende responder a três perguntas:

a) Qual é o efeito retroativo do EPLIS nas percepções e atitudes dos professores de língua inglesa do Curso de Controle de Tráfego Aéreo?
b) Qual é o efeito retroativo do EPLIS nas percepções e atitudes dos alunos do curso?
c) Qual é o efeito retroativo do EPLIS nas ações desses professores e alunos em sala de aula?

Essas perguntas permitem compreender como o efeito retroativo atua nas percepções, atitudes e ações de professores e alunos, quais são os fatores mais relevantes na sua determinação e como esses se relacionam. O efeito retroativo é, portanto, entendido como um fenômeno complexo, mediado por diversos fatores, e que envolve efeitos pretendidos, não pretendidos, indesejáveis, positivos ou negativos.

1.5 As contribuições do estudo

Mesmo com o aumento de pesquisas sobre efeito retroativo nas últimas décadas, tanto no contexto internacional quanto no contexto brasileiro, parece ainda não existir consenso sobre os fatores que interferem na sua determinação, o que dificulta a construção de um modelo

que atenda às especificidades dos exames e dos contextos em que estão inseridos. Scaramucci (2004), assim como Cheng (2008) e Tsagari (2009), mencionam a necessidade de mais estudos sobre o fenômeno que sejam longitudinais, contemplando múltiplas fases, abrangendo contextos e exames específicos, com, pelo menos, dois grupos diferentes de participantes e emprego de diversos métodos e instrumentos de coleta e análise de dados. A pesquisa ora proposta mostra-se relevante para a área de Linguística Aplicada e para a avaliação de línguas em três diferentes aspectos:

i) *Contribuições teóricas e metodológicas*. Esta pesquisa pretende não apenas explorar os fatores que interagem com o impacto do teste nas percepções, atitudes e ações de professores e alunos, mas também identificar as relações existentes entre eles. Para isso, serão empregados métodos qualitativos e quantitativos de geração de dados. Esses, por sua vez, serão analisados sob diferentes perspectivas, a fim de identificar padrões e tendências entre os participantes e compreender melhor como se apresenta a influência do exame no cenário investigado. Acreditamos que os resultados desta pesquisa irão colaborar para a construção de uma teoria mais satisfatória sobre o efeito retroativo, dado que os modelos propostos ainda não exauriram por completo sua complexidade. Isso porque as especificidades inerentes a cada exame interagem com as características do contexto em que esses estão inseridos, bem como dos participantes envolvidos, ressaltando a natureza multifacetada e situada do fenômeno. Outra contribuição importante refere-se ao fato de que, embora os professores sejam amplamente pesquisados pelo papel chave que esses profissionais exercem nesse processo, o mesmo não ocorre em relação aos alunos. Segundo Cheng et al (2011), pouco ainda se sabe sobre a influência que os exames exercem em diversos aspectos concernentes a esses participantes e como eles se preparam. Esta pesquisa, por ter também como foco de investigação o aluno, ajuda a suprir a necessidade de estudos nessa área.

ii) *Contribuições para a elaboração e implementação de políticas*. No contexto de ensino de inglês para fins específicos, mais especificamente em relação ao ensino de inglês aeronáutico no Brasil, não existem ainda investigações sistemáticas sobre o efeito retroativo do EPLIS, mesmo sendo esse um exame de alta-relevância. Scaramucci et al (2010), em seu relatório final sobre o processo de validação do EPLIS, aponta para a necessidade de estudos dessa natureza, a fim de investigar se

a política de implementação dos requisitos de proficiência linguística da OACI e o exame estão produzindo efeitos pretendidos em prol da elevação do nível de proficiência dos controladores e operadores de estação aeronáutica brasileiros. Os resultados desta pesquisa serão levados ao conhecimento das autoridades brasileiras no controle de tráfego aéreo no sentido de assessorar sobre a necessidade de reformulação da política linguística adotada, se for o caso, e de provimento de melhorias tanto para o ensino e aprendizagem de inglês aeronáutico quanto para o exame em si.

iii) *Contribuições pedagógicas*. Está prevista a realização de entrevistas individuais com os professores participantes após cada bloco de observações de aulas. Essas entrevistas têm como objetivo o aumento da confiabilidade dos dados coletados, uma vez que os professores têm a possibilidade de expor sua visão sobre os eventos ocorridos em sala de aula. Mas, funcionam também como espaços colaborativos para promoção do amadurecimento da reflexão crítica dos professores envolvidos. Liberali (2004) e Oliveira e Souza (2010) argumentam que, por meio da reflexão crítica e teoricamente informada, o professor assume um papel mais consciente e ativo sobre suas ações pedagógicas, tornando-se capaz de identificar problemas e de traçar estratégias para resolvê-los ou diminui-los. É nesse sentido que esta pesquisa colabora para o desenvolvimento profissional dos participantes e estreita a relação entre a pesquisa e a docência.

1.6 A estrutura da tese

Este trabalho está organizado em seis capítulos. Neste primeiro capítulo, contextualizei o leitor a respeito do macro contexto em que este estudo está inserido, apresentando os requisitos de proficiência linguística da OACI, as características dos exames no meio aeronáutico e os efeitos pretendidos dessa política. Em seguida, justifiquei a necessidade de pesquisas sobre efeito retroativo nesse contexto e expus minha motivação pessoal e profissional na escolha do tópico. Passei, então, a tratar das perguntas de pesquisa que guiarão a coleta e a análise dos dados e demonstrei como este estudo pode contribuir na construção de uma teoria mais satisfatória sobre o fenômeno, na formulação e reformulação de políticas linguísticas mais informadas e no desenvolvimento profissional dos professores participantes.

No Capítulo II, faço uma revisão crítica da literatura existente. Primeiramente, discuto questões terminológicas envolvendo o efeito retroativo, a complexidade do fenômeno, sua natureza, seu escopo e a diversidade de fatores que interferem na sua determinação. Em seguida, apresento as contribuições, bem como as limitações, das teorias e modelos propostos para explicar o fenômeno. Relaciono, então, os principais estudos sobre efeito retroativo no programa e no aluno e, por fim, discorro sobre os tipos de preparação nos quais professores e alunos podem se engajar, influenciados pelo exame.

No Capítulo III, discorro sobre os paradigmas metodológicos nos estudos sobre efeito retroativo e apresento a pesquisa de métodos mistos como a escolha mais apropriada para lidar com os propósitos desta investigação. Esta pesquisa é composta por duas fases sequenciais, sendo a primeira de natureza quantitativa, e a segunda, qualitativa. Para a geração de dados, foram utilizados questionários, comentários escritos, entrevistas, grupo focal e observações de aulas.

No Capítulo IV, trato da fase quantitativa do estudo, cujos dados foram coletados por meio de questionários. Utilizei diversos métodos estatísticos para a análise dos dados, tais como a estatística descritiva, a análise de componentes principais e testes de hipóteses, com o objetivo de traçar o perfil dos participantes, identificar o grau de influência do exame em suas percepções e atitudes e constatar padrões e tendências entre eles.

O Capítulo V apresenta a análise qualitativa dos dados que teve como objetivo expandir, explicar, refinar e corroborar os resultados estatísticos alcançados na fase anterior. Os dados coletados nos comentários escritos, grupo focal, entrevistas individuais e observações de aulas foram submetidos à análise de conteúdo de Bardin (1977), utilizando como categorias de análise os componentes extraídos na fase quantitativa.

O Capítulo VI é dedicado a responder as perguntas de pesquisa, fazendo uma discussão dos resultados nas análises quantitativa e qualitativa. Apresento, então, as conclusões do estudo, as contribuições em termos teóricos, metodológicos e pedagógicos, assim como as limitações e dificuldades encontradas. Por fim, traço algumas sugestões para pesquisas futuras.

CAPÍTULO II

OS ESTUDOS SOBRE EFEITO RETROATIVO

2.1 Considerações iniciais

Este capítulo tem como objetivo discutir o conceito de efeito retroativo e seu mecanismo de ação (circunstâncias, forças interagentes e maneiras de medi-lo). Para tanto, discorro, primeiramente, sobre o poder dos testes[11] no direcionamento de políticas educacionais e apresento o conceito de efeito retroativo. Passo a tratar da natureza do fenômeno e sua relação com os exames de desempenho. Em seguida, discuto algumas propostas de teorias elaboradas com o objetivo de explicar seu mecanismo e direcionar pesquisas sobre o assunto. Apresento, então, o papel do professor na sua determinação e discuto sobre a necessidade de mais estudos que foquem nos alunos. Por fim, traço um paralelo entre o efeito retroativo e a validade de um exame e argumento que a responsabilidade de investigar esses efeitos deve recair sobre todos os envolvidos – elaboradores do teste e usuários.

2.2 O efeito retroativo

Poucos discordariam de que vivemos em um mundo cada vez mais regulado por testes externos. O Exame do Ensino Médio (ENEM), por exemplo, seleciona candidatos para dezenas de universidades públicas federais no país. Nas carreiras do funcionalismo público, concursos são realizados periodicamente e contam com a inscrição de milhares de pessoas. Até mesmo para exercer uma profissão, em alguns casos, o diploma de graduação não é suficiente, fazendo-se necessária uma aprovação em um exame da especialidade, como é o caso da Ordem dos Advogados do Brasil.

Na área da educação, práticas avaliativas externas são capazes de nortear ações nos contextos de ensino e aprendizagem que as precedem. Nas palavras de Shohamy (2006, p. 103, tradução nossa), "os exames são capazes de impor o que será estudado, aprendido e ensinado

[11] "Teste" e "exame" são utilizados, neste estudo, como sinônimos para definir procedimentos formais de coleta de informações. No entanto, em inglês, os termos *testing* e *assessment* caracterizam paradigmas de avaliação bastante distintos. Enquanto *testing* evidencia uma visão estruturalista e positivista de avaliação, *assessment* descreve a avaliação como prática social, permeada de valores, o que condiz com as mudanças conceptuais ocorridas nas últimas décadas na Linguística Aplicada, como explica Scaramucci (2011a). Em português, ambos os termos são, comumente, traduzidos como avaliação.

por professores e alunos",[12] principalmente, e mais facilmente, se a relevância atribuída a eles for considerada alta. Smith (1991, p. 524, tradução nossa) define exames de alta-relevância como "aqueles que acarretam consequências ou desencadeiam ações (por exemplo, contribuem para seleção, avaliação por mérito, promoção, retenção, ou tomada de decisão)"[13], com graves consequências para os indivíduos, instituições e sociedade.

Em contextos de alta-relevância, Frederiksen e Collins (1989) explicam que as notas nos testes se tornam uma moeda de troca capaz de transformar todo um sistema educacional. Quanto maior a relevância atribuída a um teste, maior o seu potencial de induzir mudanças. Em suas palavras:

> o sistema ajusta suas práticas curriculares e instrucionais, e os alunos ajustam suas estratégias de aprendizagem e objetivos, para maximizar as notas nos exames utilizados para avaliar os resultados educacionais[14] (FREDERIKSEN; COLLINS, 1989, p. 27, tradução nossa).

A relação existente entre a avaliação e o ensino e aprendizagem é conhecida como efeito retroativo (em inglês, *backwash*, *washback effect* ou simplesmente *washback*) e tem sido objeto de estudo de vários pesquisadores na Educação e na Linguística Aplicada (ALDERSON; WALL, 1993; MESSICK, 1996; BAILEY, 1996; CHENG, 1999, 2004; WATANABE, 2004a, 2004b; SCARAMUCCI, 1999, 2004; GIMENEZ, 1999; RETORTA, 2007; CHALHOUB-DEVILLE, 2009). Os termos em inglês *backwash* ou *washback effect* são considerados sinônimos. No entanto, os linguistas aplicados preferem empregar o termo *washback* (ALDERSON; WALL, 1993).

Scaramucci (2004) aponta que, em dicionários não técnicos da língua inglesa, somente o termo *backwash* é encontrado para definir uma condição resultante de uma ação tomada previamente, porém como algo geralmente prejudicial e indesejável. Em português, o fenômeno tem sido traduzido como "efeito retroativo", enfatizando o sentido de retroação (aquele que retroage, que modifica o que está feito). Quanto à conotação negativa dada ao termo em inglês, Scaramucci (2004) não observou essa situação em língua portuguesa.

Alderson e Wall (1993, p. 117, tradução nossa), em seu artigo seminal sobre a existência do efeito retroativo, definem o fenômeno como sendo o grau com que o teste

---

[12] No original: "Tests are capable of dictating what will be studied, learned and taught by students and teachers".
[13] No original: "High stakes tests are those that have consequences or trigger actions (e.g. contribute to selection, merit evaluation, promotion, retention, or take-over decision)".
[14] No original: "The system adjusts its curricular and instructional practices, and students adjust their learning strategies and goals, to maximize the scores on the tests used to evaluate educational outcomes".

influencia "professores e alunos de línguas a fazerem coisas que eles não necessariamente fariam se não fosse pela existência do teste"[15] e sugerem que esses efeitos fiquem restritos à sala de aula.

Assim como Alderson e Wall (1993), Messick (1996), Bachman e Palmer (1996) e McNamara (2010) compreendem o efeito retroativo como algo restrito à sala de aula, em um nível micro de atuação, e propõem uma diferenciação entre "efeito retroativo" e "impacto". O "impacto" é descrito por eles como os efeitos de exames em um contexto educacional mais amplo, ou seja, na sociedade e nas políticas educacionais. Nessa perspectiva, o "efeito retroativo" está inserido dentro do escopo do impacto, limitando-se à influência do teste nos processos de ensino e aprendizagem, nos indivíduos envolvidos e nas atividades pedagógicas.

Porém, essa distinção não é compartilhada por todos os pesquisadores da área. Em diversos estudos, efeito retroativo é usado como sinônimo de impacto, englobando tanto um foco mais restrito (professores e alunos em sala de aula) quanto um foco mais amplo (o sistema educacional e a sociedade como um todo) (SCARAMUCCI, 1999, 2004; ANDREWS et al, 2002; ANDREWS, 2004; CHENG, 2004; WATANABE, 2004a). Nesse caso, "efeito retroativo" é utilizado como um termo técnico, enquanto "impacto" é considerado um termo geral e não técnico, assim como a palavra "influência".

Outros termos relacionados à influência exercida por testes são encontrados na literatura em Educação. Instrução guiada por medidas (*measurement-driven instruction*) refere-se à noção de que os exames devem direcionar o ensino. Alinhamento curricular (*test-curriculum alignment*) enfatiza a influência de exames na (re)elaboração do currículo escolar. Frederiksen e Collins (1989) defendem a noção de validade sistêmica (*systemic validity*) na qual um teste é considerado válido somente se provocar efeitos positivos. Messick (1989), por sua vez, coloca o efeito retroativo como uma das facetas da validade consequencial de exames (*consequential validity*).

Esta pesquisa utiliza efeito retroativo e impacto como termos intercambiáveis que descrevem a influência de exames nos processos de ensino e aprendizagem, em um nível micro de atuação, com potencial para afetar as percepções e atitudes daqueles envolvidos, o currículo escolar, os métodos de ensino, o conteúdo abordado, os materiais didáticos e as avaliações, bem como todo o sistema educacional e a sociedade. Essa também tem sido a decisão de outros estudiosos do tema, dentre eles, Cheng (1999), Bailey (1999), Scaramucci (1999, 2004) e Andrews (2004). Embora este estudo tenha como foco de investigação o nível micro do efeito

[15] No original: "teachers and learners do things they would not necessarily otherwise do because of the test".

retroativo, seus resultados trazem implicações também para o nível macro, principalmente, em relação à (re)formulação de políticas linguísticas.

2.3 A natureza do efeito retroativo e os exames de desempenho

Até os anos de 1990, o efeito retroativo era entendido como tendo um caráter determinista, ou seja, acreditava-se que um teste ruim resultaria em efeitos negativos enquanto um bom teste traria efeitos benéficos. Alderson e Wall (1993) argumentam que essa visão é ingênua, diante da multiplicidade de fatores que podem influenciar o ensino e aprendizagem. Como bem coloca Messick (1996, p. 242, tradução nossa), "um teste ruim pode estar associado a efeitos positivos e um bom teste com efeitos negativos devido a outras coisas que são feitas ou não no sistema educacional".[16] Os estudos sobre efeito retroativo de Alderson e Wall (1993), Cheng (1999), Scaramucci (1999) e Watanabe (2004b) são alguns exemplos de muitos que confirmaram a complexidade do fenômeno, cujo processo é de natureza multifacetada e mediado por variáveis que podem interferir de forma positiva ou negativa no impacto do exame.

Por muito tempo, a adoção de exames externos como medida para estabelecer reformas educacionais esteve associada a algo prejudicial ao ensino e aprendizagem, devido à gama de efeitos desastrosos que eles poderiam acarretar. Os estudos conduzidos por Noble e Smith (1994) e Smith (1991) sobre instrução guiada por medidas (*measurement-driven instruction*), por exemplo, apontam que os professores, pressionados para elevar a nota de seus alunos nesses exames, tendiam a moldar a sua prática pedagógica para aproximá-la do teste. Entre os aspectos negativos, estavam: o *estreitamento do currículo* que reduz a capacidade do professor de trabalhar conteúdos e tópicos que não estão no escopo do teste; a *redução dos métodos de instrução* que faz com que o professor perca sua habilidade de adaptar, criar e diversificar os modos como trabalha em sala de aula; e *a diminuição do tempo destinado à instrução* para ensinar técnicas de preparação para o teste.

Em relação aos alunos, Prodomou (1995, p. 19, tradução nossa) atesta que "a avaliação é uma camisa de força que restringe estilos de aprendizagem individuais e desencoraja o potencial de crescimento de alunos mais fracos".[17] Os testes a que Prodomou se refere são aqueles de natureza estruturalista, nos quais o foco reside na manipulação e

[16] No original: "[…] a poor test may be associated with positive effects and a good test with negative effects because of other things that are done or not done in the educational system".
[17] No original: "testing is a straitjacket that cramps personal learning styles and discourages the 'weak' learner's potential".

memorização de regras gramaticais e vocabulário. Prodomou (1995) distingue, então, dois tipos de efeito retroativo no ensino e aprendizagem: o efeito retroativo implícito (*covert washback*) e o efeito retroativo explícito (*overt washback*). A influência explícita é percebida no tratamento direto do exame em sala de aula: utilização de materiais didáticos focados no teste e uso de provas anteriores ou itens similares. O caráter implícito refere-se ao impacto do teste nas visões de professores e alunos sobre ensinar e aprender. Considerado sutil e, muitas vezes, inconsciente, esse processo afeta a sala de aula de várias formas: o erro passa a ser penalizado, alunos são desencorajados a fazerem perguntas, o tempo para a realização de um exercício torna-se cronometrado e limitado, a interação entre alunos é minimizada, o professor não oferece qualquer tipo de explicação para as respostas dos exercícios, o livro didático preparatório torna-se o direcionador do ensino.

Fredericksen e Collins (1989, p. 29, tradução nossa) explicam que, de fato, testes objetivos como, por exemplo, os testes de múltipla escolha, exercem efeitos negativos, como aqueles citados por Smith (1991), Noble e Smith (1994) e Prodomou (1995), por enfatizarem "habilidades de baixo nível, conhecimento factual, memorização de procedimentos e habilidades isoladas".[18] No entanto, afirmam que testes diretos, conhecidos hoje como testes de desempenho (*performance tests*), cujas tarefas são representativas das situações de uso da língua no mundo real, podem promover efeitos benéficos. Segundo Fredericksen e Collins (1989, p. 27, tradução nossa), exames de natureza comunicativa, autênticos e diretos possuem a capacidade de "induzir, no sistema educacional, mudanças curriculares e instrucionais que promovam o desenvolvimento de habilidades cognitivas que esses testes pretendem avaliar".[19] Assim, a adoção de tais testes levaria a um aumento da validade sistêmica[20] que, por sua vez, influenciaria o desenvolvimento de habilidades e competências de alto nível (*high level skills*) e a capacidade de solução de problemas reais.

Fredericksen e Collins (1989) não foram os primeiros e únicos a reconhecerem a capacidade existente em certos testes de promover efeitos positivos. Em 1988, Pearson (1988, p. 101, tradução nossa) já havia argumentado a favor do uso de exames como "alavancas para mudanças" (*levers for change*). No mesmo ano, Alderson (1988) fez o seguinte comentário:

> Ao invés de vermos os testes negativamente como uma restrição conservadora ao desenvolvimento, nós poderíamos reconhecer o fato de que os testes exercem

[18] No original: "Objective tests emphasize low-level skills, factual knowledge, memorization of procedures, and isolated skills".
[19] No original: "A systemically valid test is one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure".
[20] No conceito de validade sistêmica, um teste é considerado válido somente se produzir efeitos benéficos.

> poderosa influência nos alunos e na aprendizagem. Essa influência será negativa se o teste for ruim, mas pode ser positiva se o teste representar um objetivo que valha a pena para nossos alunos. Em outras palavras, os testes possuem potencial considerável para a inovação[21] (ALDERSON, 1988, p. 87, tradução nossa).

Embora essa fala de Alderson (1988) soe determinista quando diz que um teste ruim causaria, necessariamente, consequências negativas, o autor reconhece que a elaboração de bons instrumentos de avaliação é um meio de promover mudanças educacionais e se coloca ciente de que essa relação não é direta e de que outros fatores, além do exame em si, podem afetar esse processo.

Para que um exame seja capaz de influenciar positivamente o ensino e a aprendizagem, Fredericksen e Collins (1989) apresentam quatro padrões de validade a serem seguidos durante o processo de desenvolvimento: a) os exames precisam avaliar os conhecimentos e habilidades de forma direta; b) as tarefas presentes no exame devem cobrir, ao máximo, todos os conhecimentos, habilidades e estratégias necessários no desempenho da atividade no mundo real; c) a atribuição de notas deve ser confiável e utilizar critérios para avaliar os subprocessos das tarefas independentemente; e d) os critérios de avaliação devem ser claros e estar disponíveis para todos os envolvidos nos processos – examinadores, professores, alunos, entre outros.

Assim como Fredericksen e Collins (1989), Messick (1996) destaca duas características essenciais que devem ser respeitadas na elaboração de exames a fim de minimizar os efeitos negativos e otimizar o impacto positivo: ser direto e autêntico, ou seja, ter como base amostras autênticas e diretas do uso comunicativo da língua. De acordo com Messick (1996, p. 241, tradução nossa), "deve haver pouca ou nenhuma diferença entre as atividades para a aprendizagem da língua e as atividades envolvidas na preparação para o teste".[22] Dessa forma, as ameaças à validade a que os exames estão sujeitos seriam contidas, resultando no incremento de seu potencial de provocar efeitos positivos, o que Messick (1996) denominou *washback by design*. Scaramucci (2011a, p. 115) explica o conceito de *washback by design* como a busca pela "validade no projeto de elaboração do teste como base para alcançarmos esses efeitos [efeitos positivos]". Na seção 2.8, discorro melhor sobre a visão de Messick a

[21] No original: "Instead of viewing the tests negatively as a conservative restraint on development, we could acknowledge the fact that tests have a powerful influence on learners and learning. This influence will be negative if the test is bad, but it can be positive if the test represents a worthwhile goal for our students. In other words, tests have considerable potential for innovation".

[22] No original: "There should be little if any difference between activities involved in learning the language and activities involved in preparing for the test".

respeito da relação entre efeito retroativo e validade, fazendo um contraponto com conceito de validade sistemática proposto por Fredericksen e Collins (1989).

A autenticidade garante que nada relevante para o construto seja deixado de fora do processo de avaliação, minimizando o risco de sub-representação do construto (*construct under representation*). Os testes diretos, por sua vez, permitem que nada irrelevante seja acrescentado, evitando, assim, interferência e contaminação do construto a ser avaliado, a chamada "variância irrelevante ao construto" (*construct-irrelevant variance*). Sabemos, entretanto, que essas ameaças sempre estarão presentes em qualquer contexto avaliativo, uma vez que um exame tende a suscitar sentimentos de ansiedade, por exemplo, além de não ser incapaz de englobar todos os processos envolvidos na situação alvo. Porém, como pontua Messick (1996), é necessário reduzi-las e mantê-las sob controle para que efeitos benéficos sejam assim gerados.

Mcnamara (1996) também acredita na capacidade dos testes comunicativos, baseados em desempenho, de promover efeitos positivos no ensino, uma vez que os professores passariam a preparar os alunos para as tarefas comunicativas presentes nos exames. Scaramucci (2011a, p. 109) reforça essa postura ao argumentar que para promover melhorias no ensino e na aprendizagem "elaborar um bom exame, apesar de não ser condição suficiente, é uma condição necessária".

2.4 O escopo e a variabilidade do efeito retroativo

Nas últimas três décadas, diversos estudiosos se dedicaram a entender a relação existente entre a avaliação e o ensino e aprendizagem. No entanto, as primeiras pesquisas, além de compreenderem o efeito retroativo como um fenômeno determinista, careciam de dados empíricos provenientes de observações em sala de aula. A publicação do artigo de Alderson e Wall (1993) marcou uma mudança significativa nos estudos sobre o fenômeno. Sua existência e complexidade foram reconhecidas, e as pesquisas passaram a investigar seu escopo, sua natureza, as circunstâncias em que atua, os fatores intermediadores e como poderia ser medido. Dados provenientes de observações de aulas, juntamente com questionários e entrevistas com participantes, embasaram empiricamente a construção de teorias e modelos.

Alderson e Wall (1993), considerados os precursores na proposta de modelos que pudessem orientar os estudos sobre efeito retroativo, postularam quinze hipóteses, apresentadas no Quadro 1, para descrever os aspectos do ensino e aprendizagem que são, geralmente, afetados por exames. Em resumo, o efeito retroativo poderia abranger o conteúdo, a metodologia, o ritmo, a sequência, o grau e a profundidade do que será ensinado e aprendido,

sendo também capaz de influenciar as atitudes daqueles envolvidos nesses processos – professores e alunos. Apesar de terem sido bastante elucidativas em relação às áreas afetadas pelo efeito retroativo, as hipóteses de Alderson e Wall (1993) não exploraram, em detalhes, os fatores externos que interferem na determinação do fenômeno.

Quadro 1 − Hipóteses de efeito retroativo

1.	Um teste influencia o ensino.
2.	Um teste influencia a aprendizagem.
3.	Um teste influencia o que os professores ensinam.
4.	Um teste influencia como os professores ensinam.
5.	Um teste influencia o que os alunos aprendem.
6.	Um teste influencia como os alunos aprendem.
7.	Um teste influencia o ritmo e a sequência do ensino.
8.	Um teste influencia o ritmo e a sequência da aprendizagem.
9.	Um teste influencia o grau e a profundidade do ensino.
10.	Um teste influencia o grau e a profundidade da aprendizagem.
11.	Um teste influencia as atitudes para com o conteúdo e o método de ensinar e aprender.
12.	Testes que têm consequências importantes terão efeito retroativo.
13.	Testes que não têm consequências importantes não terão efeito retroativo.
14.	Testes terão efeito retroativo para todos os alunos e professores.
15.	Testes terão efeito retroativo para alguns alunos e alguns professores, mas não para outros.

Fonte: ALDERSON; WALL, 1993, p. 120-121 (tradução nossa)

Três anos mais tarde, em um estudo sobre o efeito retroativo do TOEFL (*Test of English as a Foreign Language*) nos conteúdos e métodos instrucionais em cursos preparatórios, conduzido em conjunto com Hamp-Lyons, Alderson propôs a reformulação da hipótese de número 15 para que fosse incluído o fato de que a quantidade e o tipo de efeito retroativo variam de professor para professor e de aluno para aluno, dependendo do *status* atribuído ao exame (o nível de relevância). Além da relevância atribuída ao teste, Alderson e Hamp-Lyons (1996) acrescentaram três elementos intermediadores nesse processo: a) até que ponto o exame é contrário às práticas em vigor; b) a opinião de professores e autores de livros didáticos sobre os métodos mais apropriados na preparação; e c) a disposição e capacidade de professores e autores de livros didáticos de inovar.

Hughes (1993), longe de discordar dos estudos de Alderson e Wall (1993) sobre efeito retroativo, argumentou que não deveríamos nos preocupar tanto em distinguir os tipos de efeitos e os aspectos do ensino e aprendizagem afetados. Para ele, os resultados da aprendizagem, influenciados pelo exame, devem ser nosso principal foco de interesse. Hughes (1993) ofereceu, então, um modelo tripartite no qual *participantes*, *processos* e *produtos* são passíveis de serem afetados pelo teste. Nesse modelo, o exame afeta primeiramente as

percepções e atitudes dos participantes (*stakeholders*)[23] envolvidos no processo de ensino e aprendizagem. A lista de participantes, comparada à proposta de Alderson e Wall (1993), é estendida para contemplar administradores educacionais, elaboradores de materiais didáticos, desenhistas de cursos, além de professores e alunos. As ações tomadas pelos participantes, influenciados pelo exame, são definidas como processos e englobam o desenvolvimento de materiais didáticos, a elaboração de currículos, mudanças na metodologia, usos de estratégias de aprendizagem, o emprego de técnicas para passar no exame, dentre outros. Os resultados da aprendizagem, ou seja, o que é aprendido e a qualidade dessa aprendizagem, são denominados produtos.

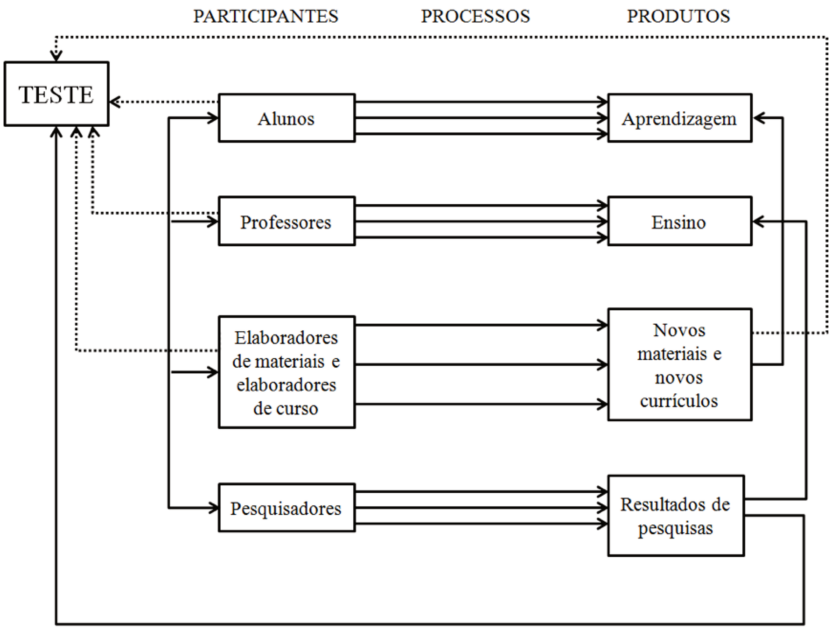
Hughes (1993, p. 2) também estabeleceu cinco condições necessárias para que possíveis efeitos positivos pudessem ocorrer: 1) é importante para os alunos serem bem-sucedidos no exame; 2) os professores devem desejar que seus alunos tenham um bom desempenho; 3) os participantes devem estar familiarizados com o exame e compreender as implicações de sua natureza e conteúdo para o ensino e aprendizagem ; 4) os participantes devem possuir ou serem capazes de obter o conhecimento cobrado pelo exame; e 5) os recursos necessários para a preparação devem estar disponíveis. Assim, interferência em qualquer uma dessas condições poderia reduzir a capacidade de um teste em exercer os efeitos esperados.

O modelo de efeito retroativo tripartite de Hughes (1993) foi expandido por Bailey (1996) que, para explicar seu mecanismo de seu funcionamento, destacou as múltiplas inter-relações entre participantes, processos e produtos (Figura 1). Nessa proposta, as setas contínuas representam a influência direta do teste nos participantes que, por sua vez, irão engajar-se em processos para a criação de novos produtos. As linhas pontilhadas representam o caminho inverso que pode ocorrer quando os participantes influenciam de alguma forma os exames.

A fim de sistematizar os estudos sobre efeito retroativo, Bailey (1996, 1999) também propôs distinguir o *efeito retroativo no aluno* (*learner washback*) do *efeito retroativo no programa* (*programme washback*), ou seja, nos professores, administradores educacionais, elaboradores de curso e materiais didáticos, com base nas hipóteses de efeito retroativo de Alderson e Wall (1993). Dessa forma, as hipóteses de Alderson e Wall (1993) foram classificadas entre as que estariam relacionadas diretamente ao efeito retroativo no aluno (2, 5, 6, 8 e 10) e aquelas com referência direta ao efeito retroativo no programa (1, 3, 4, 7, 9 e 11).

[23] O termo *stakeholders* pode ser traduzido também por "protagonistas", "atores" ou "partes interessadas". Hoje, esse termo abarca todos aqueles envolvidos no desenvolvimento e aplicação do exame, incluindo os usuários dos resultados do teste. Entre eles estão, elaboradores de exames, professores, educadores, elaboradores de material didático, alunos, candidatos, administradores e órgãos governamentais.

Figura 1 − Modelo básico de efeito retroativo



Fonte: BAILEY, 1996, p. 264 (tradução nossa)

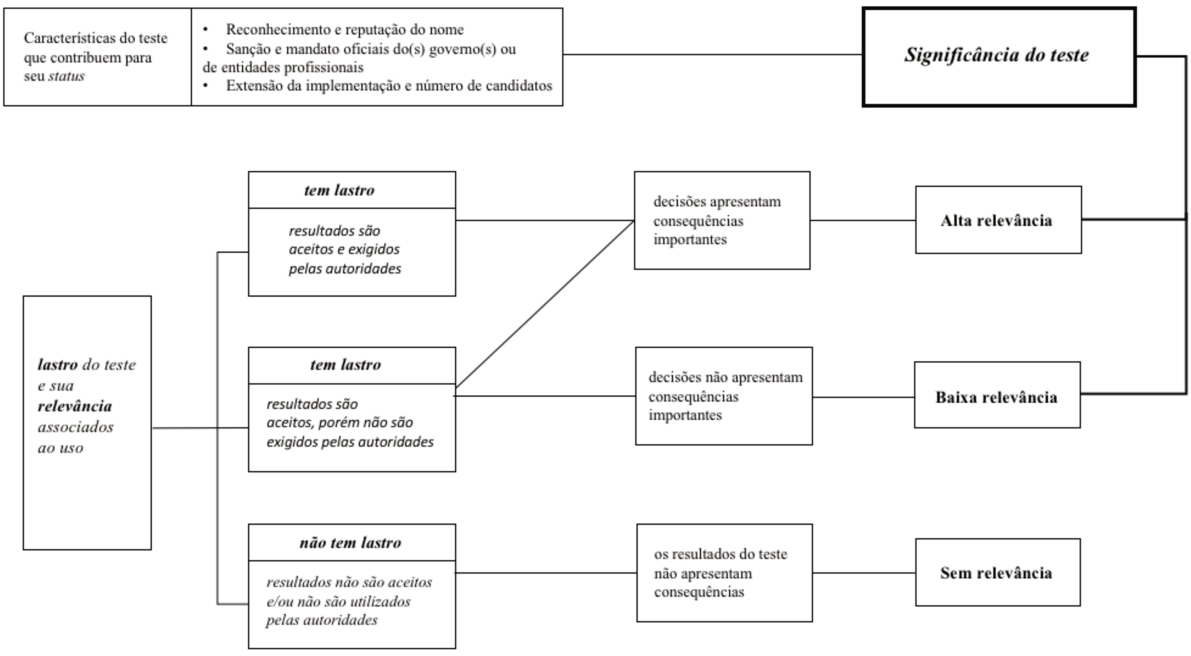
Watanabe (2004a) fez uso da distinção feita por Bailey (1996) para diferenciar o efeito retroativo no aluno − o que os alunos aprendem, como eles aprendem, o ritmo, a sequência, o grau e a profundidade da aprendizagem, do efeito retroativo no programa que afeta o conteúdo, a metodologia e a abordagem adotada pelo professor em sala de aula. Em relação aos fatores que intermediam o efeito retroativo de um exame, Watanabe (2004a) propôs organizá-los da seguinte forma: *fatores ligados ao teste* (métodos, conteúdo, habilidades, propostas e decisões a partir dos resultados), *fatores de prestígio* (relevância e *status* do teste dentro do sistema educacional), *fatores pessoais* (formação do professor, suas crenças, entre outros), e *fatores contextuais* como, por exemplo, as exigências das escolas em relação ao exame e à preparação e a pressão das autoridades por resultados mais satisfatórios.

Watanabe (2004a) apresentou, então, uma metodologia de estudo da complexidade do efeito retroativo, na qual os fatores elencados acima atuam na determinação desse efeito em cinco dimensões: especificidade, intensidade, extensão, intencionalidade e valor. Em relação à *especificidade*, o efeito retroativo pode ser geral ou específico. Por "geral", entende-se o impacto exercido por exames independente de suas características ou conteúdo, ao passo que o "específico" se refere às influências do teste no conteúdo e nos tipos de atividades realizadas em sala de aula. A *intensidade* refere-se a um *spectrum*, que vai do fraco ao forte. Se ela é forte, o exame determinará tudo que irá acontecer em sala de aula, levando o professor a trabalhar nos moldes do que é exigido pelo teste. Se a intensidade é fraca, o exame poderá afetar alguns

conteúdos e outros não, alguns professores e não outros, alguns alunos e não outros. A *extensão* diz respeito à duração desse impacto, podendo ser curta ou longa. A *intencionalidade* está ligada ao fato de o efeito retroativo do teste ser algo pretendido, ou seja, previsto intencionalmente por seus elaboradores, ou não intencional, quando o impacto é considerado inesperado. O *valor*, por sua vez, será positivo ou negativo e depende do julgamento dos diversos envolvidos no processo − um mesmo efeito pode ser considerado para alguns como algo positivo e para outros como negativo.

Os fatores de prestígio e sua relação com o efeito retroativo foram explorados por Stoneman (2006), que argumentou a favor da necessidade de diferenciar os termos *status* e *relevância,* utilizados por Alderson e Hamp-Lyons (1996). Stoneman (2006) alegou que Alderson e Hamp-Lyons (1996) não deixaram claro se os termos *status* e relevância foram empregados como sinônimos ou se seriam dois atributos relacionados entre si. Para Stoneman (2006), *status* e relevância descrevem condições ligadas à significância de um exame, porém não podem ser tratados como equivalentes, uma vez que seus papéis na mediação do efeito retroativo não são idênticos. Na tentativa de explicar como esses dois atributos intermediam a significância de um teste, Stoneman (2006, p. 93) ofereceu um novo modelo (Figura 2).

Figura 2 − Modelo de significância atribuída a um teste



Fonte: STONEMAN, 2006, p. 93 (tradução nossa)

Nessa proposta, três componentes são apresentados como contribuintes para o *status* de um teste: 1) o grau de reconhecimento e reputação atribuído ao exame por seus usuários; 2) a sanção oficial ou chancela dada ao exame por parte do governo ou de entidades profissionais; e 3) o escopo de sua implementação. O exame, no entanto, não precisa possuir todas essas características para apresentar *status*. O nível de relevância (*test stakes*), por sua vez, estaria intimamente ligado à importância que o indivíduo atribui às consequências provocadas por seus resultados (*test consequences*). Segundo Stoneman (2006), um teste cujos resultados são aceitos e utilizados por aqueles envolvidos na tomada de decisão apresentará lastro (*test currency*), porém sua relevância não será necessariamente alta, dependendo, portanto, de como os indivíduos interpretam as consequências atreladas a esses resultados.

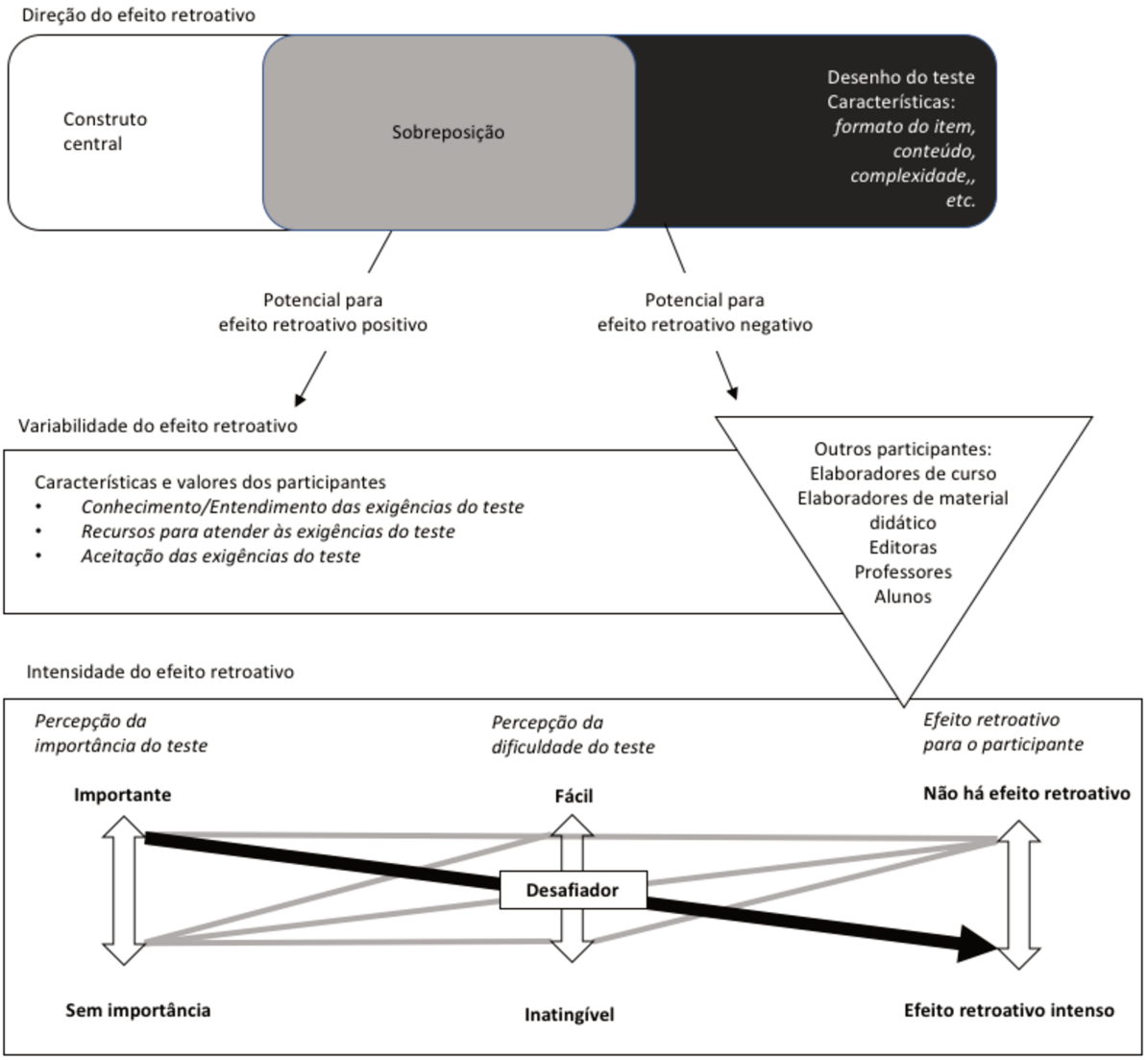
Em resumo, o *status* de um teste é medido pelo grau de aceitação e uso de seus resultados, denominado lastro (*test currency*). A relevância (*stakes*) refere-se à percepção do indivíduo das consequências dos resultados do teste. Nesse sentido, enquanto o *status* é considerado um atributo mais estável e concreto, a relevância é tida como um construto subjetivo, podendo ser interpretada diferentemente pelos participantes. Ou seja, "o que é considerado uma consequência trivial para um participante pode ser vista como fundamental por outro"[24] (ALDERSON, 2004, p. x, tradução nossa).

Embora o estudo de Stoneman (2006) tenha contribuído consideravelmente para aumentar o entendimento acerca dos diferentes elementos que compõem a significância de um teste, ainda pouco se sabe sobre os aspectos envolvidos na percepção do participante que fazem com que ele atribua alta ou baixa relevância a um exame.

Green (2007a), por sua vez, optou por um modelo unificado (Figura 3), no qual a direção, a variabilidade e a intensidade do efeito retroativo são integradas. A direção está relacionada ao quanto o exame reflete o construto que se pretende medir. Como discutido anteriormente, ser autêntico e direto são características que aumentam as chances de os exames produzirem efeitos benéficos ao ensino e aprendizagem. Nesse modelo, podemos observar a área de sobreposição entre o construto e as características do teste (itens, conteúdo, habilidades). Quanto maior a área de sobreposição, maior será o potencial de o teste exercer efeitos positivos e vice-versa. A direção do efeito retroativo está, portanto, primeiramente, associada aos fatores do teste, descritos por Watanabe (2004a).

[24] No original: "[…] what is a trivial consequence for one person may be an important matter for another".

Figura 3 − Modelo de efeito retroativo



Fonte: GREEN, 2007a, p. 24 (tradução nossa)

Sabemos, porém, que as características do exame, apesar de serem necessárias, não são suficientes para promover inovações (MESSICK, 1996; SCARAMUCCI, 2011a). O efeito retroativo é um fenômeno complexo, mediado por outras variáveis que vão além do exame em si. Sua variabilidade é determinada pelas características e valores dos participantes envolvidos no processo, os chamados fatores pessoais por Watanabe (2004a), e esses devem estar representados em uma teoria sobre o assunto. Green (2007a) destacou, em seu modelo, três variáveis, citadas por Hughes (1993) e Alderson e Hamp-Lyons (1996): o nível de entendimento sobre as exigências do exame, o nível de aceitação dessas demandas e os recursos disponíveis para atendê-las.

A noção de intensidade do efeito retroativo, citada por Watanabe (2004a) e desenvolvida por Cheng (2005) para explicar o grau com que os participantes ajustam suas ações às exigências do exame, também está presente na proposta de Green (2007a). A percepção dos participantes sobre a significância do teste irá determinar o grau de envolvimento em atividades preparatórias. Sendo assim, quanto maior for a significância (*status*, aceitação, consequências e relevância) atribuída ao exame por seus usuários, maior será a intensidade de sua influência. Outro fator contemplado no modelo está ligado à dificuldade do exame. Se os participantes percebem o exame como fácil demais, eles não dispensarão qualquer atenção especial a ele. As demandas do exame também serão ignoradas se elas foram entendidas como intangíveis por professores e alunos.

Contudo, embora Watanabe (1996) e Smith (1991) tenham mostrado que os métodos de ensino podem mudar drasticamente quando professores e alunos se veem próximos à realização do exame, sugerindo que a proximidade do exame desempenha papel importante na intensidade do efeito retroativo, bem como as crenças, valores e conhecimento do professor, esse fator não está contemplado no modelo de Green (2007a), tampouco nos modelos anteriores. Os fatores contextuais, ressaltados por Watanabe (2004a), também não são explicitamente abordados por Green, embora ele tenha afirmado o quão dependente o efeito retroativo se apresenta do contexto em que está inserido.

Portanto, ainda que os modelos apresentados mostrem uma evolução no entendimento do efeito retroativo e nos auxiliem a ter uma visão mais clara sobre seu mecanismo de funcionamento, percebemos que nenhum deles é capaz de, isoladamente, englobar toda a complexidade do fenômeno. Isso porque é bastante difícil anteciparmos todas as variáveis que irão interagir com o efeito retroativo, uma vez que elas estão intrinsicamente ligadas ao contexto no qual o exame está inserido. No entanto, esses modelos oferecem pontos de partida para a investigação. Parto, então, das suposições de que um teste é capaz de influenciar as percepções e atitudes dos participantes, fazendo com que eles se engajem em novas práticas pedagógicas. As características do exame, permeadas pelos valores, crenças e conhecimento dos participantes, podem afetar diversos aspectos do ensino e aprendizagem. Sabemos, também, que outras variáveis, tais como a importância do teste, a percepção do seu nível de dificuldade, o nível de entendimento sobre suas exigências, sua proximidade e os recursos disponíveis, atuam na intensidade do efeito retroativo.

A seguir, apresento uma seleção de estudos sobre efeito retroativo no cenário internacional e nacional, cujos resultados contribuíram para um melhor entendimento do mecanismo de funcionamento do fenômeno. Para fins de organização, divido os estudos com

base na distinção feita por Bailey (1996, 1999) e Watanabe (2004a) entre efeito retroativo nos programas (no currículo, no conteúdo, nos materiais didáticos, nos métodos de instrução e nas percepções e atitudes do professor) e efeito retroativo nos alunos, com especial interesse nos estudos nos quais professores e alunos foram o foco principal da investigação.

2.5 Estudos sobre efeito retroativo nos programas

O efeito retroativo nos programas, com ênfase nas percepções, atitudes e ações de professores, tem sido extensivamente investigado (BAILEY, 1996; WATANABE, 2004a). Esse fato não causa estranhamento dado o papel significativo do professor no processo de efeito retroativo. Bailey (1996, p. 17, tradução nossa) descreve os professores como "os participantes mais visíveis, a linha de frente na condução dos processos de efeito retroativo relacionados à instrução".[25] Os estudos apresentados a seguir exibem como foco de investigação a influência de exames de alta-relevância no professor em contextos bastante distintos de ensino e aprendizagem.

Noble e Smith (1994), em uma pesquisa sobre mudanças no sistema escolar por meio de um exame obrigatório de leitura, escrita e matemática, implantado no estado do Arizona, aferiram a variabilidade no comportamento dos professores provocada pela dissonância entre suas crenças, sua capacidade de inovação e o exame. O estudo contou com análise documental, entrevistas com os elaboradores da política e professores, além de observações em sala de aula. Eles concluíram que, apesar de o teste ter sido criado com base em pressupostos cognitivo-construtivistas, os resultados desejáveis não foram alcançados em sua plenitude. Algumas razões foram apontadas, tais como a ausência de iniciativas com foco no desenvolvimento profissional dos professores dentro desse programa. Segundo Noble e Smith (1994, p. 17), mudanças conceituais são difíceis de serem alcançadas se não forem levadas em consideração as crenças daqueles que são o alvo (professores) e as condições do meio no qual estão inseridos (escolas).[26]

Cheng (1999) também demonstrou que mudanças em um exame de alta-relevância, com a intenção de promover inovações curriculares, não foram capazes de modificar completamente a prática pedagógica dos professores investigados. Em seu estudo, Cheng

[25] No original: "Of course, the most visible participants in program washback are language teachers. It is they who are "the front-line conduits for the washback processes related to instruction".
[26] No original: "Conceptual change is seldom achieved without attending to the beliefs of those who are the targets (teachers) and the conditions of the environments in which they function (schools)" (tradução nossa).

(1999) analisou o efeito retroativo de um exame de inglês, considerado de alta-relevância para os alunos da educação secundária em Hong Kong, nas atitudes e comportamentos de professores. Esse exame havia passado por reestruturações, em 1993, com o objetivo de alavancar melhorias no sistema educacional. As mudanças concentraram-se na integração das habilidades de compreensão oral, leitura e escrita para a solução de tarefas da vida real e na introdução de atividades de simulações e discussão em grupo em substituição à leitura em voz alta e conversa guiada. Esperava-se, então, que o novo exame transformasse um ensino pouco interativo e com foco no professor em um ensino de inglês mais comunicativo e com foco no aluno.

Foram conduzidas entrevistas e observadas as aulas de três professoras em dois momentos: antes da implementação do novo exame e após a consolidação do exame, quando as novas diretrizes curriculares passaram a ser válidas. Algumas mudanças foram identificadas pela pesquisadora que sugerem influência da nova avaliação, tais como o aumento nas atividades de simulação, mais oportunidades de prática oral com discussões em grupos nas aulas das três professoras, aumento do tempo de fala do aluno e diminuição do tempo de fala do professor nas aulas de duas professoras. Entretanto, o padrão com que as atividades eram conduzidas ainda permanecia bastante controlado pelos professores, iniciando com a explicação do professor, seguida de atividades práticas para os alunos e finalizando com comentários do professor. Cheng (1999) concluiu que, apesar das mudanças nos tipos de atividades desenvolvidas em sala de aula, as crenças e atitudes das professoras sobre ensinar e aprender não foram alteradas.

No contexto brasileiro, o estudo de Scaramucci (1999) investigou o impacto da prova de inglês do vestibular da Universidade Estadual de Campinas (Unicamp), considerada de alta-relevância, na prática pedagógica de uma professora de escola pública de nível médio. A prova de inglês da Unicamp, a partir de 1987, começou a abordar a leitura como um processo ativo de construção de significados, contrária à visão tradicional de que a leitura é um processo passivo de decodificação de palavras. O objetivo dessa proposta era selecionar alunos com capacidade de "organizar suas ideias e exprimir-se com clareza, estabelecer relações, interpretar dados e elaborar hipóteses" (SCARAMUCCI, 1999, p. 9), além de promover inovações no ensino médio brasileiro. Ao contrário de Cheng (1999), com um período mais extenso entre a modificação do exame e a coleta de dados, esperava-se que o efeito retroativo fosse considerável.

Todavia, os dados obtidos por meio entrevistas, questionários e observações em sala de aula revelaram baixa intensidade do efeito retroativo. Scaramucci (1999) apontou que,

embora a professora tenha afirmado possuir conhecimentos sobre a proposta da prova de inglês da Unicamp e utilizar provas anteriores em sala de aula, o que ocorria, na prática, eram aulas ainda bastante fundamentadas em uma visão estruturalista de leitura, com excessos de exercícios de tradução, foco em gramática e pouca ou nenhuma variedade de gêneros discursivos, temas e textos autênticos.

Scaramucci (1999) concluiu que o impacto da proposta da prova de inglês da Unicamp ficava restrito ao conteúdo trabalhado em sala de aula, ou seja, a leitura. Não foram observados impactos na abordagem empregada pela professora para o ensino dessa habilidade. Esses resultados corroboram aqueles obtidos por Cheng (1999). No entanto, além das crenças e da formação do professor, Scaramucci (1999) acrescentou outros fatores que podem ter limitado o efeito retroativo: material didático pouco atraente, sobrecarga de trabalho do professor, desinteresse dos alunos, ausência de uma proposta de ensino de inglês pela escola e inexistência de uma coordenação pedagógica efetiva.

Assim como Cheng (1999) e Scaramucci (1999), Qi (2004) também verificou limitações no efeito retroativo de um exame de alta-relevância para àqueles que anseiam entrar em uma universidade chinesa e que havia sido reformulado na intenção de promover melhorias pedagógicas na educação secundária. As modificações no exame abrangeram a avaliação das habilidades de leitura, a compreensão oral e a escrita, inexistente na versão anterior, e a diminuição no peso destinado a questões de gramática e vocabulário. A intenção era que essas mudanças fomentassem o uso da língua com propósitos reais de comunicação.

Participaram da pesquisa oito elaboradores do exame, dez professores de inglês e três inspetores educacionais. Os dados foram coletados por meio de entrevistas conduzidas com esses participantes e os resultados revelaram que o efeito retroativo do exame ainda era bastante limitado. O ensino descontextualizado de aspectos linguísticos da língua, gramática e vocabulário continuava desempenhando papel crucial, enquanto as habilidades de leitura, compreensão oral e escrita eram praticadas, individualmente, como dever de casa, diferentemente do que os elaboradores do exame pretendiam. Além disso, grande parte das atividades com foco no desenvolvimento das habilidades eram elaboradas no formato múltipla escolha, o que fez supor que não apenas o conteúdo do teste, mas também o seu formato, exercia impacto no ensino e aprendizagem.

O estudo de Qi (2004), comparado aos estudos de Cheng (1999) e Scaramucci (1999), é limitado por não contemplar dados relativos a observações diretas em sala de aula. Mesmo tendo incluído vários grupos de participantes em sua pesquisa, as afirmações feita por Qi (2004) não puderam ser confrontadas com o que ocorria na prática.

Retorta (2007), por sua vez, além de coletar dados de diferentes participantes (professores, alunos, diretores, coordenadores, autores de materiais didáticos adotados pela escola, autoridades e elaboradores do exame), também realizou observações em sala de aula. A pesquisadora investigou o impacto da prova de inglês do vestibular da Universidade Federal do Paraná no ensino e aprendizagem de escolas de ensino médio em Curitiba. Em 1995, a prova de inglês da universidade paranaense havia passado a avaliar a leitura, por meio de textos autênticos, em substituição à avaliação de conhecimento gramatical. Assim como na pesquisa de Scaramucci (1999), observamos aqui a existência de um período de tempo considerável entre a modificação do exame e a pesquisa, o que favorece, a princípio, a ocorrência de efeito retroativo. Participaram da pesquisa de Retorta (2007) sete escolas de nível médio, sendo duas públicas, duas particulares e três cursos preparatórios, com atendimento a diferentes classes sociais.

Os resultados mostraram que, em relação às escolas públicas, não havia efeito retroativo do exame. O currículo era guiado pelo material didático adotado pela escola. Esse, por sua vez, era pautado em uma visão tradicional de ensino voltada para gramática e tradução. Nas entrevistas com os diretores e professores, percebeu-se que o exame vestibular tinha baixa ou nenhuma relevância nessas escolas por ser considerado uma meta inatingível. Os diretores, professores e alunos mostravam-se resignados e conformados com essa situação.

Nas escolas particulares e nos cursos preparatórios, foi observado efeito retroativo positivo do exame no conteúdo (leitura), na escolha do material didático, na motivação de professores e alunos, porém em níveis diferentes de intensidade. Contudo, efeitos negativos também foram constatados nesses dois contextos, tais como resquícios de um ensino descontextualizo e focado na gramática e no vocabulário. Esses efeitos podem estar relacionados a problemas estruturais da prova. Retorta (2007), ao analisar amostras de provas, descobriu que, embora o teste avaliasse a leitura, era privilegiado o modelo de leitura ascendente, com foco na reconstituição da informação e com itens no formato múltipla escolha, o que impedia a construção de sentidos na interpretação do texto.

A pesquisadora concluiu que o fator classe social relacionado à percepção de dificuldade do exame pareceram mediar o efeito retroativo. Para ela, o efeito retroativo ocorria com mais intensidade e mais positivamente em contextos privilegiados, afetando a busca de informações sobre o exame, a motivação e o interesse da escola e dos professores em preparar seus alunos e a adoção de materiais didáticos mais adequados.

Gimenez (1999) e Watanabe (2004b), por sua vez, analisaram o efeito retroativo negativo em contextos de ensino e aprendizagem de exames de língua inglesa com ênfase no

conhecimento de gramática e vocabulário. O estudo de Gimenez (1999) teve como foco o impacto da prova de inglês do exame vestibular da Universidade Estadual de Londrina no ensino de inglês de escolas de segundo grau e de um curso preparatório e sua relação com a proposta curricular do estado.

Primeiramente, Gimenez (1999) investigou se o exame de inglês de entrada na universidade londrinense estava condizente com proposta curricular do estado do Paraná, modificada em 1988, para o ensino de língua inglesa no segundo grau. A proposta curricular defendia um ensino voltado para a dimensão sociocultural da comunicação e para o uso efetivo da língua, com ênfase na leitura. A visão de leitura pautava-se na construção de sentidos, com uso de textos autênticos que promovessem a participação do aluno em sala de aula. A prova de inglês da universidade alinhava-se a essa proposta parcialmente, uma vez que também abordava o conhecimento de gramática e de vocabulário de forma isolada. Nesse exame, 50% dos itens envolviam questões gramaticais em frases curtas e descontextualizadas. Os demais 50% compreendiam a avaliação da leitura com ênfase apenas na retomada de informações.

Nas observações de aulas de uma professora em um curso preparatório, Gimenez (1999) constatou predomínio de atividades de gramática e tradução ditadas, geralmente, pelo material didático, e pouco ou nenhum envolvimento com a leitura. O papel do professor limitava-se a fornecer explicações gramaticais e traduções de palavras-chave, ensinar macetes para acertar as questões da prova, incitar a repetição de termos pelos alunos e esclarecer dúvidas.

Os resultados indicaram que havia impactos da prova de inglês da universidade londrinense no ensino de inglês no curso preparatório, porém esse efeito restringia-se a algumas tarefas da prova com foco na análise gramatical. A compreensão de leitura, que compreendia 50% do exame, não era trabalhada pela professora que, por sua vez, enfatizava atividades não presentes no teste como, por exemplo, a repetição. A pesquisadora concluiu seu estudo afirmando que algumas partes do exame terão efeito retroativo e outras não, e que o efeito retroativo é dependente das crenças dos professores e de suas percepções sobre as chances de sucesso de seus alunos.

Watanabe (2004b) investigou o impacto dos exames de entrada às universidades japonesas em cursos preparatórios, a fim de verificar se as alegações sobre seu efeito maléfico nas atividades de ensino eram pertinentes. Seis alegações guiaram o estudo de Watanabe: a) referência constante aos exames; b) predomínio de métodos gramática-tradução; c) explicações sobre estruturas gramaticais em detrimento do uso da língua para propósitos comunicativos; d) pouca ênfase nas habilidades orais; e) pouca participação dos alunos; e f) modelos limitados de

interação. Foram observadas aulas em três escolas de ensino médio que ofereciam cursos preparatórios para os exames e eram conhecidas por aprovar um número expressivo de alunos em universidades de prestígio. No total, cinco professores foram entrevistados, todos ministravam aulas em turmas regulares e preparatórias.

Os resultados apontaram que as alegações eram muito simplistas face à complexidade da realidade da sala de aula. Observou-se que, realmente, as aulas de alguns professores eram influenciadas negativamente pelos exames. Porém, alguns deles foram influenciados positivamente ao promoverem o desenvolvimento de habilidades comunicativas. Watanabe (2004b) concluiu que a familiaridade do professor com diferentes métodos de ensino, o nível de culpa atribuída pelo professor ao exame e o apoio da escola para o desenvolvimento da competência comunicativa dos alunos pareciam mediar o tipo de efeito retroativo em sala de aula. Para Watanabe, as inovações não dependem exclusivamente de mudanças no sistema de avaliação. Se não houver um trabalho de pesquisa-ação que motive a reflexão crítica com os professores, mudanças não serão efetivamente alcançadas.

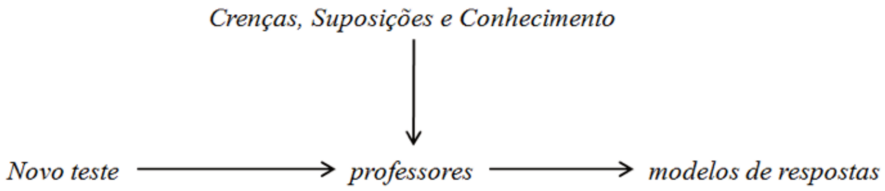
Em suma, os estudos de Noble e Smith (1994), Scaramucci (1999), Cheng (1999), Qi (2004 Retorta (2007), Gimenez (1999) e Watanabe (2004b) analisaram o efeito retroativo de exames de alta-relevância nos sistemas educacionais nos quais estavam inseridos. Em relação à capacidade de exames em promoverem inovações, os resultados revelaram que os efeitos pretendidos não foram alcançados conforme esperado. Mudanças no conteúdo abordado em sala de aula foram verificadas, porém a abordagem e os métodos de ensino permaneciam inalterados. Em relação aos exames considerados problemáticos, efeitos retroativos negativos foram, de fato, constatados, porém efeitos positivos também estavam presentes.

Em todos esses estudos, um nível alto de variação foi encontrado nas ações dos professores observados. Diversos fatores relacionados aos professores foram apontados como possíveis causas, tais como crenças, valores, capacitação, motivação, contexto institucional. O modelo abaixo (Figura 4), proposto por Burrows (2001, 2004), ilustra a intermediação feita por esses participantes na determinação do efeito retroativo e reconhece que eles irão agir diferentemente ao impacto de um teste, influenciados por suas crenças, suposições e conhecimento.

Burrows (2001, 2004) também acredita que, embora os professores sejam afetados de diferentes maneiras pelo exame, suas respostas podem revelar certos padrões de comportamento. Esses modelos de resposta são importantes de serem identificados, pois colaboram para uma teoria geral mais satisfatória sobre o efeito retroativo. Em seu estudo, Burrows (2001) identificou três diferentes tipos de comportamento entre os professores: os

resistentes (*the resisters*), cujas crenças, suposições e conhecimento conflitam com os propósitos do exame; aqueles que adotam totalmente as mudanças por estarem de acordo com suas bases teóricas e educacionais (*the adopters*); e aqueles que se adaptam (*the adapters*) ou que decidem adotar parcialmente as mudanças (*the partial adopters*), ou seja, apesar de não estarem em total consonância com o que é proposto eles não o rejeitam por completo.

Figura 4 − Modelo de inovação curricular



Fonte: BURROWS, 2004, p. 126 (tradução nossa)

As crenças, suposições e conhecimentos dos professores funcionam, portanto, como um filtro. Qualquer conflito com a natureza e as exigências do teste podem dificultar a ocorrência dos efeitos pretendidos. Alderson (2004), ao afirmar que tanto o professor quanto o teste são responsáveis pelo efeito retroativo, justifica a necessidade de mais estudos que investiguem a fundo o comportamento desses participantes. Em suas palavras:

> muito mais atenção deve ser dada às razões pelas quais os professores ensinam do modo como ensinam. Nós precisamos entender suas crenças sobre ensino e aprendizagem, seu grau de profissionalismo, a adequação de sua capacitação e o seu entendimento sobre a natureza e os fundamentos do exame[27] (ALDERSON, 2004, p. xi, tradução nossa).

Borg (2003, p. 81) explica que a cognição do professor (*teacher cognition*), termo cunhado por ele para se referir a "o quê os professores sabem, acreditam e pensam",[28] é moldada pela sua experiência como aprendiz, sua formação profissional, sua prática de sala de aula e fatores contextuais. Nesse sentido, a forma como o professor lida com um exame é fruto de suas experiências em avaliar e ser avaliado, dos conhecimentos adquiridos por ele sobre essa área durante seu processo de formação e das restrições impostas pelo contexto em que atua.

---

[27] No original: "Much more attention needs to be paid to the reasons why teachers teach the way they do. We need to understand their beliefs about teaching and learning, the degree of their professionalism, the adequacy of their training and their understanding of the nature of and rationale for the test".

[28] No original: "[…] what teachers know, believe and think".

Para Scaramucci (2016), o grau de formação do professor na área de avaliação em contexto de línguas também influenciará na promoção de efeitos positivos. Assim, expõe:

> para explorar o potencial que exames externos poderiam trazer para as práticas na escola, o professor necessita estar capacitado para analisar profundamente esses exames e avaliar os resultados de seus alunos, o que pressupõe entendimento das matrizes de referência em que se fundamentam [os exames] e da relação que elas têm com o currículo e programas de ensino" (SCARAMUCCI, 2016, p. 157).

A demanda por um entendimento mais profundo, por parte de professores e outros envolvidos no processo de avaliação de línguas, das questões teóricas, práticas e sociais que permeiam o desenvolvimento e uso de testes (FULCHER, 2012; TAYLOR, 2009, 2013; SCARAMUCCI, 2016; QUEVEDO-CAMARGO; SCARAMUCCI, 2018) adveio do uso exponencial de exames nas últimas décadas, principalmente, como meio de elevar os padrões educacionais, com consequências graves para os envolvidos (SPOLSKY, 2008). No caso dos professores, essa necessidade se mostra ainda mais latente devido ao duplo papel que desempenham como facilitadores e avaliadores do processo de aprendizagem de línguas (SCARAMUCCI, 2006, 2016; INBAR-LOURIE, 2008a; SCARINO, 2013). Ademais, Fulcher (2012) é incisivo quando diz que os professores devem tomar consciência do poder investido nos testes e sobre como eles afetam suas práticas. Taylor (2013) define "letramento em avaliação" como parte de um conjunto de letramentos múltiplos que nós como indivíduos ou comunidades precisamos desenvolver, ou seja, "conhecimento, habilidades e competências relacionadas a um número crescente de domínios no nosso dia-a-dia, incluindo a língua e os padrões de discurso pertencentes a esses domínios"[29] (TAYLOR, 2013, p. 404).

Inbar-Lourie (2008a, p. 389) caracteriza o "letramento em avaliação" como a "habilidade de compreender, analisar e aplicar informações sobre o desempenho do aluno para a melhoria da instrução"[30] ou para a tomada de decisões de forma informada e consciente de suas consequências. Para Inbar-Lourie (2008a, 2013), o "letramento em avaliação em contexto de línguas" engloba, além das habilidades requeridas para o "letramento em avaliação", um componente linguístico e cultural que inclui teorias sobre aquisição de línguas, ensino e aprendizagem e uso da língua em diversos contextos.

Por serem conceitos recentes e em construção, ainda não se estabeleceu um consenso sobre quais seriam os conhecimentos, habilidades e competências necessárias para os

[29] No original: "[…] knowledge, skills and competence relating to a growing number of domains in everyday life; this includes the language and discourse patterns belonging to these domains".
[30] No original: "the ability to understand, analyze and apply information on student performance to improve instruction".

diversos envolvidos (protagonistas) nas práticas de avaliação e se haveria componentes essenciais ou uma base comum que todos deveriam possuir. No entanto, parece haver certa concordância de que o "letramento em avaliação em contexto de línguas" é um conceito multidimensional que envolve diversas camadas e níveis de competência, fazendo com que sua profundidade e extensão dependam das circunstâncias contextuais e das funções e responsabilidades dos envolvidos (INBAR-LOURIE, 2013).

Sobre as necessidades de professores de línguas em relação à área de avaliação, Fulcher (2012) analisou as respostas de 278 professores de várias regiões do mundo a um questionário, com o objetivo de fornecer informações para o desenvolvimento de cursos e materiais voltados para essa finalidade. Os resultados mostraram a necessidade de integrar princípios e conceitos a uma abordagem procedural (como fazer) na elaboração, desenvolvimento, manutenção e aplicação de exames em larga escala e em sala de aula, incluindo o aumento da conscientização sobre suas origens, razões e impactos.

Sobre o contexto brasileiro, Scaramucci (2016) pontua que os níveis de letramento em avaliação em geral e em contexto de línguas ainda são baixos. Até mesmo no campo da pesquisa em linguística aplicada, Scaramucci encontra deficiências no embasamento teórico de estudos que discutem questões referentes à avaliação, demonstrando desconhecimento de seus autores sobre a mudança de paradigma pela qual a área da avaliação passou nos últimos anos. Nas salas de aula, a avaliação ainda é essencialmente "somativa, classificatória, de produto, mecanicista, autoritária e excludente (SCARAMUCCI, 2016, p. 154), ignorando o seu papel de fornecer informações sobre a aprendizagem a fim de identificar falhas e corrigi-las. Para Scaramucci (2016), o problema reside na pouca importância dada à avaliação nos cursos de formação de professores em geral e de Letras. Esse óbice, no entanto, não é próprio do Brasil, ocorrendo também em outros contextos, como pontuado por Taylor (2013), que, por sua vez, acrescenta que o fato de muitos educadores perceberem a avaliação de línguas como altamente técnica e especializada contribui para a pouca disseminação da área.

Podemos dizer que a avaliação no Brasil, como em outros contextos, de modo geral, é ainda fortemente dominada pelo paradigma estruturalista, psicométrico e positivista, no qual a aprendizagem é vista como o acúmulo de conhecimentos e a avaliação tem o objetivo exclusivo de testar o que foi aprendido, restringindo seu papel ao de "avaliação *da* aprendizagem" (*assessment of learning*), o que Inbar-Lourie (2008b) denomina "cultura da testagem" (*testing culture*). Na cultura da testagem, impera a confiabilidade e os testes baseados em norma, nos quais a nota obtida pelo aluno é interpretada com base no desempenho de outros indivíduos (INBAR-LOURIE, 2008b; SCARAMUCCI, 2006).

Em oposição à cultura da testagem, a "cultura da avaliação" (*assessment culture*) está embasada na teoria construtivista, na qual a língua e a avaliação são entendidas como práticas sociais, situadas culturalmente e permeadas de valores. Nessa perspectiva, a avaliação e a aprendizagem estão interligadas e avaliar é uma forma não apenas de monitorar a aprendizagem, mas principalmente de promovê-la, a chamada "avaliação *para* a aprendizagem" (*assessment for learning*). Assim, são estimuladas múltiplas formas avaliativas para a coleta de informações sobre o progresso do aluno, retomando os objetivos da aprendizagem, identificando óbices, propondo correções e valorizando tanto o processo da aprendizagem (avaliação formativa) quanto o seu produto (avaliação somativa) (INBAR-LOURIE, 2008b; SCARINO, 2013; SCARAMUCCI, 1997, 2006, 2016).

O letramento em avaliação de professores pode ser considerado, portanto, um outro fator na determinação da intensidade e do tipo de influência que exames, principalmente, externos, exercem nos eventos e avaliações em sala de aula (FROETSCHER, 2017).

2.6 Estudos sobre efeito retroativo nas percepções e atitudes de alunos

Apesar de oito das quinze hipóteses de efeito retroativo, formuladas por Alderson e Wall (1993), fazerem referência direta aos alunos e ao processo de aprendizagem, pouco ainda se sabe sobre como opera o efeito retroativo nesses participantes, mesmo sendo eles uns dos principais afetados pelas consequências do uso de testes. Os alunos tendem a ser menos estudados como principal foco de investigação, quando comparados aos professores. Longe, porém, de ignorar o papel chave do professor na determinação do efeito retroativo, faz-se mister maiores investigações sobre como os alunos se comportam em relação a um exame (ALDERSON; BANERJEE, 2001; SCARAMUCCI, 2004; STONEMAN, 2006; CHENG et al, 2011).

São vários os focos de pesquisa sobre o efeito retroativo nesses participantes. Tendo como ponto de partida o modelo básico de efeito retroativo proposto por Bailey (1996) e fundamentado em Hughes (1993), o pesquisador pode, por exemplo, optar por restringir seu estudo a compreender as percepções e atitudes dos alunos, ou focar sua análise nos processos de aprendizagem (como eles se preparam), ou, ainda, investigar os produtos gerados (o desempenho obtido).

Bailey (1996) listou dez processos nos quais os alunos poderiam se engajar em decorrência da influência do exame, são eles: 1) praticar itens similares aos do teste; 2) estudar vocabulário e regras gramaticais; 3) conversar na língua alvo; 4) ler extensivamente na língua

alvo; 5) ouvir programas de rádio ou televisão na língua alvo; 6) aplicar estratégias para realização do exame; 7) matricular-se em cursos preparatórios; 8) solicitar orientação para estudos e *feedback* no desempenho; 9) participar, solicitar ou exigir aulas extras preparatórias; e 10) faltar às aulas de inglês para estudar para o teste.

Pearson (1988), por sua vez, apontou para uma possível relação entre o teste e a motivação do aluno. Hughes (1993) argumentou que o entendimento dos alunos sobre a natureza do exame influencia o modo como eles se preparam. Bachman e Palmer (1996) sugeriram que esse grupo poderia ser afetado de três maneiras: pela experiência na preparação para exames, pelo *feedback* recebido em testes anteriores e pelas decisões tomadas com base nos resultados obtidos.

Discuto aqui três estudos sobre efeito retroativo conduzidos no cenário internacional e três estudos brasileiros que tiveram como principal foco de investigação as percepções e atitudes de alunos, bem como os processos nos quais esses participantes se engajam, interesses desta pesquisa. Cheng (1998), por exemplo, analisou as percepções e atitudes de alunos frente a mudanças em um exame de alta relevância. Stoneman (2006) estudou a relação existente entre as percepções dos alunos sobre o *status* de um teste e a preparação. Cheng et al (2011) optaram por delimitar seu escopo de pesquisa às percepções de alunos e pais sobre o efeito retroativo de um exame em sala de aula. No Brasil, Bartholomeu (2002) e Correia (2003) exploraram as percepções e atitudes de alunos sobre a influência de exames vestibulares, e Lanzoni (2004) comparou os resultados obtidos em um exame de proficiência em leitura acadêmica a candidatos a programas de mestrado com o tipo de preparação realizada e suas percepções.

O estudo de Cheng (1998) buscou investigar em que grau as mudanças ocorridas em um exame de proficiência em inglês para alunos da educação secundária em Hong Kong influenciavam suas percepções e atitudes em relação às atividades realizadas dentro e fora do ambiente de sala de aula, as oportunidades de práticas da língua e as estratégias empregadas. Dois grupos de alunos participaram da pesquisa. O primeiro grupo era formado por alunos que realizariam o exame ainda no formato antigo, e o segundo grupo realizaria pela primeira vez o novo exame. Foram aplicados o mesmo questionário aos dois grupos. Os dados foram analisados comparativamente por meio do emprego de testes de hipóteses.

Os resultados revelaram que o efeito retroativo era ainda limitado e superficial. Aumento em atividades relacionadas à nova avaliação foram percebidas, tais como discussões em grupo e em pares, assim como aumento em atividades de produção oral. No entanto, a motivação e o uso de estratégias permaneceram inalterados, bem como o tempo de fala do

professor. Cheng (1998) concluiu que mudanças em um exame não garantem que inovações no ensino e aprendizagem sejam alcançadas. Normalmente, o que se veem são mudanças no conteúdo com pouca ou nenhuma influência do exame na abordagem de ensinar ou aprender. Todavia, acredito que o pouco tempo de implementação das novas mudanças (um ano) tenha contribuído significativamente para o baixo impacto encontrado.

Ainda no contexto de Hong Kong, a pesquisa de Stoneman (2006) investigou o impacto de dois exames de proficiência em inglês, aplicados aos graduandos de uma universidade em Hong Kong, e a relação existente entre o *status* dos exames e a preparação dos alunos. Stoneman (2006) beneficiou-se de uma oportunidade ímpar, uma vez que, durante a coleta de dados, o governo havia decidido substituir o exame desenvolvido e adotado pela universidade por um exame que fosse internacionalmente reconhecido − o IELTS (*International English Language Testing System*). Acreditava-se que pelo fato de o IELTS apresentar *status* maior que o exame interno, o impacto na aprendizagem dos alunos seria elevado.

Os dados revelaram que, de fato, o *status* do IELTS influenciou positivamente as percepções dos alunos quanto à importância do teste e à intenção de engajar-se em atividades de preparação, assim como a quantidade de horas dedicadas a essas atividades. Ou seja, os alunos que prestaram o IELTS externaram opiniões mais positivas sobre a importância do exame e da preparação, quando comparados com o primeiro grupo que prestou o exame interno. No entanto, o fato de o IELTS ter reconhecimento internacional e ser recomendado pelo governo em âmbito nacional, seu resultado não afetava as notas finais dos alunos, sem falar no seu caráter opcional. Tanto o IELTS quanto o exame interno não acarretavam consequências graves para os estudantes e, portanto, ambos careciam em termos de relevância. Stoneman (2006) concluiu que, para os alunos, a relevância de um exame é um fator de motivação mais importante que o *status* atribuído a ele. No entanto, Stoneman (2006) não realizou observações em sala de aula.

Alunos e pais foram o foco do estudo de Cheng et al (2011) que analisou, estaticamente, o impacto de uma avaliação de produção oral em inglês, conduzida em sala de aula nas percepções desses participantes. Essa avaliação tinha sido incorporada a um exame nacional de certificação em Inglês, considerado de alta-relevância na educação secundária de Hong Kong, com um peso equivalente a 15%. 389 alunos de escolas secundárias e 315 pais responderam aos questionários.

Os resultados apontaram que os alunos eram afetados de maneiras diferentes dependendo do grau de competência linguística percebida, ou seja, alunos com níveis mais altos

de proficiência apresentavam maior motivação para o estudo, realizavam mais atividades dentro e fora da sala de aula, possuíam visões mais positivas sobre exames externos e sentiam-se menos preocupados com os desafios impostos pelo teste conduzido em sala de aula, quando comparados aos alunos com competência julgada menor. Também foi apontado que, quanto mais perto da data do exame, maior era o impacto deste sobre as situações de aprendizagem. No entanto, dados qualitativos de sala de aula não foram coletados. Em relação aos pais, o estudo indicou que o nível de educação, o tempo compartilhado em família e o conhecimento sobre o sistema de avaliação da escola desempenhavam papel importante no apoio que eles ofereciam aos seus filhos em relação ao exame.

No contexto brasileiro, Bartholomeu (2002) investigou o efeito retroativo das provas de inglês do vestibular de três universidades públicas do Estado de São Paulo nas percepções, atitudes e motivações de alunos no último ano do ensino médio de uma escola particular. Os dados foram gerados por meio de questionários, entrevistas individuais e observações de aulas.

Os resultados dos questionários e entrevistas mostraram que os alunos, de modo geral, tinham conhecimento limitado e superficial sobre as provas, restritos ao formato (múltipla escolha) e ao conteúdo abordado (gramática, vocabulário e leitura). No entanto, características pessoais apresentadas por alguns alunos (autoconhecimento, autoeficácia[31] e autonomia) contribuíram para a promoção de atitudes mais positivas em relação aos exames. Também foi verificada associação entre a percepção de dificuldade das provas e a atitude dos participantes.

Em relação às aulas observadas, nenhuma evidência de efeito retroativo foi encontrada. As aulas da professora seguiam o conteúdo previsto na apostila e estavam voltadas para o ensino de gramática, com pouco interesse e participação dos alunos, provavelmente, devido à falta de prestígio da língua estrangeira em relação às demais disciplinas do currículo escolar e à crença de que não se aprende inglês na escola regular. Bartholomeu (2002) também mostrou evidências de que o ensino tem papel determinante nas percepções e atitudes dos alunos para com o exame, uma vez que a maioria deles depende do professor e da escola como fontes exclusivas de informação.

[31] Bandura (1982) argumenta que a relação entre conhecimento e ação é mediada pelo nível de autoeficácia do indivíduo, ou seja, sua percepção de o quão bem pode executar ações que lhe são apresentadas. Em outras palavras, a forma como o indivíduo julga sua capacidade na execução de uma tarefa pode afetar sua motivação e comportamento.

Com uma metodologia bastante semelhante à de Bartholomeu (2002), Correia (2003) analisou as percepções de alunos sobre a prova de inglês do vestibular da Universidade Estadual de Campinas (Unicamp) e seu efeito retroativo no modo como eles se preparavam para o exame dentro e fora da sala de aula. O contexto investigado foi um curso preparatório comunitário, na cidade de Campinas, para atender alunos de baixa renda que desejavam ingressar em uma universidade. Esperava-se um impacto mais intenso, com ênfase na leitura como construção de sentido, uma vez que esse curso era mantido pela Unicamp e ministrado por alunos e ex-alunos dessa instituição.

No entanto, os resultados revelaram baixa intensidade de efeito retroativo do exame, influenciando alguns alunos e não outros. Três diferentes perfis de alunos foram levantados, indicando que as características pessoais, tais como autonomia, motivação, autoeficácia interagem com o efeito retroativo. São eles: a) alunos que estudaram em cursos livres de inglês sentiam-se preparados, tinham expectativas de bons resultados e se dedicavam ao estudo da língua; b) alunos que estudaram inglês em cursos livres, porém não se sentiam preparados nem tinham expectativa de sucesso e, por essas razões, optaram por não se dedicar; e c) alunos que não estudaram em cursos livres, não se sentiam preparados ou com boas expectativas, porém continuavam se dedicando ao estudo.

Quanto ao conhecimento que os alunos possuíam sobre a prova de inglês, esse se restringia ao método (perguntas discursivas em português) e à habilidade exigida (leitura). Em relação às observações de aulas, embora tenha sido verificado efeito retroativo no material elaborado pelas professoras e utilizado em sala de aula, as aulas ainda estavam calcadas em uma leitura linear, com tradução integral ou parcial do texto e ensino de gramática descontextualizado, demonstrando que o exame não foi capaz de modificar as concepções das professoras sobre ensino de leitura.

Correia (2003) justificou que, ao demonstrar uma visão de leitura distinta daquela que fundamenta o exame de inglês da Unicamp, as professoras podem ter sido influenciadas por outros vestibulares, o que me parece, no entanto, pouco provável, uma vez que grande parte dos alunos (94%) manifestou preferência pelo vestibular dessa universidade. A meu ver, a visão ascendente de leitura das professoras pode, também, estar relacionada a experiências passadas de aprendizagem, capazes de enraizar crenças sobre ensinar e aprender difíceis de serem reconstruídas. Outros fatores contribuintes são salas numerosas, falta de recursos, sobrecarga de trabalho.

Assim como no estudo de Bartholomeu (2002), Correia (2003) também constatou que as percepções, atitudes e motivações dos alunos são fortemente dependentes do professor

e do ensino, reforçando a crença de que o professor é o responsável pela aprendizagem. Ademais, ambos corroboram resultados anteriores de Watanabe (2004b), Scaramucci (1999), Gimenez (1999) e Cheng (1999) de que mudanças conceituais no modo de ensinar do professor são difíceis de serem alcançadas somente por meio da avaliação, necessitando, portanto, de capacitação profissional contínua e voltada para a reflexão crítica.

O estudo de Lanzoni (2004) teve como foco de investigação o efeito retroativo de um exame de proficiência em leitura acadêmica nas percepções e atitudes de candidatos a programas de mestrado de uma universidade paulista. Primeiramente, foi aplicado um questionário a 90 candidatos, e os dados coletados foram triangulados com os resultados obtidos por eles no exame. Lanzoni (2004) dividiu, então, os candidatos em três grupos: aprovados, aprovados limítrofes e reprovados.

Por meio de análise estatística descritiva, foi possível observar uma relação entre o nível de proficiência demonstrado no exame e o efeito retroativo. Alunos com alto desempenho disseram ser pouco ou nada afetados pelo exame e apresentaram níveis mais baixos ou nulos de ansiedade. Alunos com baixo desempenho, por outro lado, apresentaram serem mais afetados pelo exame e relataram níveis mais altos de ansiedade. Esses resultados contrastam, em parte, com aqueles alcançados por Cheng et al (2011), em que a percepção de maior proficiência influenciava positivamente a intensidade do efeito. No caso de Lanzoni (2004), o que parece estar em jogo é a percepção de dificuldade do exame. Para o grupo com alto desempenho, a prova pode ter sido considerada fácil demais, o que não os mobilizou em termos de preparação.

Lanzoni (2004) também mostrou evidências de que quanto maior o desempenho do aluno menos dependente ele aparenta ser de estratégias de tradução palavra por palavra, o que é considerada uma prática ineficaz. Em relação ao tipo de efeito retroativo, evidências de efeito positivo e negativo foram encontradas entre os participantes, tais como propensão a ler textos em língua inglesa, atitude mais positiva para com o inglês, aumento excessivo da ansiedade e envolvimento em práticas preparatórias com provas anteriores. Para a análise qualitativa, Lanzoni (2004) coletou, por meio de questionários e entrevistas por telefone, dados de três candidatos que haviam prestado o exame duas ou três vezes antes de serem aprovados. Esses participantes reportaram aumento na ansiedade, dedicação à preparação para o exame, matriculando-se em aulas particulares ou em grupos, e mudança no uso de estratégias de tradução. Novamente, a hipótese de que o exame afeta alguns alunos e não outros é corroborada, bem como a de que características pessoais colaboram no tipo e na intensidade do efeito retroativo.

O exame analisado por Lanzoni (2004), como pontuado pelo próprio pesquisador, priorizava um modelo ascendente de leitura com ênfase na decodificação, localização de informação específica e tradução de trechos escritos, como modo de aumentar a confiabilidade de seus resultados. Assim sendo, apesar de Lanzoni (2004) sugerir que houve evidências de que a concepção de leitura dos participantes da fase qualitativa da pesquisa se modificou, a meu ver, essas mudanças podem também estar mais atreladas às estratégias empregadas para a realização do teste. Acredito que os participantes substituíram uma estratégia de tradução literal e sem sentido por uma tradução em que o produto é mais significativo e compreensível (WATANABE, 1996), uma vez que a visão de leitura que fundamentava o teste em questão permaneceu inalterada. Contudo, o fato de efeitos positivos terem sido constatados por Lanzoni (2004) corrobora o entendimento de que o fenômeno não é determinista, ou seja, um bom teste pode apresentar efeitos negativos, assim como um teste ruim pode gerar efeitos benéficos.

Estudos como os de Cheng (1998), Bartholomeu (2002), Correia (2003), Lanzoni (2004), Stoneman (2006) e Cheng et al (2011) atendem à necessidade de expandir as investigações sobre o efeito retroativo nos alunos como protagonistas. Outro fato interessante é a inclusão de mais participantes no processo de coleta de dados de alguns estudos como, por exemplo, pais e professores. Quevedo-Camargo e Scaramucci (2016) explicam que a participação de mais de um protagonista possibilita uma visão mais completa do fenômeno, assim como o emprego de múltiplos instrumentos de coleta de dados. No entanto, as pesquisas de Lanzoni (2004), Stoneman (2006) e Cheng et al (2011) se apoiaram em dados de entrevistas e questionários, o que limita o alcance de seus resultados, uma vez que as afirmações não puderam ser confrontadas com o que realmente ocorria em sala de aula ou na preparação para o exame. Isso, porém, não tira a importância desses estudos em alimentarem novas pesquisas. Podemos depreender de seus resultados alguns padrões de respostas entre os alunos, assim como sugeriu Burrows (2004) em relação aos professores. As percepções e atitudes dos alunos parecem ser determinadas não apenas pelo nível de entendimento que possuem sobre o exame, sua relevância e proximidade, conforme os estudos de Hughes (1989, 1993), Watanabe (2004a) e Smith (1991) sugeriram, mas também pelo nível de proficiência e algumas características pessoais como a motivação, a autonomia e a autoeficácia.

2.7 Preparação para o exame

Smith (1991) argumenta que os participantes investem o que for necessário para se prepararem ou preparar aqueles que irão realizar um exame frente às graves consequências que

seus resultados podem acarretar. Messick (1982, p. 9, tradução nossa) define a preparação para o teste (*coaching*) como "qualquer intervenção especificamente feita com o objetivo de aumentar as notas obtidas".[32] Apesar do termo "preparação para o exame" suscitar, à princípio, a ideia de efeito negativo, essas intervenções podem ser consideradas apropriadas. Messick (1982) explica que existe uma progressão nos tipos de preparação, variando de ações de curto prazo, com ênfase na manipulação simples de itens do exame, a instruções de qualidade de longo prazo, com o intuito de desenvolver o conhecimento e a capacidade do aluno nos domínios e habilidades que o teste se dispõe a avaliar.

Com o objetivo de distinguir práticas preparatórias adequadas e práticas preparatórias indesejáveis, Messick (1982, p. 32) listou três tipos de preparação. O tipo 1 desenvolve as capacidades e habilidades cobradas no teste, resultando no aumento das notas obtidas, porém não compromete a validade do exame. O segundo tipo de preparação melhora as estratégias para a realização do exame ao ensinar, por exemplo, técnicas de gerenciamento da ansiedade ou do tempo, o que reduz a variância irrelevante ao construto, propiciando uma avaliação mais precisa da capacidade do aluno. O tipo 3 de preparação visa ao ensino de estratagemas e artifícios para a identificação da resposta correta, o que prejudica a validade do exame, uma vez que causa ganho irreal nas notas obtidas.

O uso de artifícios para se chegar à resposta correta é explicado por Cohen (2006) como estratégias para responder testes (*test wiseness*). Em provas de múltipla escolha, por exemplo, o candidato pode procurar por sinais de correspondência entre o enunciado e as alternativas, sinais gramaticais (o enunciado corresponde gramaticalmente a uma alternativa), alternativas similares (eliminação de opções incorretas por conter informações semelhantes entre si) e itens que respondem outros itens. Cohen (1998) também considera a existência de outros dois grupos de estratégias: as estratégias de aprendizagem e as estratégias de uso da língua. Ambas consistem em ações selecionadas pelos indivíduos de forma consciente para melhorar a aprendizagem e/ou o uso da língua. As estratégias de aprendizagem são técnicas utilizadas para identificação e armazenamento do material linguístico a fim de direcionar o estudo. Podemos citar, por exemplo, a criação de mapas semânticos ou a associação de novo vocabulário a imagens e sons pelos aprendizes. As estratégias de uso relacionam-se ao emprego desse material linguístico em uma situação de compreensão e produção linguísticas e podem também ser aplicadas na realização de uma tarefa avaliativa. Elas se dividem em: estratégias de resgate do material linguístico a ser utilizado, estratégias de treinamento para praticar o que foi

[32] No original: "Any intervention procedure specifically undertaken to improve test scores".

aprendido, estratégias de proteção que dão a impressão de que o material linguístico é maior do que ele realmente é, e estratégias de comunicação utilizadas na produção de uma mensagem significativa.

Popham (1991) associou a proliferação de práticas preparatórias inadequadas e antiéticas à pressão sofrida por alunos, professores e diretores na obtenção de resultados satisfatórios em exames públicos obrigatórios. Nesse sentido, a relevância do exame torna-se uma faca de dois gumes: ao mesmo tempo que é vista como fator impulsionador na direção de inovações curriculares, pode ocasionar danos graves ao sistema educacional. Qi (2004) chamou a atenção para a necessidade de se manter a relevância de um exame em um nível ideal (*right level*) a fim de evitar o aumento de práticas inadequadas. Hamp-Lyons (1998) solicitou a elaboração de um código de ética que proíba esses tipos de preparação e argumenta que a responsabilidade por seus danos recai sobre todos os envolvidos no processo.

Popham (1991) apresenta cinco tipos de preparação para o exame: preparação com provas anteriores (a instrução se baseia no uso de versões anteriores já publicadas do exame); preparação em estratégias gerais para realização de testes (a instrução apresenta estratégias para lidar com vários formatos de provas e itens como, por exemplo, como gerenciar o tempo e compreender enunciados), preparação em mesmo formato (a instrução foca diretamente o conteúdo coberto pelo exame e pratica itens no mesmo formato da prova); preparação em formatos variados (a instrução foca diretamente no conteúdo abordado pelo exame, porém emprega uma gama de itens em diferentes formatos) e preparação com prova atual (a instrução é voltada para a prática dos itens que serão empregados no teste).

Para determinar a adequação dos diferentes tipos de preparação, Popham (1991) propôs a consulta a duas normas: a ética profissional e a defesa educacional. A primeira faz referência à violação da responsabilidade ética, ou seja, nenhum professor pode utilizar práticas preparatórias que violem os princípios éticos, tais como roubo, fraude e falsidade. A defesa educacional, por sua vez, indica que nenhuma prática preparatória pode aumentar os resultados do exame sem, simultaneamente, desenvolver a competência do aluno no domínio avaliado.

Somente dois tipos de preparação são considerados éticos e adequados de acordo com Popham (1991): a preparação em estratégias gerais para a realização de um exame, pois faz com que os alunos estejam familiarizados com o formato do teste e os tipos de tarefas, e a preparação em formato variado que aumenta as notas no exame ao mesmo tempo que desenvolve a proficiência do aluno na língua. Esses dois tipos de preparação assemelham-se, respectivamente, aos Tipo 2 e Tipo 1, propostos por Messick (1982).

A preparação com provas anteriores e a preparação no mesmo formato, na opinião de Popham (1991), ferem a defesa educacional ao inflar as notas sem, ao mesmo tempo, promover o desenvolvimento da competência. Além disso, essas práticas também são capazes de estreitar o currículo, ocasionando, assim, um efeito indesejado. A preparação com prova atual, além de ser considerada uma prática ilegal ocasionada pelo vazamento ou roubo de uma cópia da prova, fere também a defesa educacional inflando as notas sem demonstrar qualquer ganho na proficiência, uma vez que os alunos podem, facilmente, memorizar as respostas.

Haladyna et al (1991) também veem como indesejáveis práticas preparatórias cujos currículos são direcionados para o teste, em termos de conteúdos, habilidades e tipos de questões, por constituírem uma possível ameaça à validade. Haladyna e Downing (2004) explicam que, pelo fato de um teste ser capaz de avaliar apenas uma pequena parcela das habilidades que fazem parte de seu construto, a prática exclusiva com itens dessa amostra pode aumentar as notas obtidas, porém não garante que o aluno tenha desenvolvido sua competência no que não foi diretamente testado. Como práticas válidas, Haladyna et al (1991) citam as estratégias gerais que familiarizam os alunos com o formato dos itens de prova, o ensino de procedimentos no preenchimento do formulário de respostas e as técnicas de motivação e estímulo.

Com base nessas discussões, Xie (2013) sugere delimitar o primeiro tipo de preparação para o teste (Tipo 1) de Messick (1982) ao desenvolvimento de uma ampla gama de habilidades que fazem parte do que se pretende avaliar, o que é visto como adequado, e acrescenta a essa proposta do pesquisador a preparação Tipo 4, caracterizada pelo uso de itens idênticos ou similares àqueles encontrados no teste, sendo essa considerada inadequada.

Fica evidente certa contradição nessas classificações quanto à adequação das práticas preparatórias: é visto como necessário e positivo promover a familiaridade do aluno com o formato do exame, porém a preparação com provas anteriores e a preparação no mesmo formato são consideradas práticas inadequadas ou até mesmo antiéticas. A meu ver, seria prematuro definir como inadequado o uso em sala de aula de itens idênticos ou similares aos do teste, sem ponderar, por exemplo, sobre a extensão desse uso. Dessa forma, tomo como prejudiciais ao ensino somente práticas preparatórias voltadas exclusivamente para o formato contido no exame, pois apresentam o potencial de estreitar o currículo e os métodos de instrução. Porém, em pequena escala, acredito que elas podem e devem fazer parte do repertório do professor. Em outras palavras, o professor pode utilizar, em sala de aula, itens de teste disponíveis ou elaborar questões similares às do exame, sem que isso seja considerado como algo indesejável, mas como forma de familiarizar os alunos com o teste que irão realizar.

Ao considerar que o exame investigado neste estudo é de alta-relevância, acredito que professores e alunos irão se engajar em práticas preparatórias para obtenção de resultados desejáveis. As classificações aqui apresentadas serão, então, utilizadas como ponto de partida para a análise de como os alunos estão sendo preparados ou se preparam para o exame. Julgamentos quanto à adequação dessas práticas e seus efeitos, tais como o estreitamento do currículo e a redução dos métodos de instrução, serão feitos com base na variedade e extensão com que são aplicadas e não apenas na sua presença ou não em sala de aula.

Esse posicionamento encontra embasamento em pesquisas como a de Gebril e Eid (2017) que demonstrou que as percepções de adequação das práticas preparatórias não são consensuais. Os participantes desse estudo eram professores do 2º grau em escolas públicas do Egito e estavam envolvidos diretamente com a preparação de alunos para o *thanaweya amma* – um exame de inglês considerado de alta-relevância por determinar a conclusão dos estudos secundários e a entrada na universidade. Os resultados apontaram que a grande maioria desses professores consideraram apropriados o ensino de estratégias para a realização do exame, a prática com exames anteriores, a prática no mesmo formato do teste e a revisão de conteúdo. De forma geral, eles externaram opiniões positivas acerca de práticas preparatórias para o exame, uma vez que promovem a familiarização com o teste e seu formato trazendo resultados melhores.

Outros estudos têm se dedicado à análise da relação entre o tipo de preparação recebida e o desempenho do aluno no teste. Green (2007b), por exemplo, investigou se os alunos que frequentavam cursos preparatórios com foco no IELTS (o que ele denominou curso Tipo A) obtinham notas mais altas na parte escrita do exame, quando comparados a alunos que frequentaram cursos de inglês para fins acadêmicos (Tipo B) e cursos que combinavam a escrita acadêmica e a preparação para o exame (Tipo C). A princípio, acreditava-se que os alunos que optaram por uma preparação focada no IELTS alcançariam melhores resultados, mesmo sem, necessariamente, adquirirem algumas das habilidades-chave na escrita acadêmica.

Para tanto, foram analisadas as relações entre variáveis ligadas aos participantes, ao curso e ao desempenho (resultados obtidos no pré-teste e no pós-teste e em testes de gramática e vocabulário), consideradas capazes de mediar o efeito retroativo do teste. Os resultados mostraram que os alunos que frequentaram cursos do Tipo A ou C não obtiveram nenhuma vantagem em termos de notas no exame em relação ao Tipo B. Pelo contrário, as características individuais dos alunos, tais como motivação, autoconfiança, conhecimento das exigências do teste pareceram exercer maior influência sobre o desempenho do que o tipo de curso realizado. Green (2007b) não analisou, porém, se, em áreas não testadas pelo teste, esses diferentes alunos

alcançariam o mesmo grau de competência. Resultados como esses colocam a afirmação de que a preparação em mesmo formato ou com provas anteriores promove notas acima da capacidade dos alunos sob questionamento.

Além disso, embora esta pesquisa não tenha como propósito analisar os resultados obtidos pelos alunos no exame e compará-los com o tipo de preparação oferecida, as conclusões de Green (2007b) também nos chamam a atenção para as características individuais dos alunos e suas atitudes na preparação para o teste, mostrando que, se quisermos compreender melhor como o efeito retroativo atua no aluno e em sua aprendizagem, é importante considerarmos esses aspectos.

Xie e Andrews (2012), por sua vez, analisaram a influência de um exame de proficiência em inglês, aplicado a universitários chineses, nas percepções de alunos em relação aos conhecimentos e habilidades exigidos pelo teste (*test design*), aos seus usos (*test uses*) e suas relações com a preparação para o exame. Para tanto, os autores se apoiaram na teoria de expectativa-valor (*expectancy-value theory*), oriunda da Psicologia da Aprendizagem e relacionada ao conceito de autoeficácia, em que "a escolha dos indivíduos, a persistência e o desempenho podem ser explicadas por suas crenças sobre o quão bem-sucedidos eles serão na atividade e pelo grau com que valorizam a atividade"[33] (WIGFIELD; ECCLES, 2000, p. 68 apud XIE; ANDREWS, 2012, p. 52). Em outras palavras, essa teoria pressupõe que as ações de um indivíduo na realização de uma tarefa são influenciadas pelo grau de confiança que ele deposita sobre sua capacidade e pela importância que atribui ao seu sucesso na atividade.

Os resultados desse estudo apontaram que as percepções sobre as demandas do teste (*test design)* tendem a impactar mais o engajamento em atividades preparatórias do que as percepções acerca dos usos do teste (*test uses*). A influência das demandas do exame sobre a preparação parece ser mediada significativamente tanto pela expectativa de sucesso quanto pelo valor atribuído ao teste. Assim, os alunos que reconhecem as habilidades e conhecimentos exigidos pelo teste como necessários tendem a demonstrar maior grau de confiança e, consequentemente, maior envolvimento em sua preparação. Xie e Andrews (2012) ainda explicam que, mesmo que os alunos não possuam as habilidades e conhecimentos avaliados pelo exame, conhecer o que é necessário para a realização dessa tarefa pode dar a eles condições de direcionar e controlar a preparação. Quanto ao valor do teste, os autores (2012) constataram

---

[33] No original: "Individuals' choice, persistence, and performance can be explained by their beliefs about how well they will do on the activity and the extent to which they value the activity".

que atitudes positivas para com as demandas do teste relacionam-se com altos graus de importância atribuída a ele, o que pode levar esses alunos a demonstrarem maior engajamento com práticas preparatórias. Já, em relação aos usos do teste, somente o fator "valor" apresentou associação significativa com a preparação. Isso significa que a percepção de valor atribuído ao teste parece endossar ou não os seus usos, o que, por sua vez, irá resultar em maior ou menor engajamento do aluno na preparação.

O estudo conduzido por Stenlund et al (2017) também nos diz bastante a respeito de como alunos de diferentes níveis de proficiência se comportam em relação à preparação para um exame considerado de alta-relevância, ao analisar as diferenças entre examinandos, no emprego de estratégias durante a realização do teste em função do desempenho obtido. Seus resultados mostraram que examinandos com alto desempenho tendem a se educar mais a respeito do teste e seu conteúdo, enquanto examinandos com baixo desempenho relatam utilizar, exclusivamente, versões anteriores do exame como fonte de informações sobre suas demandas.

Os resultados de todos esses estudos somam-se àqueles apresentados anteriormente e sua importância para esta pesquisa reside nas observações feitas sobre os fatores que intermediam as atitudes dos participantes em relação ao exame e sua preparação. Vimos, por meio de Stoneman (2006), que a relevância atribuída ao exame é um fator de motivação maior do que o seu *status*. O estudo de Green (2007b) revelou que características individuais dos alunos (motivação, autoconfiança e conhecimento do exame) apresentam maior influência sob o resultado de um exame do que o tipo de preparação oferecida. Correia (2003) e Bartholomeu (2002) também observaram que fatores individuais interagem na intensidade do efeito retroativo. Cheng et al (2011), por sua vez, nos mostraram que o nível de proficiência dos alunos se relaciona à motivação e ao engajamento em práticas preparatórias. Resultados semelhantes foram encontrados por Stenlund et al (2017) ao concluir que examinandos com baixa proficiência tendem a utilizar estratégias menos eficazes, apresentam níveis mais altos de ansiedade e conhecem menos o teste. Por fim, Xie e Andrews (2012) ressaltaram que as percepções dos alunos sobre as exigências do exame determinam um maior ou menor engajamento em práticas preparatórias.

Esta pesquisa propõe-se a analisar as percepções, atitudes e ações de dois grupos de participantes, alunos e professores, sobre a influência de um exame considerado de alta-relevância, com o objetivo de compreender os fatores que atuam no efeito retroativo e fazem com que esses participantes ajam de determinada maneira. Serão analisadas, primeiramente, mudanças nas percepções e atitudes de professores impulsionadas pela implementação do

exame no contexto escolar. As percepções e atitudes dos alunos serão investigadas em termos de proficiência, conhecimento sobre o exame e sua proximidade. Cientes de que os modelos de efeito retroativo disponíveis ainda não exauriram completamente toda a complexidade do fenômeno, a identificação de padrões é bastante relevante para esta pesquisa, pois colabora para a construção de uma teoria mais satisfatória sobre o efeito retroativo, além de mostrar com mais precisão onde estão as deficiências que impedem que mudanças de fato se concretizem.

2.8 A relação entre validade e efeito retroativo

Os anos de 1970 e 1980 presenciaram mudanças profundas na visão de língua, ensino e aprendizagem e avaliação com a emersão de novos construtos como a competência comunicativa (HYMES, 1972; CANALE; SWAIN, 1980; BACHMAN 1990), língua e avaliação como práticas sociais (SHOHAMY, 1993) e testes comunicativos e de desempenho (FREDERICKSEN; COLLINS, 1989). Dentro desse cenário, os conceitos tradicionais de validade e confiabilidade mostraram-se insuficientes para avaliar a adequação dos métodos de avaliação que surgiam e começaram a ser bastante questionados por pesquisadores tanto na área de avaliação educacional quanto em contexto de línguas. Esses debates formaram a base para uma revolução no conceito de validade, ocorrida no final dos anos de 1980 e início dos anos de 1990 (SCARAMUCCI, 2011a).

A validade, até os anos de 1970, era entendida como uma característica inerente ao teste utilizada para determinar sua qualidade e, consequentemente, sua aceitabilidade (CHAPELLE, 1999; SCARAMUCCI, 2011a). Assim, um teste era considerado válido se medisse o que se propunha a medir. A confiabilidade, ou a consistência dos resultados, era outra importante característica, sendo vista, muitas vezes, como um pré-requisito para a validade. Nessa visão, um teste era considerado válido se os seus resultados se correlacionassem com os de outros testes, alcançando, assim, a denominada validade de critério. Dois outros tipos de validade costumavam ser investigados: a validade de conteúdo consistia no julgamento de especialistas sobre a adequação do conteúdo abordado e a validade de construto definia, por meio de estudos empíricos, se o teste estava de acordo com suas bases teóricas (CHAPELLE, 1999).

No final dos anos de 1970, iniciaram-se os primeiros movimentos de pesquisadores envolvidos com a validação de construto em testes de competência comunicativa que abririam espaço para a reconstrução do conceito de validade. Dentre eles, Chapelle (1999) destaca o trabalho de Henning (1987) que expandiu o trio de validade (critério, construto e conteúdo) ao

acrescentar a validade de resposta – relacionada ao estudo dos processos cognitivos nos quais os candidatos se engajam ao responder os itens de um teste, e ao dividir a validade de critério em paralela e preditiva, em função do tempo em que foram coletadas as informações. Assim, para se chegar à validade paralela, compara-se os resultados obtidos no teste com outros testes (ou outros instrumentos de medida) realizados pelo mesmo candidato em um espaço bem curto de tempo. A validade preditiva envolve a comparação do resultado obtido no teste com o desempenho do candidato após certo período de tempo.

As contribuições de Henning (1987), entretanto, não se limitaram à expansão dos métodos para a investigação da validade, mas também introduziram a ideia de que um teste era válido para certos propósitos e não para outros e de que existia um contínuo, discordando, portanto, da visão tradicional de que a validade era um atributo de tudo ou nada e extremamente dependente da confiabilidade.

Para explicar a relação entre os conceitos de validade e confiabilidade, em que o aumento de uma poderia ocasionar a diminuição da outra, Scaramucci (2011a) cita, como exemplo, a tradução de uma lista de palavras como teste cujos resultados são bastante confiáveis, porém considerado inválido se o propósito for o de avaliar a produção escrita do candidato, uma vez que essa habilidade vai além da tradução pura e simples de palavras. Da mesma forma, redigir um texto pode ser tido como um teste válido para esse propósito, todavia seus resultados podem apresentar inconsistências caso não haja critérios de correção claros e treinamento apropriado para os avaliadores (HUGHES, 1989; SCARAMUCCI, 2011a). Bachman (1990) pontua que quanto mais autênticos são os testes, mais difícil ou impossível se torna separar evidências de confiabilidade de evidências de validade, uma vez que as características avaliadas (*traits*) podem se mesclar aos métodos empregados para medi-las. Essas observações retratam, claramente, o quão inadequados se apresentavam esses conceitos para lidar com os novos instrumentos de avaliação – os testes de desempenho.

Outras contribuições para a posterior ressignificação do conceito de validade incluíram questões ligadas ao afeto (ansiedade provocada pelo teste e sua influência no ensino e aprendizagem), à ética no processo de avaliação de línguas e à influência do exame no ensino e aprendizagem de línguas (efeito retroativo). Hughes (1989), por exemplo, apresentou o conceito de efeito retroativo como uma das qualidades que deveriam ser demonstradas pelo teste e sugeriu que as percepções dos usuários (professores, alunos, pais, entre outros) sobre a credibilidade e aceitação do exame deveriam ser também investigadas, uma vez que poderiam influenciar as ações realizadas em sala de aula, atestando, assim, validade de face ao teste.

Termos distintos são utilizados por outros autores para categorizar os diversos tipos de validade discutidos na época. Alderson et al (1995), por exemplo, os dividem em interna, externa e de construto. A validade interna refere-se ao conteúdo e ao efeito do teste, subdividindo-se em validade de conteúdo, de face e de resposta. A validade externa ou de critério pode ser determinada pela validade paralela e pela validade preditiva. Por sua vez, a validade de construto é considerada por Alderson et al (1995) a mais complexa e difícil de ser explicada por abranger as demais. A palavra construto refere-se a um aspecto ou característica psicológica que não pode ser medido ou observado diretamente. Por exemplo, a proficiência linguística é considerada um construto, e os testes de proficiência são instrumentos que operacionalizam esse construto. Para demonstrar a validade de construto de um teste, coletam-se evidências que permitam afirmar que o desempenho demonstrado é uma medida significativa das características que o teste se propõe a medir.

Embora as discussões sobre as qualidades dos exames e os diferentes métodos para determinar sua validade tenham se intensificado entre os teóricos da avaliação de línguas, principalmente, nos anos de 1980, Chapelle (1999) relata que pouquíssimas questionavam diretamente e explicitamente o conceito de validade. Já na área de avaliação educacional, o debate sobre validade tomava grandes proporções, o que resultou em três importantes ações responsáveis pela sua reconstrução. São elas: a substituição do trio de validade por um conceito unificado em que a validade de construto toma posição central nas Normas para Avaliação Educacional e Psicológica da AERA/APA/NCME,[34] os questionamentos acerca dos fundamentos tradicionais do processo de validação, e a publicação do artigo seminal de Messick (1989) intitulado Validade.

O conceito tradicional de validade, fragmentado e incompleto, no sentido de não abranger às implicações de valor às quais as interpretações dos resultados dos testes estavam sujeitas, tampouco as consequências sociais advindas de seu uso, é, então, substituído por um conceito fundamentado e unificado em torno da validade de construto, e cuja coleta de evidências perpassa a natureza social da avaliação. Nas palavras de Messick (1989), a validade é definida como "um julgamento avaliativo integrado do grau com que as evidências empíricas e as fundamentações teóricas sustentam a adequação e apropriação das inferências e ações baseadas nos resultados dos testes"[35] (MESSICK, 1989, p. 13, tradução nossa).

---

[34] AERA: *American Educational Research Association*, APA: *American Psychological Association*, e NCME: *The National Council on Measurement in Education*.

[35] No original: "Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on the test scores".

Essa definição retirou, definitivamente, o entendimento de validade como uma característica do teste, ao afirmar que o que devem ser considerados válidos ou não são a interpretação e o uso que se faz de seus resultados. Dessa forma, o processo de validação de um teste deve coletar evidências de diversas fontes que fundamentem as inferências que estão sendo feitas e as ações que estão sendo tomadas (MESSICK, 1989; BACHMAN, 1990; CHAPELLE, 1999; SCARAMUCCI, 2011a). A confiabilidade, vista, tradicionalmente, como outra característica do teste e, muitas vezes, como pré-requisito para a sua validade, se integra ao conceito de validade como um tipo de evidência, fazendo-se, portanto, necessária a coleta de informações sobre o quão consistentes são os resultados de um teste, porém não suficiente para se afirmar que as inferências e os usos são adequados e apropriados (CHAPELLE, 1999; SCARAMUCCI, 2011a). Decisões tomadas em relação à praticidade do teste − definida por Bachman e Palmer (2010) como sendo a diferença entre os recursos necessários para a elaboração de um exame e os recursos disponíveis, também passam a serem consideradas como tipo de evidências em um processo de validação.

Para compreendermos melhor o novo conceito de validade e o processo de validação, Messick (1989) propôs uma matriz de validade (Figura 5), na qual conecta as bases evidencial e consequencial ao processo de levantamento das informações necessárias à interpretação e ao uso dos resultados do teste, enfatizando, assim, a natureza social da avaliação.

Figura 5 − Matriz progressiva de validade

	INTERPRETAÇÃO DO TESTE	USO DO TESTE
BASE EVIDENCIAL	Validade de Construto	Validade de Construto + Relevância / Utilidade
BASE CONSEQUENCIAL	Implicações de Valor	Consequências Sociais

Fonte: MESSICK, 1989, p. 20 (tradução nossa)

Essa matriz apresenta característica progressiva no sentido de que a célula seguinte contém as características da célula anterior, a fim de "retratar a validade como um conceito unitário, porém multifacetado" (CHAPELLE, 1999, p. 258). Sobre as questões ligadas à natureza social da avaliação, Scaramucci (2011a, p. 112) esclarece que:

> a dimensão social da avaliação está representada em duas das células da matriz: naquela que considera as implicações, focalizando o caráter social e cultural dos significados atribuídos aos escores do teste; e naquela que leva em conta as

> consequências sociais relativas ao uso prático dos testes (SCARAMUCCI, 2011a, p. 112).

Bachman (1990, p. 242) esmiúça o mecanismo de funcionamento dessa matriz da seguinte forma: se a nossa intenção é somente justificar determinada interpretação de um resultado, devemos nos atentar para a primeira coluna da matriz, coletando evidências para a validade de construto e considerando as implicações de valor que envolvem essa interpretação. Se, no entanto, nosso propósito vai além desse e inclui utilizar o resultado para tomar uma decisão, devemos justificar esse uso, não investigando somente a validade de construto e as implicações de valor das interpretações que fazemos dos resultados, mas também a relevância e utilidade desse uso, bem como as consequências sociais advindas, representadas na segunda coluna da matriz.

As facetas presentes no conceito de validade de Messick (1989), para as quais é necessário coletar evidências de validade que embasem determinada inferência ou decisão em um processo de validação de um teste, são: de conteúdo, substantiva, estrutural, passível de generalização, externa e consequencial. Essas seis facetas contribuem para a validade de construto do teste, como forma de estabelecermos em que medida as interpretações e os usos que fazemos com base no desempenho obtido em determinado exame condizem com as habilidades que pretendemos avaliar (BACHMAN, 1990). Assim, a análise de conteúdo avalia a relevância, representatividade e qualidade do conteúdo avaliado no teste. A análise substantiva fornece evidências sobre os processos e conhecimentos envolvidos na solução dos itens do teste pelos candidatos. A estrutural investiga a estrutura interna do exame e a evidência passível de generalização estuda o grau com que o desempenho no teste pode ser generalizado para outras situações. A consequencial refere-se à análise das implicações de valor advindas das interpretações dos resultados do teste e das consequências sociais de seus usos, por meio da coleta de evidências de seus impactos, sejam esses intencionais ou não intencionais. O efeito retroativo de testes nos contextos específicos de ensino e aprendizagem, foco de investigação desta pesquisa, é considerado um tipo de consequência social do uso do teste, assim como o são as questões ligadas a viés, ética e justiça (SCARAMUCCI, 2011a).

Cabe aqui uma ressalva quanto à validade de face, discutida anteriormente. Embora, Bachman e Palmer (2010) argumentem que esse tipo de validade não pode ser aceito como evidência para analisarmos a adequação das inferências feitas com bases nos resultados do teste, ou seja, a validade das interpretações, a investigação das percepções e atitudes de professores e alunos em relação a um teste traz contribuições importantes para um estudo de efeito

retroativo, uma vez que o modo como esses participantes veem o instrumento de avaliação pode influenciar suas ações em sala de aula.

Devido, portanto, às múltiplas facetas que compõem o conceito de validade, a validade de um teste só pode ser demonstrada ou estabelecida, em um processo de validação, se forem coletadas informações de naturezas diversas, incluindo evidências consequenciais, que fundamentem determinada interpretação ou uso. Nas palavras de Scaramucci (2011a), "os aspectos consequenciais deveriam ser considerados em conjunto com outras evidências propostas por Messick em sua matriz; um teste válido seria aquele que apresentaria um conjunto de evidências convergentes" (SCARAMUCCI, 2011a, p. 113).

Messick (1996) discorda, portanto, da proposta de Fredericksen e Collins (1989) de que o efeito retroativo de um exame no contexto educacional, positivo ou negativo, deveria, por si só, ser capaz de determinar a validade de um teste, e acrescenta que o efeito retroativo somente pode ser considerado como tal se as evidências estiverem ligadas diretamente à sua introdução e uso. Para Messick (1996, p. 245, tradução nossa), os exames devem ser elaborados com o objetivo de promover efeitos retroativos positivos (*washback by design*), por meio do controle das ameaças à validade: sub-representação do construto e variância irrelevante ao construto. Feito isso, caso ainda existam efeitos negativos em jogo, a validade do teste não pode ser totalmente comprometida, já que "o efeito retroativo parece depender de uma série de fatores importantes no sistema educacional, além da validade dos testes".[36] Scaramucci (2011a, p. 115) concorda com Messick (1996) e sugere que não deveríamos investigar o efeito retroativo com o objetivo de determinar ou não a validade do teste, mas para "melhorá-lo nos aspectos que podem, potencialmente, apresentar problemas ou levar a consequências ruins ou que estão causando consequências ruins".

No entanto, demonstrar a causalidade entre teste e efeito constitui uma grande dificuldade para pesquisadores, que nem sempre possuem dados sobre o contexto antes da implementação do teste (SCARAMUCCI, 2004). A alternativa defendida por Scaramucci (2004) e por outros teóricos, como Qi (2004), é a de contrapor as intenções divulgadas pelos elaboradores de exames com o teste e com os eventos em sala de aula, a fim de identificar congruências e discrepâncias entre eles.

A proposta de Messick (1989) de integrar o estudo das consequências sociais ao conceito de validade e, consequentemente, ao processo de validação de um teste, encontra resistência. Muitos teóricos da avaliação defendem que uma investigação sobre esses impactos

[36] No original: "Washback appears to depend on a number of important factors in the educational system in addition to the validity of tests".

não deveria fazer parte da validade de um teste por ultrapassarem os limites de seu construto. Os apoiadores dessa visão, segundo Scaramucci (2011a), costumam justificar que as consequências sociais fogem da alçada e do controle do elaborador do teste, devendo recair sua investigação sobre os responsáveis pelos usos dos resultados.

Se por um lado seria ingênuo de nossa parte acreditar que o elaborador do exame é capaz de controlar todas as consequências advindas de seus usos (SCARAMUCCI, 2004), por outro, Scaramucci (2011a) ressalta que a introdução de testes, principalmente aqueles que carregam alto grau de relevância, implica, inevitavelmente, em tomada de decisão que pode afetar indivíduos ou sociedade, sendo não apenas necessário, mas também fundamental, investigar as consequências sociais dos usos de seus resultados em um projeto de validação. Assim, também, pontua Bachman (1990):

> Dado o importante papel que o teste agora desempenha em influenciar decisões educacionais e sociais sobre os indivíduos, nós não podemos mais limitar nossa investigação de validade à coleta de evidência factual para sustentar uma dada interpretação ou uso[37] (BACHMAN, 1990, p. 237).

Kane (2002), do mesmo modo, vê como deficientes muitos processos de validação de exames de alta-relevância por terem negligenciado ou subestimado as consequências sociais pretendidas e não-pretendidas de seus testes. McNamara (2010) é outro grande defensor da necessidade de um olhar mais crítico para os valores e propósitos político-sociais implícitos nos exames e reforça que as pesquisas de validade não devem se restringir à qualidade do teste em si, atribuindo maior relevância às considerações de ordem prática em detrimento de sua importância social.

A meu ver, esse impasse sobre considerar ou não as consequências sociais como parte da investigação da validade, decorre, parcialmente, da forma como Messick (1989) desenhou sua matriz. Embora ele tenha salientado que sua proposta de validade é um conceito unificado ao redor da validade de construto e essa, por sua vez, contempla as consequências sociais advindas do uso dos resultados do teste, sua matriz parece não retratar, exatamente, isso. Quando Messick (1989) particiona as evidências de validade em duas bases, evidencial e consequencial, temos a impressão de que a validade de construto independe da coleta de evidências sobre as consequências sociais, apesar de juntas formarem um todo. Isso pode fazer com que elaboradores de testes se atenham à linha superior da matriz no processo de validação

[37] No original: "Given the significant role that test now plays in influencing educational and social decisions about individuals, however, we can no longer limit our investigation of validity to collecting factual evidence to support a given interpretation or use".

e releguem a outros a responsabilidade das duas dimensões encontradas na linha inferior, ou seja, a base consequencial. McNamara e Roever (2006) também pontuam que o problema envolvendo o modo como as duas camadas da matriz se relacionam nunca foi resolvido por Messick.

Fundamento minha pesquisa na visão de que as consequências sociais advindas do uso de um teste devem fazer parte do projeto de validação desse instrumento, fazendo-se, portanto, necessária e fundamental a coleta de evidências sobre os impactos, juntamente com outros tipos de evidências, no estudo da validade de um exame. Sobre a quem caberia a responsabilidade pela coleta de evidências de consequências, concordo com a colocação de Scaramucci (2011a, p. 116) de que "se a validação é um processo contínuo e sem fim, caberia a todos a coleta de dados sobre esse processo". Assim sendo, no contexto deste estudo, eu me vejo como representante de diversos protagonistas: pertenço ao quadro de servidores do DECEA, instituição responsável pela implementação dos requisitos de proficiência linguística da OACI para controladores de tráfego aéreo e operadores de estação aeronáutica brasileiros; sou uma das examinadoras do EPLIS pelo ICEA, instituição responsável pelo desenvolvimento, manutenção e aplicação do exame; atuei como professora de inglês aeronáutico na EEAR, escola responsável pela formação dos controladores militares; e, atualmente, ministro cursos de inglês aeronáutico presenciais e a distância, oferecidos pelo ICEA, para os profissionais em serviço; além de ser usuária do transporte aéreo.

É necessário esclarecer que este estudo não tem como propósito tecer considerações sobre a validade do EPLIS, uma vez que considero as consequências sociais apenas como uma das diversas facetas que integram o conceito de validade, informando sobre a validade de construto do exame, e que necessitam serem investigadas. Ademais, o impacto do exame nos processos realizados em sala de aula é um tipo de evidência de consequência dentre as diversas possíveis de serem coletadas como, por exemplo, as relacionadas a viés, ética e justiça no uso de testes. Para delimitar a abrangência desta pesquisa e de seus resultados, me apoio nas palavras de Scaramucci (2011a) a respeito dos estudos sobre efeito retroativo.

> analisar as consequências de um teste não para concluir se é válido ou não, mas apenas para melhorá-lo nos aspectos que podem, potencialmente, apresentar problemas ou levar a consequências maléficas ou que já estão apresentando consequências maléficas (SCARAMUCCI, 2004, p. 117).

Dessa forma, esta pesquisa é vista como parte integrante de um estudo maior de validação do EPLIS, de natureza contínua, e que envolve a coleta e análise de evidências de

diversas fontes (SCARAMUCCI et al, 2010; ANDRADE, 2018). Acredito que os resultados aqui apresentados irão contribuir para a melhoria do exame, para o aperfeiçoamento dos processos de ensino e aprendizagem, bem como para (re)adequação da implementação da política linguística da OACI no contexto investigado. No viés político, endosso a proposta de Chalhoub-Deville (2009, 2016) sobre a necessidade de uma visão mais proativa na investigação do impacto, na qual elaboradores de exames, juntamente com usuários e, principalmente, pesquisadores, devem ter mais espaço e voz na formulação e reformulação de políticas linguísticas, não apenas para garantir melhores ações, mas também para avaliar, com maior precisão, os impactos potenciais, direcionando, assim, medidas mitigadoras para minimizar possíveis efeitos negativos.

2.9 Resumo do capítulo

Neste capítulo, busquei retratar as principais contribuições de estudos sobre efeito retroativo, dando ênfase àqueles conduzidos a partir da publicação do artigo seminal de Alderson e Wall (1993) que discute questões relacionadas à existência do fenômeno. Apresentei os modelos teóricos para direcionamento de pesquisas sobre o assunto propostos por Alderson e Wall (1993), por Hughes (1993) e Bailey (1996), por Watanabe (2004a), Stoneman (2006) e Green (2007a) e argumentei que eles ainda não exauriram por completo toda a complexidade envolta do fenômeno e o entendimento sobre as relações entre os fatores que interferem na sua determinação. Discuti alguns estudos cujos resultados colaboraram para a construção de uma teoria sobre o efeito retroativo e mostrei que, em comparação com as pesquisas com foco no professor, ainda pouco conhecemos sobre como o efeito atua nos alunos, mesmo sendo eles um dos principais envolvidos. Dado que a presença de um exame desencadeia uma série de processos, abordei os tipos de preparação em que alunos e professores se engajam visando ao sucesso em um teste e sua natureza. Por fim, defendi que estudos sobre efeito retroativo devem fazer parte do processo de validação de um teste, recaindo, portanto, sobre todos os envolvidos a responsabilidade de coletar evidências dessa natureza.

CAPÍTULO III

A METODOLOGIA DA INVESTIGAÇÃO

3.1 Considerações iniciais

Neste capítulo, apresento a pesquisa de métodos mistos como a abordagem metodológica escolhida para investigar o efeito retroativo do EPLIS nas percepções, atitudes e ações de professores e alunos. Discuto, primeiramente, os paradigmas metodológicos tradicionais na investigação do fenômeno e pontuo sobre a necessidade de empregar métodos qualitativos e quantitativos para atender satisfatoriamente meus objetivos de pesquisa. Em seguida, abordo o surgimento da pesquisa de métodos mistos, considerada, atualmente, a terceira maior corrente de pesquisa, bem como sua base filosófica e principais propósitos, e proponho um desenho metodológico para a coleta e análise de dados neste estudo. Passo, então, a tratar das características do contexto estudado. Apresento o ensino de língua inglesa na formação do controlador de tráfego aéreo, os participantes da pesquisa e as especificações do EPLIS. Por fim, detalho as etapas, os instrumentos de geração de dados utilizados e os procedimentos adotados para a coleta e análise dos dados.

3.2 Os paradigmas metodológicos nos estudos sobre efeito retroativo

Por três décadas, mais precisamente de 1950 a 1980, a corrente positivista dominou a concepção e o desenvolvimento de pesquisas na área de avaliação de línguas, em razão da forte influência das teorias educacionais da época que valorizavam métodos psicométricos na investigação da validade, confiabilidade e praticidade de testes (MCNAMARA, 2001; STONEMAN, 2006; WINKE, 2011). Nesse paradigma de pesquisa, é defendida a objetividade na investigação de hipóteses e/ou teorias geradas previamente, em um processo descendente (*top down*), no qual a realidade é tida como estável e isenta de qualquer envolvimento entre o pesquisador e os sujeitos da pesquisa, a fim de assegurar a neutralidade e imparcialidade no estudo. Os resultados dessas investigações são julgados independentemente do momento e do contexto e, por isso, podem ser generalizáveis. O efeito retroativo, nesse período, tinha caráter determinista, conforme discutido no Capítulo II, ou seja, acreditava-se que um bom teste

produziria efeitos retroativos positivos, enquanto um teste ruim prejudicaria o ensino e a aprendizagem (ALDERSON; WALL, 1993; SCARAMUCCI, 2004).

Todavia, a partir de 1980, esse cenário de pesquisa começou a mudar profundamente por grande influência dos trabalhos de Hymes (1972) e Canale e Swain (1980), que expandiram o construto de competência linguística, a fim de contemplar, além do conhecimento linguístico, as competências sociolinguística e estratégica, conforme colocam Scaramucci (2000), McNamara (2001) e Jang et al (2014). Em paralelo, a avaliação passou a ser reconhecida como uma prática social, permeada de valores, e não mais como evento isolado, estável, previsível e imune às questões sociais e contextuais (SHOHAMY, 1993; MCNAMARA, 2001; INBAR-LOURIE, 2008b). A visão determinista sobre o efeito retroativo de um teste no ensino e aprendizagem foi substituída, então, por uma abordagem construtivista que reconhece a complexidade do fenômeno e a diversidade de fatores individuais, sociais, contextuais e educacionais que podem interagir na sua determinação.

Os teóricos da abordagem construtivista, ao contrário dos positivistas, contestam a existência de uma realidade única e universal. Para eles, as realidades são múltiplas, dinâmicas e construídas no tempo, no espaço e na interação entre o pesquisador e os sujeitos. O pesquisador observa o fenômeno e as explicações são geradas, indutivamente, a partir dos dados, seguindo o modelo ascendente (*bottom up*). Nessa visão de mundo, é impossível dissociar o olhar do pesquisador do fenômeno estudado e, por essa razão, os resultados são sempre holísticos, carregados de valores e, portanto, dependentes do contexto (JOHNSON; ONWUEGBUZIE, 2004).

Juntamente com essas mudanças epistemológicas, a proposta de Messick (1989, 1996) de unificação do conceito de validade, na qual as consequências sociais de um exame aparecem como uma de suas facetas e o efeito retroativo como sendo um aspecto dentro da validade de consequência, impulsionou o crescimento de pesquisas de natureza construtivista sobre o assunto. Estudos mais exploratórios sobre efeito retroativo começam, portanto, a florescer e a ganhar espaço na área de avaliação de línguas, com o objetivo de descrever a complexidade do fenômeno, seu mecanismo de funcionamento e impacto naqueles envolvidos no processo, com destaque para os professores como foco principal dessas investigações.

No entanto, Alderson e Wall (1993) reportaram deficiências nesses estudos no que se refere a dados de observações de aulas. Segundo eles, essas pesquisas apoiavam-se pesadamente em entrevistas e questionários para analisar as percepções e atitudes dos participantes, sem confrontá-las com a realidade da sala de aula.

Mais tarde, Xie e Andrews (2012) fizeram outra crítica às pesquisas sobre efeito retroativo que utilizavam somente métodos qualitativos para a geração e análise de dados. Embora esses estudos tenham sido capazes de identificar múltiplos fatores atuantes na determinação do efeito retroativo, segundo esses autores, a falta de dados quantitativos impedia que fossem traçadas relações entre esses elementos.

Esta pesquisa busca suprir essas necessidades apresentadas por Xie e Andrews (2012). Seu objetivo é compreender os fatores que interagem na determinação do efeito retroativo do EPLIS e suas relações, por meio da observação e da descrição detalhada do cenário no qual esse fenômeno atua, buscando conhecer a perspectiva dos envolvidos, suas crenças, seus valores e suas razões para agirem da forma como agem. A intenção é investigar o que eles fazem e o porquê por detrás de suas ações e, assim, oferecer subsídios, provenientes de um contexto específico de investigação, para a consolidação da teoria sobre o fenômeno.

Para atender satisfatoriamente a esse propósito, empreguei, portanto, métodos de coleta e análise de dados qualitativos e quantitativos. Os métodos quantitativos possibilitaram a identificação de padrões e tendências nas percepções e atitudes de alunos e professores, uma vez que, influenciados pelo exame, as respostas desses participantes poderiam apresentar certos modelos (BURROWS, 2004). Explorei, também, as relações de significância entre o efeito retroativo, o nível de conhecimento de alunos sobre o exame, seus níveis declarados de proficiência e o quão próximos eles estavam da data de realização do teste. Os métodos qualitativos, por sua vez, expandiram, refinaram e confirmaram os resultados quantitativos, a fim de que fosse obtido, assim, um panorama mais completo de como o efeito retroativo do exame interage com as forças presentes no contexto escolar.

Os métodos quantitativos e qualitativos são vistos, portanto, como métodos complementares, na busca de uma compreensão mais rica das relações entre as diferentes facetas que envolvem um fenômeno tão complexo como o efeito retroativo. Outra opção metodológica que não contemplasse a combinação de dados quantitativos e qualitativos se apresentaria falha para atender os propósitos do estudo.

A escolha por métodos mistos na análise do efeito retroativo é considerada inovadora do ponto de vista metodológico, principalmente ao tomar o contexto brasileiro. No Brasil, as pesquisas em avaliação, em geral, são voltadas à coleta e à análise qualitativa dos dados. Scaramucci (1995) é uma das poucas pesquisadoras brasileiras que quebrou essa tradição ao mostrar a possibilidade de combinação de métodos qualitativos e quantitativos em um estudo de natureza exploratória, ainda que, na época, a denominação “pesquisa de métodos mistos” não fosse empregada para caracterizar tais investigações.

A pesquisa de métodos mistos tem sido bastante aceita entre os estudiosos da área de avaliação de línguas como uma metodologia legítima de investigação. Segundo Jang et al (2014), nos últimos 10 anos, constata-se uma tendência crescente no número de estudos que adotam essa visão pluralista, com o objetivo de buscar novos conhecimentos sobre a complexidade dos fenômenos que envolvem as práticas avaliativas.

3.3 O surgimento da pesquisa de métodos mistos

A pesquisa de métodos mistos (PMM doravante), em inglês *mixed methods research*, tem sido considerada o terceiro principal paradigma de pesquisa (JOHNSON et al, 2007), a lado das pesquisas quantitativa e qualitativa. Seu início é datado entre o final dos anos de 1980 e o começo dos anos de 1990, quando diversos teóricos começaram a escrever, de maneira sistemática, sobre uma abordagem de pesquisa distinta, que fosse capaz de integrar dados quantitativos e qualitativos e de se estabelecer como um modelo próprio de investigação.

Apesar de ser recente o movimento para a criação dessa abordagem, o interesse em empregar, intencionalmente, métodos quantitativos e qualitativos dentro de um mesmo estudo tem sido percebido há, pelo menos, 60 anos nas Ciências Sociais. Todavia, esses primeiros estudos são difíceis de serem localizados na literatura devido ao fato de não existir, até então, um termo em comum pelo qual eles pudessem ser identificados. Durante esse período, diversos nomes foram usados, tais como pesquisa integrada, métodos quantitativos e qualitativos, pesquisa híbrida, triangulação entre métodos, metodologia mista, entre outros (CRESWELL; CLARK, 2011).

Esses estudos eram marcados pelo rompimento com a forte associação, tradicionalmente existente entre "paradigma" e "método". Scaramucci (1995) explica que os métodos quantitativos e qualitativos, por terem estado, por muito tempo, tão intimamente conectados aos paradigmas positivista e interpretativista, respectivamente, passaram a designar posicionamentos filosóficos distintos. Assim, retrata:

> o método é de tal maneira ligado a um paradigma que ao se escolher um paradigma se está inevitavelmente definindo o método. Na prática, identifica-se uma associação entre método quantitativo e paradigma positivista […], por um lado, e método qualitativo e paradigma interpretativista de outro […]. Como os dois paradigmas são considerados mutualmente exclusivos, mostrando visões de mundo consideradas antagonistas, tem-se a impressão de que uma reconciliação entre dois métodos de paradigmas opostos em uma mesma pesquisa se torna impossível (SCARAMUCCI, 1995, p. 96).

Outro ponto que reforça a dicotomia qualitativo/quantitativo é a ideia de que as características ou atributos de um paradigma de pesquisa são fixas, excluindo, assim, a possibilidade de se utilizar certo método. Nessa perspectiva, a natureza exploratória de um estudo, por exemplo, seria vista como fator impeditivo na adoção de técnicas estatísticas que visassem computar a frequência com que determinado evento ocorre. Allwright e Bailey (1991), assim como Scaramucci (1995), são contrárias a esse posicionamento e defendem que são as perguntas de pesquisa que devem prevalecer na escolha dos métodos de coleta e análise de dados. Em relação às salas de aulas, contextos nos quais estudos interpretativistas predominam, Allwright e Bailey (1991) ressaltam que a combinação de métodos é bem-vinda sempre que o pesquisador entender que dados mistos atendem melhor o seu problema de pesquisa. Nas suas palavras: "nós valorizamos as investigações que combinam elementos objetivos e subjetivos, que quantificam o que pode ser beneficamente quantificado, e que utilizam coleta de dados e procedimentos de análise qualitativos sempre que conveniente" (ALLWRIGHT; BAILEY, 1991, p. 67, tradução nossa).[38]

Os conceitos de objetividade e subjetividade a que se referem Allwright e Bailey (1991), no trecho acima, são também marcas desse conflito paradigmático em que foram revestidos os métodos quantitativo e qualitativo. A concepção tradicional de ciência, na qual o conhecimento científico é visto como objetivo, imparcial, neutro, linear, generalizável e universal, impedia o uso de métodos que não fossem de natureza matemática. Nessa linha de pensamento, os estudos de base qualitativa eram considerados subjetivos no sentido de que a relação entre pesquisador e objeto poderia produzir resultados tendenciosos e parciais, carecendo, portanto, de rigor científico. A esses dois conceitos também estão embutidos, tradicionalmente, julgamentos de valor, no sentido de que a objetividade é tida como positiva e confiável, enquanto a subjetividade indica algo instável e, portanto, deve ser evitada. No trecho acima, essa conotação positiva e negativa não existe. Allwright e Bailey (1991) usam os termos objetivos e subjetivos para fazerem referência a elementos que podem ser quantificáveis e elementos que requerem um grau maior de inferência do pesquisador.

É a visão pluralista e complementar de pesquisadores como Allwright e Bailey (1991) e Scaramucci (1995) que fomentou a criação de um projeto metodológico que abrangesse estudos que se colocavam ao longo desse contínuo polarizado em quantitativo e

[38] No original: "We see most value in investigations that combine objective and subjective elements, that quantify only what can be usefully quantified, and that utilise qualitative data collection and analysis procedures whenever they are appropriate".

qualitativo, mostrando as inúmeras possibilidades de combinação de métodos para atender adequadamente a diferentes propósitos de pesquisa.

No entanto, para que a PMM se estabelecesse como uma abordagem metodológica distinta das tradicionais faltava a adoção de um princípio filosófico que norteasse toda a concepção do estudo. O pragmatismo apresentou-se como grande parceiro e aliado. A doutrina pragmática compreende a realidade como singular e múltipla. Para essa corrente de pensamento, a relevância está em entender os problemas de ordem prática, empregando, para isso, os métodos mais adequados, estejam eles tipicamente associados à pesquisa quantitativa ou qualitativa. O pragmatismo abandona, portanto, a dicotomia entre o positivismo e o interpretativismo e se coloca como paradigma prático e aplicado, no qual as perguntas de pesquisa são os fatores decisivos na escolha dos métodos.

Essa nova abordagem de pesquisa, porém, não tem a pretensão de substituir estudos puramente quantitativos e qualitativos, mas fornece os meios para responder questões de pesquisa que, em razão de sua natureza, ficariam incompletas e falhas se somente um método fosse adotado (JOHNSON; ONWUEGBUZIE, 2004; CRESWELL; CLARK, 2011). Muitos pesquisadores nas Ciências Sociais e na Educação passaram a adotar a PMM, argumentando que um problema de pesquisa não precisaria, necessariamente, ser estudado sob o viés de uma metodologia em detrimento de outra, mas pelo emprego de múltiplos métodos (TERREL, 2012).

O termo *mixed methods research* tem sido amplamente utilizado nessas áreas para descrever estudos caracterizados pelo posicionamento pragmático do pesquisador na escolha dos métodos de coleta e análise mais adequados aos seus propósitos como, por exemplo, as pesquisas de Allwright e Bailey (1991) e Scaramucci (1995). Pode-se, então, dizer que, com o advento da PMM como abordagem metodológica, o conflito paradigmático, polarizado e incompatível com o mundo contemporâneo em que vivemos foi, enfim, superado.

Este estudo fundamenta-se, portanto, na visão de mundo múltipla, prática e aplicada, na qual a escolha dos métodos partem da natureza das perguntas de pesquisa, dissociando-se da visão dualista tradicional em que métodos quantitativos pertencem ao paradigma positivista e métodos qualitativos ao paradigma construtivista, assim como seus atributos.

3.3.1 A definição da PMM e seus propósitos

A fim de ter uma visão mais clara sobre as razões que levaram diversos pesquisadores a optarem pela combinação de métodos, Greene et al (1989) revisaram 57 estudos de natureza mista e identificaram cinco propósitos principais: triangulação (*triangulation*), complementaridade (*complementarity*), desenvolvimento (*development*), iniciação (*initiation*) e expansão (*expansion*).

A triangulação refere-se ao emprego de métodos quantitativos e qualitativos para analisar um mesmo fenômeno O objetivo é a convergência e a corroboração dos resultados. Esse termo também é encontrado em pesquisas de natureza exclusivamente interpretativista para designar a análise qualitativa de dados obtidos por diversas fontes como, por exemplo, diários, entrevistas, observações, protocolos verbais. Nesses estudos, os dados são triangulados com o objetivo de aumentar a validade e a confiabilidade dos resultados encontrados.

A complementaridade, diferentemente da triangulação, busca analisar as diversas facetas de um fenômeno, a fim de alcançar uma compreensão mais rica e elaborada. Enquanto a triangulação busca a corroboração dos resultados, a complementaridade busca aprimorar e clarificar os resultados obtidos inicialmente com os resultados obtidos no emprego do segundo método.

No propósito de desenvolvimento, os resultados obtidos com o emprego do primeiro método irão trazer informações para o desenvolvimento do segundo método. Já no caso da iniciação, o uso de métodos mistos leva à descoberta de inconsistências e contradições, sendo necessária a reformulação das perguntas de pesquisa. Por fim, o propósito de expansão refere-se à necessidade de ampliar o escopo e o alcance da investigação.

Na revisão de Greene et al (1989), os propósitos da complementaridade e da expansão foram os mais comuns em termos de justificativa para o emprego de métodos mistos. Porém, podemos notar que esses dois objetivos se sobrepõem, uma vez que ambos envolvem o uso de um segundo método para ampliar os resultados obtidos inicialmente. Essa é, também, a opinião de Jang et al (2014), que propõem apenas quatro propósitos para os estudos mistos: triangulação, complementaridade, desenvolvimento e iniciação. Isso, porém, não quer dizer que todos os estudos considerados mistos devem apresentar todos esses objetivos em conjunto. Pelo contrário, para Jang et al (2014), a complementaridade e triangulação são os que mais se destacam.

A notoriedade desses dois propósitos pode ser confirmada na definição de PMM proposta por Johnson et al (2007). Ainda sem uma definição que englobasse todos os critérios

considerados essenciais para caracterizar os estudos de natureza mista, Johnson et al (2007) pediram aos pesquisadores mais influentes na área de metodologia que lhe enviassem uma sugestão. Com base na análise dos textos recebidos, a PPM foi definida como:

> um tipo de pesquisa na qual o pesquisador ou grupo de pesquisadores combinam elementos das abordagens de pesquisa qualitativa e quantitativa (ex.: uso de pontos de vista qualitativos e quantitativos, coleta de dados, análises e técnicas de inferência) para propósitos gerais de ampliação e aprofundamento da compreensão e corroboração [de uma hipótese][39] (JOHNSON et al, 2007, p. 123, tradução nossa).

Conforme podemos observar na citação acima, Jonhson et al (2007) destacam como justificativas da combinação de métodos a necessidade do pesquisador de ampliar, aprofundar e corroborar seus resultados. A ampliação e o aprofundamento definem o propósito da complementaridade, enquanto a corroboração refere-se à triangulação.

Creswell e Clark (2011) concordam com a definição proposta por Johnson et al (2007), uma vez que ela enfatiza a posição da PMM como uma metodologia distinta, e resumem as principais características desse novo paradigma de pesquisa em: a) a coleta e análise de dados qualitativos e quantitativos; b) a combinação das diferentes fontes de dados; e c) a opção de priorizar um conjunto de dados (qualitativos ou quantitativos) ou de tratar ambos igualmente, em um mesmo estudo ou em programa de estudos com múltiplas fases.

É interessante notar que o fato de a PMM permitir que um método seja priorizado em relação ao outro faz com que se abra ao pesquisador um leque de possibilidades de combinação e interação entre os dados. Anseio esse demonstrado por diversos teóricos que viam no dualismo paradigmático o engessamento das práticas de investigação. Os modelos mais frequentes de PMM serão apresentados a seguir.

3.3.2 Modelos de PMM e suas principais características

Para definir o desenho metodológico que será utilizado em determinada pesquisa, Creswell e Clark (2011) argumentam que decisões cruciais devem ser tomadas considerando o nível de interação, o grau de prioridade, a relação temporal e o ponto de interface entre os métodos quantitativo e qualitativo.

---

[39] No original: "Mixed methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration".

Em relação ao nível de interação, os métodos quantitativo e qualitativo podem ser implementados independentemente ou podem interagir ao longo do estudo. Quando os métodos são independentes, os dados são coletados e analisados separadamente. A combinação dos dados acontece somente na etapa final da pesquisa, no momento de interpretação dos resultados. Quando existe uma interação entre os métodos, o desenho ou o andamento de uma fase pode depender dos resultados gerados na fase anterior e a análise dos dados coletados em cada vertente é feita conjuntamente.

O grau de prioridade refere-se ao peso e à importância dados aos diferentes métodos para responder às perguntas de pesquisa. O pesquisador pode atribuir o mesmo grau de relevância aos dois métodos, priorizar o método quantitativo ou enfatizar o método qualitativo. Quanto à relação temporal entre os métodos quantitativo e qualitativo dentro de um estudo, o pesquisador pode optar por utilizá-los concomitantemente, sequencialmente ou em combinação (quando o estudo apresenta múltiplas fases).

A integração dos dados obtidos pela combinação de métodos pode ocorrer em diversos momentos da pesquisa. Na coleta, por meio da técnica chamada conexão (*connecting*), a geração do segundo conjunto de informações depende do primeiro. Quando os elementos são coletados separadamente e independentemente, pode-se empregar a técnica de fusão (*merging*) durante o processo de análise. A integração pode também ocorrer somente na interpretação dos resultados do estudo. E, ainda é possível optar pela combinação somente na conclusão da pesquisa, quando se comparam os resultados alcançados.

Creswell e Clark (2011) e Terrel (2012) identificaram, pelo menos, seis tipos principais de pesquisas de métodos mistos, resumidos no Quadro 2. O "desenho paralelo convergente" é caracterizado pela coleta de dados quantitativos e qualitativos concomitantemente e pela análise independente dos dados. Os dois métodos recebem o mesmo grau de importância, e os resultados obtidos são combinados somente na fase de conclusão ou interpretação do estudo. O propósito da combinação entre métodos nesse modelo é confirmar ou corroborar os dados (triangulação). No "desenho sequencial explicativo", os métodos são aplicados em sequência, iniciando com a coleta dos dados quantitativos e, em seguida, dos qualitativos. Os dados qualitativos são usados para explicar os resultados iniciais (quantitativos) com maior profundidade e riqueza de detalhes, atendendo, assim, ao propósito de complementaridade. Nesse modelo, ambas as fases são tratadas de forma igualitária, mas é possível priorizar a fase qualitativa, e a essa variante é dada o nome de "desenho preliminar quantitativo". No "desenho sequencial exploratório" os dados também são coletados sequencialmente, porém inicia-se com a coleta de dados qualitativos. Após a análise desses

dados, segue-se para a fase quantitativa com o objetivo de testar as hipóteses levantadas ou generalizar os resultados iniciais (desenvolvimento). Nesse modelo, ambas as fases recebem a mesma consideração ou prioriza-se uma delas.

Quadro 2 − Modelos de pesquisa de métodos mistos

Tipos de desenho	Grau de interação	Grau de prioridade	Relação temporal	Ponto de interface
Paralelo convergente	independente	ênfase igual	concomitante	na interpretação ou conclusão
Sequencial explicativo	interativo	ênfase igual (ou prioridade pode ser dada a uma fase)	sequencial (quantitativo primeiro)	na coleta ou análise de dados
Sequencial exploratório	interativo	ênfase igual (ou prioridade pode ser dada a uma fase)	sequencial (qualitativo primeiro)	na coleta ou análise dados
Incorporado	interativo	ênfase na fase quantitativa ou qualitativa, dependendo do foco do estudo	concomitante ou sequencial	na análise dos dados
Multifásico	interativo	ênfase igual, ênfase na fase quantitativa ou na fase qualitativa	concomitante, sequencial ou ambos	em todo o desenho da pesquisa
Emancipatório	interativo	ênfase igual, ênfase na fase quantitativa ou na fase qualitativa	concomitante ou sequencial	em todo o desenho da pesquisa

Existe, também, a possibilidade de utilizar um desenho de pesquisa tradicional, quantitativo ou qualitativo, e incluir dados qualitativos ou quantitativos dentro do estudo. Esse modelo é conhecido por "desenho incorporado" (*embedded design*). Nele, o pesquisador pode incluir dados qualitativos dentro de um experimento ou incluir dados quantitativos em um estudo etnográfico, por exemplo. O método incorporado recebe menos proeminência e os dados são combinados somente na análise dos resultados. O propósito desse modelo é ganhar uma perspectiva maior sobre o fenômeno (complementaridade). Nas pesquisas compostas por múltiplas fases, a coleta de dados pode acontecer tanto em sequência ou concomitantemente e os diferentes conjuntos de dados são tratados igualmente.

Por fim, alguns pesquisadores adotam a pesquisa de métodos mistos em estudos que dão voz a grupos marginalizados socialmente e excluídos dos círculos de poder (mulheres, minorias étnicas, religiosas, raciais e culturais, pessoas com deficiências, grupos LGBT, pessoas com baixo poder aquisitivo, entre outros grupos) (JANG et al, 2014). Nesse caso, o que define esses tipos de pesquisa são os seus propósitos emancipatórios e de transformação social, ao invés das características de como foi feita a combinação entre os métodos quantitativos e qualitativos empregados no estudo.

3.4 O desenho metodológico da pesquisa

Para investigar o efeito retroativo do EPLIS nas percepções, atitudes e ações de alunos e professores, adotei a pesquisa de métodos mistos dentro de um desenho sequencial explicativo, dividido em duas fases, sendo a primeira quantitativa e a segunda qualitativa, com o propósito de complementaridade e de triangulação dos dados. O efeito retroativo é considerado um fenômeno complexo de natureza altamente variável e dependente do contexto. A combinação de métodos mostrou-se a mais adequada para identificar as variáveis significativas na influência do exame no contexto escolar e explicar como elas se relacionam. Essa decisão encontra respaldo nas palavras de Creswell e Clark (2011):

> resultados quantitativos podem tecer explicações para as relações entre as variáveis, porém falta uma compreensão mais detalhada sobre o que os testes estatísticos realmente querem dizer. Dados e resultados qualitativos podem ajudar a construir esse conhecimento[40] (CRESWELL; CLARK, 2011, p. 9, tradução nossa).

Primeiramente, conduzi um estudo preliminar que contemplou a análise de documentos, aplicação de questionários curtos e visitas técnicas. Ressalto que o estudo preliminar não deve ser compreendido como um estudo piloto pelo fato de não se basear na aplicação, em menor escala, dos procedimentos e métodos propostos na pesquisa, mas, sim, na coleta de informações para a concepção da pesquisa. Dessa forma, a análise minuciosa de documentos que embasaram o desenvolvimento do EPLIS ajudou a traçar as características da linguagem avaliada pelo exame e os efeitos pretendidos com a sua implementação. O questionário inicial propiciou a coleta de dados preliminares sobre a influência do exame nas percepções de alunos e professores. As visitas técnicas, por sua vez, possibilitaram conversas informais com professores, coordenadora pedagógica, adjunto e chefe da Subseção de Idiomas da escola e alunos, além de observações de aulas não estruturadas. Esse estudo preliminar é considerado de grande importância, pois me permitiu, além de um melhor entendimento sobre o cenário de pesquisa, escolher e construir instrumentos de coleta de dados mais apropriados e situados para o estudo.

A primeira fase da pesquisa objetivou-se a investigar padrões e tendências entre os participantes em relação ao efeito retroativo do EPLIS em diversos aspectos do ensino e aprendizagem de inglês na EEAR. Por meio de um questionário, dados quantitativos sobre as

[40] No original: "Quantitative results can net general explanations for the relationships among variables, but the more detailed understanding of what the statistical tests or effect sizes actually mean is lacking. Qualitative data and results can help build that understanding".

percepções e atitudes de professores e alunos foram coletados. Em relação aos dados referentes aos professores, foi realizada uma análise estatística descritiva. Os resultados foram, então, confrontados com dados iniciais coletados durante estudo preliminar, a fim de investigar possível aumento na intensidade do efeito retroativo, ocasionado pela implementação do EPLIS no contexto escolar e pela melhoria em sua divulgação. As respostas dos alunos ao questionário foram analisadas, primeiramente, por meio da análise estatística descritiva. Em seguida, aplicou-se a técnica de análise fatorial, por meio do método de componentes principais, que consiste na redução de um grande número de variáveis (representadas pelos itens no questionário) a um número menor de elementos facilitando, assim, a interpretação dos dados. Foram, então, investigadas as relações de (in)significância entre os componentes extraídos e três fatores: o conhecimento dos alunos sobre o exame, o nível de proficiência linguística e o quão próximo eles estavam da realização do teste.

A segunda fase visou à coleta de dados qualitativos por meio de comentários escritos deixados pelos participantes no preenchimento do questionário, grupo focal com alunos, entrevistas individuais com chefia, coordenação, professores e instrutor e observações de aulas de três professores e um instrutor. Essa fase teve como objetivo explicar, expandir, aprofundar e corroborar os resultados quantitativos para uma compreensão mais rica sobre como atua o efeito retroativo do EPLIS no cenário investigado. Para isso, os dados do grupo focal e das entrevistas foram transcritos e, juntamente com os comentários, foram submetidos à análise de conteúdo. Os dados foram, então, rotulados utilizando-se as categorias extraídas na análise de componentes principais conduzida na fase I do estudo. As observações de aulas foram registradas por vídeo ou áudio de acordo com a opção escolhida pelo professor ou instrutor participante. Utilizou-se uma ficha de observação com categorias que representavam áreas que poderiam ser, potencialmente, afetadas pelo exame. Os dados da observação foram confrontados com entrevistas posteriores para fins de confiabilidade.

O Quadro 3 fornece um panorama das etapas desta pesquisa, seus objetivos, instrumentos utilizados e tipo de análise conduzida. O cenário da investigação, assim como os participantes, o exame, a escolha dos instrumentos, os procedimentos de coleta e a análise adotados serão discutidos detalhadamente nas seções 3.5, 3.7 e 3.8.

Informo que este estudo foi, devidamente, aprovado pelo Comitê de Ética em Pesquisa da Unicamp em 03 de março de 2016, por meio do parecer 1.436.512 (Anexo A), sob o número de protocolo 52971216.9.00000.5404.

Quadro 3 − Panorama das etapas da pesquisa

Etapas	Objetivos	Instrumentos	Análise
Estudo preliminar	Analisar a natureza da linguagem avaliada pelo EPLIS e os efeitos retroativos pretendidos; Coletar dados inicias sobre as percepções e atitudes de professores e alunos em relação ao EPLIS Construir os instrumentos de coleta de dados.	Doc 9835 AN/453 Especificações do exame Manual do candidato Questionário inicial Visitas técnicas	Análise documental e descritiva
Fase I Quantitativa	Identificar mudanças na intensidade do efeito retroativo nas percepções e atitudes dos professores após implementação do exame; Investigar o efeito retroativo nas percepções e atitudes dos alunos em função dos seguintes fatores: conhecimento geral sobre o EPLIS, nível de proficiência declarada e proximidade do exame. Comparar a influência do exame nas percepções e atitudes de professores e alunos.	Questionários fechados	Análise estatística descritiva Análise fatorial de componentes principais Testes estatísticos de significância
Fase II Qualitativa	Ampliar, expandir e aprofundar os resultados quantitativos; Investigar o efeito retroativo do EPLIS nas ações de alunos e professores; Analisar o tipo de preparação para o exame desenvolvida em sala de aula.	Comentários escritos por alunos e professores Grupo focal com alunos Entrevistas individuais com professores, instrutor, coordenadora e chefe Observações de aulas	Análise de Conteúdo

3.5 O cenário da pesquisa

O estudo foi conduzido na Escola de Especialistas de Aeronáutica (EEAR doravante), localizada na cidade de Guaratinguetá, interior do estado de São Paulo. Essa instituição pertence à Força Aérea Brasileira, responsável pela formação dos controladores de tráfego aéreo brasileiros e principal empregadora nessa área. Participaram da pesquisa coordenação pedagógica, chefia da Subseção de Idiomas, professores, instrutores de inglês aeronáutico e alunos das quatro séries do Curso Básico de Controle de Tráfego Aéreo dessa escola.

3.5.1 O Curso Básico em Controle de Tráfego Aéreo

O Curso Básico em Controle de Tráfego Aéreo da EEAR (BCT doravante) possui a duração de dois anos, divididos em quatro séries, sendo ministrado em regime de internato. O currículo do curso abrange disciplinas em três diferentes campos: no campo Geral, são oferecidos conteúdos básicos comuns com o objetivo de nivelar o conhecimento dos alunos oriundos de diversas regiões do país; no Campo Militar, os alunos recebem formação nos postulados do militarismo; e no Campo Técnico, é oferecida instrução específica para que o futuro controlador de tráfego aéreo exerça as atribuições de sua especialidade.[41] É importante salientar aqui que a disciplina de Língua Inglesa para o curso BCT pertence ao Campo Técnico, ou seja, o conteúdo ministrado deve estar relacionado ao exercício das funções do especialista. Outras disciplinas que pertencem ao Campo Técnico são: Fundamentos de Voo, Meteorologia, Aeronaves, Regras Básicas de Tráfego Aéreo, Fraseologia etc.

O ingresso na EEAR é feito por meio de concurso público de provas. Os candidatos devem possuir o 2º grau de escolaridade completo e ter entre 17 e 24 anos de idade. Além das provas escritas, os candidatos devem passar por uma inspeção de saúde, e serem aprovados em um exame de aptidão psicológica e em um teste de condicionamento físico. O concurso para a especialidade BCT acontece semestralmente, com o provimento de 128 vagas,[42] totalizando 256 vagas no ano. Após a conclusão do curso de formação, esses profissionais são promovidos à Graduação de Terceiro Sargento da Aeronáutica e designados a atuar nos diversos órgãos de controle de tráfego aéreo espalhados no território nacional, conforme necessidade do serviço e classificação no curso.

A escolha em conduzir este estudo na EEAR deu-se pelo grande número de controladores de tráfego aéreo formados nessa instituição, além da regularidade de concursos para essa especialidade. Cerca de 80% desses profissionais recebem sua formação na EEAR. Os 20% restante são capacitados pelo ICEA e pertencem a quadros de carreira de natureza civil no âmbito do Comando da Aeronáutica ou na Empresa Brasileira de Infraestrutura Aeroportuária (Infraero).

---

[41] Informações obtidas no sítio eletrônico da EEAR. Disponível em: < http://www2.fab.mil.br/eear/>. Acesso em: 20 maio 2017.

[42] Dado obtido no Edital de Concurso – Instruções Específicas para o Exame de Admissão ao Curso de Formação de Sargentos da Aeronáutica para o Primeiro Semestre do ano de 2018, aprovado pela Portaria DEPENS Nº587-T/DPL, de 14 de dezembro de 2016. Disponível em <https://www.acheconcursos.com.br/edital-concurso/edital-concurso-aeronautica-2016-cfs-2018-sargento-de-trafego-aereo> Acesso em: 20 abr. 2017.

3.5.2 Os participantes da pesquisa

Atualmente, a EEAR possui um corpo docente permanente de Inglês, formado por professores civis e militares, que atua no curso BCT. Os professores civis foram classificados por meio de concurso público de provas e títulos. Os professores militares são do Corpo de Oficiais Convocados selecionados por meio de análise curricular. Todos possuem, como qualificação mínima, o diploma de Letras com licenciatura em língua inglesa.

Devido ao grande volume de aulas, a EEAR também conta com um corpo docente volante de instrutores de inglês aeronáutico. Esse grupo é formado, principalmente, por controladores de tráfego aéreo em serviço que obtiveram níveis de proficiência 5 ou 6 no EPLIS.[43] Em relação à qualificação na área de ensino de línguas, a maioria desses profissionais realizou um treinamento básico voltado para inglês aeronáutico. Essa capacitação tem duração de 3 semanas e é considerada equivalente a um curso TEFL (*Teaching English as a Foreign Language*). Por não terem licenciatura em língua inglesa, esses profissionais são denominados "instrutores", diferenciando-os, assim, do grupo de professores.

Para o estudo principal, foram convidados a participar os professores e instrutores de inglês aeronáutico que atuaram no curso BCT, no ano de 2016[44], totalizando 21 profissionais, 17 professores e 04 instrutores de inglês aeronáutico. Desses, 16 responderam ao questionário. Três professores e um instrutor tiveram suas aulas observadas no ano seguinte. Em relação ao corpo discente, a EEAR contou com 407 alunos BCT, no ano de 2016, distribuídos em quatro séries. Como mencionado anteriormente, para serem admitidos na EEAR, os alunos devem ter concluído com aproveitamento o ensino médio e serem aprovados em um processo seletivo. A faixa etária desses alunos varia de 17 a 24 anos. 387 alunos responderam ao questionário e 12 alunos (três alunos por série) participaram do grupo focal.

3.5.3 As disciplinas de língua inglesa no curso BCT

A fim de cumprir as recomendações da OACI em relação à proficiência de controladores de tráfego aéreo e operadores de estação aeronáutica, o Comando da Aeronáutica publicou, em 2008, o seu Plano de Implementação dos Requisitos de Proficiência em Inglês

[43] O EPLIS avalia seis níveis de proficiência. O nível 5 é denominado Avançado, e o nível 6, Experto, conforme apresentado no Capítulo I.

[44] Foi feito um estudo preliminar com o grupo de docentes permanente da EEAR no ano de 2014, antes da implementação do EPLIS na escola, para coleta de dados iniciais. Essa etapa da pesquisa será detalhada mais adiante na seção 3.6.

para o período de 2008 a 2010 (BRASIL, 2008), o que pode ser considerado o documento de política linguística da OACI no Brasil. Nesse documento, foram elencadas as ações de capacitação que já vinham sendo realizadas ou aquelas previstas para a elevação da proficiência desses profissionais nas escolas de formação e pós-formação, bem como nos órgãos regionais de controle de tráfego aéreo. Em termos de formação inicial, as ações abrangeram a introdução de uma prova de inglês específica no processo seletivo para a especialidade BCT, a mudança de metodologia de ensino de línguas com foco na abordagem de inglês para fins específicos (ESP), e a ênfase no desenvolvimento da produção e compreensão orais dos alunos na língua alvo (CHINI, 2014; TOSQUI-LUCKS et al, 2016).

Apoiada por esse Plano, em 2008, a EEAR iniciou o processo de reestruturação das disciplinas de inglês da especialidade BCT. Nesse ano, foi incluída uma prova escrita de inglês de nível intermediário no processo de ingresso na especialidade em questão no formato múltipla escolha. A partir de 2009, a EEAR passou a adotar novos materiais didáticos voltados para o ensino de inglês aeronáutico, com ênfase no desenvolvimento das habilidades orais e no ensino comunicativo da língua para propósitos específicos. Atualmente, esse curso de inglês contempla 500 tempos de carga horária (CH), distribuídos em quatro disciplinas: Inglês I, Inglês II, Inglês III e Inglês ATC (*Air Traffic Control*), ministradas nas quatro respectivas séries do curso. As informações aqui apresentadas acerca das disciplinas de inglês foram obtidas no Plano de Unidade Didáticas do curso BCT referente ao ano de 2016. Esse documento contém o conteúdo programático das quatro disciplinas que compõem o curso de inglês, os objetivos de aprendizagem e as orientações metodológicas. Sua elaboração é de responsabilidade dos professores da EEAR, em conjunto com o seu corpo de pedagogos.

É importante pontuar que, antes da reestruturação, o curso de inglês para alunos BCT adotava um conjunto de livros didáticos intitulado *American Language Course*, desenvolvido pelo *Defense Language Institute*, do Departamento de Defesa dos Estados Unidos da América, durante a II Guerra Mundial. Apesar de ter como foco o desenvolvimento das habilidades orais, esse curso não atendia às necessidades do controlador de tráfego aéreo em formação por estar voltado para o inglês geral. Além disso, seguia uma metodologia audiolingual de ensino de línguas, valorizando o vocabulário e a precisão gramatical por meio da repetição (formação de hábitos) e automatização, o que permitiu e incentivou práticas de ensino centradas no professor, consideradas hoje ultrapassadas. Sá (2010) e Chini (2014, 2018) também argumentam que esse curso, mantido por longos 20 anos, não satisfazia às necessidades dos aprendizes, além de contemplar uma metodologia inadequada e conteúdos distantes daqueles vivenciados no ambiente profissional.

3.5.3.1 Inglês I

No currículo da disciplina Inglês I, ministrada para a 1ª série BCT, estão previstos 140 tempos de aula.[45] Essa disciplina tem como foco principal o ensino do inglês geral com a justificativa de oferecer aos alunos base linguística para as disciplinas que a sucedem e que estão voltadas para o ensino de inglês para fins específicos (inglês aeronáutico). Dos 140 tempos de aula, 20 são dedicados a atividades avaliativas. Os objetivos, de acordo com o Plano de Unidades Didáticas, são: a) empregar, oralmente e por escrito, vocabulário e estruturas gramaticais convencionais da língua inglesa; b) praticar a pronúncia correta de termos e palavras da língua inglesa; e c) interpretar textos em inglês geral e de aviação. Esses objetivos parecem denotar uma visão ainda estruturalista de linguagem, com foco no ensino de vocabulário e gramática, bem como na habilidade de leitura de forma isolada, o que contradiz as necessidades de uso da língua no exercício da profissão. Além disso, apesar de constar como cerne do currículo da disciplina o desenvolvimento das habilidades de compreensão e produção orais, nenhum dos objetivos mencionados as abordam, pelo contrário, enfatizam a leitura.

Em termos de conteúdo, o programa contempla temas gerais (descrição física e de personalidade de pessoas, estados e capitais dos EUA, pontos turísticos, autoridades e famosos) e temas relacionados à aviação. Os temas relacionados à aviação são bastante abrangentes e extrapolam o universo da comunicação piloto-controlador. Em uma unidade, por exemplo, são trabalhadas questões relacionadas a refeições a bordo de aeronaves, na qual os alunos devem identificar diferentes tipos de refeições servidas de acordo com a classe do voo; identificar tipos de alimentos não apropriados; e selecionar tipos de refeições especiais por motivo de saúde, crença religiosa e cultural ou por solicitação do passageiro. Esses tópicos são incomuns na comunicação piloto-controlador, denotando divergência com o domínio de uso da língua.

No início da disciplina, é realizado um teste diagnóstico para avaliar a proficiência do aluno, porém não há separação dos alunos em turmas de acordo com o nível de proficiência obtido. É adotado o volume 2 do livro didático *Listen in*, da editoria Cengage, com foco no inglês geral. Os alunos realizam duas provas escritas de caráter somativo, com 40 questões de múltipla escolha cada: 20 questões de compreensão escrita e 20 questões de compreensão oral. É frisado que essas avaliações devem demandar habilidades e capacidades superiores, evitando, assim, práticas de memorização do conteúdo ensinado. Também são aplicadas duas provas orais, porém o desempenho do aluno nessas avaliações não é considerado na somatória das

[44] Cada tempo de aula tem a duração de 45 minutos.

notas atribuídas nas provas escritas. Na prova oral, o professor oferece um *feedback* aos alunos nos cinco dos seis critérios da tabela de níveis da OACI: pronúncia, estrutura, vocabulário, fluência e interação. Não é oferecido *feedback* em relação à compreensão oral pelo entendimento dos professores e coordenação de que essa habilidade já foi avaliada nas provas escritas.

3.5.3.2 Inglês II e III

Para as disciplinas de inglês das 2ª e 3ª séries do curso, estão previstos 120 tempos de aulas em cada, sendo 20 tempos destinados à avaliação. Os objetivos das disciplinas são muito semelhantes aos da 1ª série e estão ligados ao ensino de vocabulário, pronúncia e gramática. Porém, diferentemente do currículo anterior, maior ênfase é dada no desenvolvimento das habilidades orais. São esperados dos alunos da 2ª e 3ª séries produzir gêneros orais (diálogos, exposição de argumentos ou soluções para situações problemas, descrição de imagens e relatos sobre incidentes) e aplicar técnicas para o desenvolvimento da compreensão oral. No entanto, essas técnicas não são especificadas.

Na 2ª série, é adotado o livro *English for Aviation*, da Editora Oxford. A 3ª série utiliza o livro *Aviation English*, da Editora Macmillan. Ambos os livros foram desenvolvidos para contemplar as necessidades do piloto e do controlador de tráfego aéreo dentro da comunicação aeronáutica. Segundo o currículo, essas duas disciplinas estão fundamentadas na abordagem ESP, na qual as necessidades linguísticas reais do aprendiz são fator central na elaboração de um curso (HUTCHINSON; WATERS, 1987). Porém, pelo fato de os materiais utilizados focarem também no piloto, espera-se que haja adaptação por parte do professor e instrutor, uma vez que pilotos e controladores apresentam necessidades que nem sempre se sobrepõem.

No início dessas duas disciplinas, é aplicado um teste de nivelamento para a distribuição dos alunos em turmas com níveis semelhantes de proficiência. Em relação às avaliações de rendimento, é seguido o padrão da série anterior com duas provas escritas somativas e duas provas orais cuja nota não entra no cômputo da média final, ou seja, é apenas formativa. No final de cada disciplina, os alunos desenvolvem uma atividade não avaliativa, denominada “projeto”. Na 2ª série, os alunos, divididos em grupos, assistem a diversos filmes sobre acidentes aéreos. Cada grupo prepara uma apresentação oral sobre a história do filme que assistiu. Os alunos da 3ª série participam de palestras proferidas, em inglês, por profissionais da aviação e do controle de tráfego aéreo. O objetivo, segundo o currículo da disciplina, é

promover a conscientização sobre a importância do inglês no controle de tráfego aéreo, além de desenvolver as habilidades de compreensão e produção orais. No entanto, em relação à 3ª série, não há nenhuma menção à realização de atividade posterior às palestras proferidas, exceto perguntas feitas pelos alunos durante o evento.

Nas recomendações metodológicas do currículo dessas duas disciplinas, o professor é instruído a planejar aulas voltadas à prática da língua, explorando diversos padrões de interação em sala de aula: aluno-aluno, aluno-grupo e professor-aluno, além de monitorar os alunos durante as atividades comunicativas para, posteriormente, corrigir os erros produzidos e oferecer *feedback* apropriado. Essa informação, porém, não consta no currículo da disciplina de Inglês I.

3.5.3.3 Inglês ATC

A disciplina de inglês da 4ª série, chamada de Inglês ATC, possui 110 tempos de instrução e 20 tempos de avaliação. Ela também está ancorada na abordagem ESP e seu objetivo geral faz referência explícita à tabela de níveis de proficiência da OACI: empregar as habilidades expressas pelos descritores holísticos, a saber, pronúncia, estrutura, vocabulário, fluência, compreensão e interação. Como objetivos específicos, são repetidos aqueles presentes nas disciplinas da 2ª e 3ª séries: produção de gêneros orais e aplicação de técnicas para desenvolvimento da compreensão oral. O conteúdo programático contempla temas relacionados à rotina profissional de pilotos e controladores de tráfego aéreo. É adotado o livro didático *Check your Aviation English*, da Editora *Macmillan*. Em relação ao processo de avaliação, duas provas escritas são aplicadas, como nas séries anteriores, para avaliação da leitura e compreensão oral separadamente. Como prova oral, os alunos são instruídos a criar um pôster com figuras relacionadas a um dos temas abordados na disciplina e apresentá-lo à turma. Primeiramente, eles fazem uma descrição das figuras contidas no pôster e, em seguida, contam uma história relacionando-as. Nessa série, a prova oral é considerada um projeto pelo fato de a elaboração do pôster ser feita ao longo da disciplina. Em relação às recomendações metodológicas ao grupo docente, são mantidas as orientações descritas no currículo das 2ª e 3ª séries. O Quadro 4 resume as principais características das disciplinas de língua inglesa nas quatro séries do curso.

A carga horária das disciplinas não é distribuída uniformemente durante o semestre, fazendo com que haja períodos razoavelmente longos sem aulas de inglês. A disciplina da 4ª série, Inglês ATC, é a mais concentrada, sendo ministrada no formato intensivo, com duração

de quatro a cinco semanas. Em 2016, por exemplo, essa disciplina foi ministrada no período de 27 de junho a 19 de julho, com uma média de sete tempos diários de aula. Em contraposição, a 3ª série do curso teve, em média, dois a três tempos de aulas de inglês por dia, no período de fevereiro a abril.

Em todas as disciplinas, as aulas são ministradas em sistema de rodízio de professores e instrutores, ou seja, os professores e instrutores revezam-se nas diversas turmas, geralmente, ministrando uma (1) unidade completa do material didático em cada grupo. Além disso, as unidades didáticas não são ministradas em sequência, o que permite ao professor/instrutor ministrar um mesmo conteúdo em mais de uma turma.

Quadro 4 − O ensino de inglês no curso BCT

DISCIPLINA	CH	FOCO	SEPARAÇÃO POR NÍVEL	LIVRO DIDÁTICO	AVALIAÇÃO	PROJETO
INGLÊS I	140	Inglês Geral	Não	Listen In, Volume 2, Ed.Cengage Learning	2 provas práticas com 40 questões cada: 20 com foco na compreensão oral e 20 de compreensão escrita (caráter somativo) + 2 provas orais (caráter formativo)	Não há.
INGLÊS II	120	Inglês Aeronáutico	Sim	English For Aviation, Ed. Oxford		Apresentação oral sobre um filme
INGLÊS III	120	Inglês Aeronáutico	Sim	Aviation English, Ed. Macmillan		Participação em Palestras
INGLÊS ATC	130	Inglês Aeronáutico	Sim	Check your Aviation English, Ed. Macmillan		Criação de um pôster e apresentação oral (teste oral)

3.5.4 O uso do EPLIS no contexto escolar

De 2007 a 2013, o EPLIS tinha como público-alvo somente controladores de tráfego aéreo e operadores de estação aeronáutica que se encontravam em pleno serviço. Assim sendo, os alunos procedentes da EEAR só se submetiam ao EPLIS após a formatura, quando já estavam atuando nos órgãos de controle para os quais foram designados, ou seja, seis a doze meses após a conclusão do curso, com raríssimas exceções. Porém, em 2014, o ICEA desenvolveu um novo conjunto de provas, denominado EPLIS *ab-initio*, voltado exclusivamente para controladores de tráfego aéreo em processo de formação. O EPLIS *ab-*

initio, a partir daquele mesmo ano, passou a ser aplicado, semestralmente, na EEAR, aos alunos da 4ª série BCT.

A versão *ab-initio* do EPLIS diferencia-se da versão elaborada para profissionais em serviço por não abordar questões que requerem experiência profissional em determinado órgão de controle para serem solucionadas.[46] Não há diferenças em termos estruturais, de *status* ou de resultados, sendo, portanto, considerado um exame oficial, cujo resultado obtido constará no registro do futuro profissional junto ao DECEA. Dessa forma, o aluno que foi considerado aprovado apresenta proficiência mínima para atuar no órgão de controle para o qual será designado após a conclusão do curso. Não existe ainda nenhum tipo de sanção imediata ao aluno que não atinge o nível mínimo, fora a necessidade de realizar novamente o exame no ano seguinte. Por outro lado, para os aprovados, benefícios como, por exemplo, maior flexibilidade nas trocas de turnos são oferecidos pelos órgãos de controle nos quais irão atuar.

Conforme explicitado no Capítulo I, apesar de não ter sido divulgada oficialmente a razão que levou à implementação do EPLIS nos cursos de formação, acredita-se que ela tenha sido motivada pelo interesse das autoridades em aumentar o grau de relevância do exame na escola, na expectativa de que houvesse aumento no nível de proficiência desses alunos. Sabe-se, também, que existe a intenção de fazer com que o EPLIS se torne um teste de saída, obrigatório para a conclusão do curso na EEAR, com o objetivo de elevar ainda mais as estatísticas de aprovação no exame.

3.5.4.1 As especificações do EPLIS

O ICEA é a instituição responsável pelo planejamento, desenvolvimento e aplicação de métodos, técnicas e atividades de avaliação de inglês aeronáutico no âmbito do COMAER. Como mencionado anteriormente, esse instituto desenvolveu o EPLIS e o aplicou pela primeira vez em 2007 para controladores de tráfego aéreo e operadores de estação aeronáutica em serviço. Desde então, o EPLIS é aplicado anualmente a esses profissionais, que estejam envolvidos com aviação civil internacional no Brasil e possuam a Carteira de Habilitação Técnica válida para trabalhar nos órgãos operacionais de tráfego aéreo (estações aeronáuticas, torres de controle, controles de aproximação e centros de controle de área). Em relação aos alunos da EEAR, desde o segundo semestre de 2014, o EPLIS é aplicado duas vezes ao ano, acompanhando o calendário escolar.

[46] O EPLIS para os profissionais em serviço apresenta três versões de acordo com o órgão operacional no qual o candidato atua: Torre de Controle, Controle de Aproximação e Centro de Área.

O EPLIS foi elaborado seguindo as recomendações feitas pela OACI no Manual de Implementação dos Requisitos de Proficiência Linguística (Doc 9835 AN/453), publicado com o objetivo de dar um suporte aos países em relação a aspectos de capacitação e avaliação de proficiência de pilotos, controladores de tráfego aéreo e operadores de estação aeronáutica envolvidos com aviação civil internacional. O objetivo do exame é avaliar o uso do inglês comum para comunicação aeronáutica (*plain English*) na solução de situações incomuns, inesperadas e/ou emergenciais que podem ocorrer na prestação do serviço de controle de tráfego aéreo. Pelo fato de a comunicação aeronáutica ser feita via radiotelefonia, o exame tem como foco as habilidades de compreensão e produção orais. O conteúdo do exame aborda temas, tópicos e funções pertinentes às comunicações aeronáuticas. O Quadro 5 descreve a estrutura do exame, suas fases, o modo de aplicação, o formato, o objetivo de cada fase, os critérios de avaliação utilizados e os tipos de tarefas a serem realizadas.

Quadro 5 − A estrutura e as características do EPLIS

Estrutura	Aplicação	Formato	Objetivo	Caráter	Tarefas
Fase I	a distância	30 questões de múltipla escolha	avaliar a compreensão oral do candidato	eliminatório − o candidato deve acertar, no mínimo, 20 questões para realizar a Fase II	O candidato escolhe uma única alternativa correta de acordo com o áudio. Os enunciados das questões e as alternativas estão em português.
Fase II	presencial	entrevista	avaliar a compreensão e a produção oral do candidato	é aferido ao candidato um nível de proficiência conforme descrito na tabela de níveis de proficiência da OACI	**Parte 1**: Quatro perguntas abertas sobre a profissão. **Parte 2**: Dez enunciados são lidos. O candidato explica a situação e oferece uma sugestão. **Parte 3**: Quatro perguntas abertas sobre aviação e controle de tráfego aéreo. **Parte 4**: O candidato descreve uma figura e elabora uma história.

O EPLIS pode ser considerado um exame de desempenho com propósitos específicos mesmo possuindo uma fase de múltipla escolha para avaliar a compreensão oral do candidato. McNamara (1997) explica que a avaliação de desempenho se apresenta em um contínuo que varia de uma abordagem mais forte para uma abordagem mais fraca em termos de representatividade do cenário real de uso da língua. A versão mais forte, conhecida como *work sample*, consiste em simular em alto grau as tarefas comunicativas reais que serão

desempenhadas pelo candidato no uso futuro da língua. Na tradição mais cognitiva, a tarefa em si é menos importante do que o quê o desempenho do candidato revela sobre sua habilidade linguística. Os testes fundamentados nessa abordagem apresentam uma orientação comunicativa, porém não estão preocupados em representar o cenário de uso da língua com um alto nível de verossimilhança.

No caso do EPLIS, sua fase de múltipla escolha, apesar de afastá-lo da abordagem mais forte, contém itens que abordam situações autênticas de interação entre piloto e controlador. Ademais, na fase II do exame, é avaliado o uso integrado de habilidades linguísticas na solução de problemas de tráfego aéreo, o que o caracteriza como um exame de desempenho com propósitos específicos. Segundo Scaramucci et al (2010), no EPLIS, conhecimentos específicos necessários para o desempenho da profissão interagem com os conhecimentos linguísticos exigidos em uma situação comunicacional, distanciando-o, assim, de um teste de conhecimento apenas.

O ICEA tem buscado submeter continuamente o exame a processos de validação com o objetivo de aprimorar sua qualidade, seguindo, assim, as recomendações da OACI quanto à coleta e à análise de evidências de validade que sustentem as interpretações e os usos de seus resultados e que demonstrem que o exame está alinhado com os critérios estabelecidos no Doc 9835 AN/ 453. Em 2009, o EPLIS foi enviado à análise de uma equipe de especialistas em avaliação de línguas da Unicamp, liderada pela profa. Dra. Matilde Virgínia Ricardi Scaramucci, pesquisadora de alto renome na área de avaliação em contexto de línguas. Novamente, em 2014, o exame foi apreciado pelo Serviço de Testes de Inglês Aeronáutico da OACI.

Essas averiguações promoveram melhorias no exame, tais como a criação de uma página na *web* exclusiva para o teste, a elaboração de um manual para o candidato, a substituição de áudios da fase I do exame cujos gêneros divergiam da comunicação aeronáutica e a criação de provas específicas por órgão de controle de tráfego aéreo para os profissionais em serviço e de uma versão *ab-initio* para aqueles ainda em formação. No entanto, as funções comunicativas avaliadas, os tipos de tarefas e formatos não sofreram alterações em relação à primeira versão aplicada em 2007, com exceção da inclusão das tarefas de descrever uma imagem e relatar uma história sobre ela, ocorrida em 2009. As principais críticas ao exame envolvem divergências entre a natureza da interação em contextos reais de uso da língua e aquela propiciada na fase II e a necessidade de maior correspondência entre os propósitos comunicativos das situações autênticas e aqueles suscitados na entrevista oral. Essas questões

serão retomadas nos Capítulos IV e V, dedicados à análise dos dados. A seguir, detalho as fases do exame, os tipos de tarefas e seus propósitos.

3.5.4.2 A fase I do EPLIS

A compreensão oral é essencial para as comunicações aeronáuticas, porque as interações são feitas, em sua grande maioria, por rádio frequência. Uma das características chave do uso da língua nesse contexto é a habilidade de entender diversos sotaques, uma vez que essa comunicação abrange uma comunidade internacional de usuários. Por essa razão, a primeira fase do EPLIS é dedicada, exclusivamente, a avaliar a compreensão oral do candidato por meio da escuta de gravações sobre situações de controle de tráfego aéreo dentro de uma gama variada de sotaques. Essa fase é realizada, a distância, por meio do computador que gera, aleatoriamente, as questões a serem respondidas por cada candidato, com base no grau de dificuldade dos itens. São, ao todo, 30 questões de múltipla escolha e cada item refere-se a um áudio diferente. É considerado apto para a segunda fase do EPLIS o candidato que obtiver um mínimo de 20 questões corretas. Os enunciados e as alternativas de resposta estão em português, pois o que está sendo avaliado é apenas a compreensão oral do candidato.

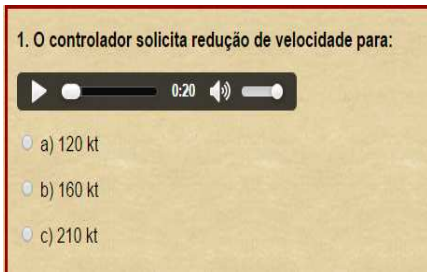
As gravações trazem situações verossímeis da comunicação piloto-controlador via rádio, com duração entre 10 e 45 segundos, podendo ser o áudio real ou uma regravação feita em laboratório. Os temas abrangem tanto situações rotineiras (mudança de nível, redução da velocidade etc) quanto situações não-rotineiras (fogo a bordo da aeronave, falhas de comunicação, falhas mecânicas, problemas com combustível, entre outras).[47] A escolha dos áudios busca contemplar sotaques presentes no controle de tráfego aéreo internacional, sem priorizar uma nacionalidade em detrimento de outra. Além disso, também são levados em consideração os diferentes órgãos de controle de tráfego aéreo (torre, controle de aproximação e centro de área) onde as interações ocorreram, a fim de manter um equilíbrio entre elas, uma vez que cada um desses órgãos apresenta características particulares.

Em relação aos itens dessa fase, para cada questão existem três alternativas de resposta, das quais apenas uma é considerada correta. As perguntas podem focar na ideia geral da interação, em informações específicas ou em detalhes. Veja abaixo dois exemplos de questões da fase I.[48]

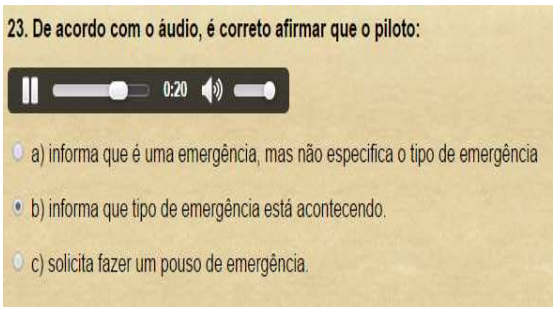
---

[47] A lista completa dos temas e domínios encontra-se no Apêndice B, Parte II, do Doc 9835 AN/453.

[48] Todos os exemplos de tarefas do exame apresentados nesta pesquisa são públicos e podem ser encontrados no seguinte endereço eletrônico: eplis.icea.gov.br.



Fonte: https://eplis.icea.gov.br



Fonte: https://eplis.icea.gov.br

Como na fase II do EPLIS, o candidato interage apenas com o entrevistador, justifica-se, portanto, a realização da fase I para avaliar sua habilidade na compreensão de diferentes sotaques. Além disso, por questões de logística, praticidade e limitação de recursos humanos, a fase I funciona também como um filtro, selecionando apenas os candidatos com maior potencial para atingir o nível mínimo exigido. Cerca de 3.000 profissionais realizam a primeira fase do EPLIS por ano, o que impossibilita a aplicação da segunda fase a todos os candidatos. Por fim, vale ressaltar que a nota obtida na fase I não é contabilizada no critério de compreensão para a fase seguinte.

3.5.4.3 A fase II do EPLIS

A segunda fase do EPLIS consiste na avaliação integrada das habilidades de compreensão e produção orais do candidato, no formato de uma entrevista. Foi elaborada com o propósito de obter uma amostra significativa do desempenho oral do candidato. São quatro conjuntos de prova: 1) EPLIS *ab-initio*,[49] voltado para os controladores de tráfego aéreo em formação; 2) EPLIS TWR, para os profissionais que atuam na torre de controle ou estação aeronáutica; 3) EPLIS APP, para os que trabalham no centro de controle de aproximação e; 4)

[49] Desde o final de 2014, o EPLIS tem sido aplicado aos alunos concludentes dos cursos de formação em Controle de Tráfego Aéreo como um exame obrigatório. Os resultados informam os órgãos operacionais que receberão esses profissionais sobre sua capacidade linguística. Por enquanto, a não aprovação no EPLIS não impede o aluno de concluir o curso e iniciar sua carreira profissional.

EPLIS ACC, para os controladores no centro de controle de área. A aplicação do teste dura de 15 a 30 minutos e apenas um candidato é avaliado por vez.

A avaliação é feita por dois profissionais. Um deles fica a cargo da interlocução e atribui, ao final da entrevista, uma nota geral ao desempenho demonstrado pelo candidato, utilizando uma grade de avaliação holística, elaborada com base na escala de níveis de proficiência da OACI (Anexo B). O segundo examinador posiciona-se fora do campo de visão do candidato e é responsável por atribuir uma nota analítica, ou seja, uma nota para cada um dos aspectos previstos nessa tabela. Para embasamento de sua nota, esse avaliador deve preencher um formulário de desempenho, com amostras da produção do candidato. Não há qualquer tipo de interação entre esse avaliador e o candidato.[50] Caso não haja concordância na nota final atribuída por esses dois avaliadores, um terceiro avaliador é requisitado para que seja feita uma nova avaliação do desempenho do candidato.

A entrevista está dividida em quatro partes. Na primeira parte, o candidato responde a quatro perguntas sobre sua carreira e rotina de trabalho. É permitido ao candidato negociar significado caso seja necessário, solicitando ao interlocutor repetir a pergunta ou parafraseá-la. São avaliados, nesse momento, a compreensão face a face, o uso da língua, a clareza com que o candidato se expressa e o emprego adequado das estratégias de negociação de sentido. As perguntas são, geralmente, abertas e abrangem situações no presente, passado e futuro. Caso haja questões no formato sim/não, é pedido ao candidato que justifique sua resposta. Seguem abaixo modelos de questões utilizadas nessa parte do exame.

	Part I
1.	When did you start your career and why?
2.	What's the importance of your job to aviation?
3.	What was the most difficult part of becoming an air traffic controller?
4.	What are your plans for the future?

Fonte: https://eplis.icea.gov.br

A parte II é composta por 10 situações incomuns, inesperadas ou de emergência plausíveis de ocorrerem no controle de tráfego aéreo. Primeiramente, o interlocutor posiciona uma barreira visual entre ele e o candidato, simulando, assim, uma característica importante da interação piloto-controlador: a ausência de contato visual. Para cada situação lida pelo

[50] Desde 2016, por restrição orçamentária, o avaliador responsável pela avaliação analítica não tem estado presente durante a entrevista, atribuindo, portanto, sua nota posteriormente, por meio do áudio gravado e inserido no sistema de gerenciamento de avaliações.

interlocutor, o candidato é instruído a explicar o que entendeu e, em seguida, a fazer uma sugestão para a solução do problema apresentado. Nessa parte, o candidato pode pedir uma repetição para cada situação lida. Porém, não é permitido ao interlocutor parafrasear, caso o candidato apresente problemas de compreensão.

Os objetivos são avaliar a compreensão somente por voz, o uso da língua e a clareza com que o candidato se expressa, assim como sua capacidade em transitar entre a fraseologia e o inglês comum para comunicação aeronáutica. As situações abordam os temas e domínios presentes no Doc 9835 AN/453 e estão ordenadas de acordo com o grau de dificuldade. Elas são, relativamente, curtas para que a memória não interfira na recuperação das informações. Seguem abaixo alguns exemplos de situações:

	Part II
1.	There are flocks of birds flying in the vicinity of the airport.
2.	The pilot lost radio contact after reaching cruising altitude.
3.	The rain has stopped, but the runway is still contaminated.
4.	The cabin crew is trying to control an aggressive passenger so the pilot requested diversion.

Fonte: https://eplis.icea.gov.br

Na parte III, o bloqueio visual é retirado e o candidato deve responder a quatro perguntas abertas sobre aviação e controle de tráfego aéreo. Elas estão ordenadas em grau de dificuldade. As duas primeiras costumam conter assuntos mais concretos, relacionados ao trabalho, ao passo que as duas últimas são mais abstratas. Nessa parte, o candidato pode, além de pedir uma repetição para cada uma das perguntas, solicitar paráfrase. São avaliados a compreensão face a face, o uso da língua em situações concretas e abstratas e o emprego das estratégias de negociação de sentido. Nessa parte, é permitido ao candidato demonstrar uso mais sofisticado da língua, característica presente nos candidatos de nível 5 e 6. Seguem exemplos de questões da parte III:

	Part III
1.	How do different weather conditions affect aviation where you work?
2.	Do you think that computers will replace air traffic controllers in the future? Why? Why not?

Fonte: https://eplis.icea.gov.br

Na parte IV, é apresentada ao candidato uma figura relacionada a situações de tráfego aéreo. O candidato tem 15 segundos para observar a imagem e, após esse tempo, ele é

instruído a descrevê-la por um (1) minuto. Em seguida, são dados ao candidato mais 30 segundos para pensar em uma história sobre a figura que ele deverá contar em um (1) minuto e 30 segundos. Os objetivos dessa parte são avaliar o uso da língua, a clareza com que o candidato se expressa, assim como a capacidade de produzir turnos de fala mais longos. Segue abaixo um modelo de imagem utilizada nessa parte do exame:



Fonte: https://eplis.icea.gov.br

3.6 O estudo preliminar

O estudo preliminar foi realizado no período de julho de 2014 a março de 2016. Como exposto anteriormente, essa etapa abordou um estudo minucioso da literatura existente sobre efeito retroativo, a análise dos documentos que embasaram a elaboração do EPLIS, a coleta de dados de base com os professores por meio de um questionário inicial e visitas técnicas ao cenário pesquisado, bem como observações de aulas.

Conforme já descrito, no final de 2014, as autoridades decidiram implementar o EPLIS *ab-initio* na 4ª série do curso BCT. Abriu-se, então, uma grande oportunidade para a coleta de dados preliminares sobre as percepções e atitudes dos professores em relação ao exame. Para isso, foi elaborado um questionário inicial (Apêndice A), fundamentado na revisão da literatura feita previamente, principalmente, no estudo de Cheng (2005). Antes de sua aplicação, o questionário foi submetido à revisão de dois especialistas da área de Linguística Aplicada. Responderam ao questionário 12 professores, o que correspondeu a um retorno de 60%. É importante pontuar que, na época em que o questionário foi preenchido, os professores não tinham ainda conhecimento de que o exame seria implementado na EEAR.

Acredito que esses dados iniciais sobre as percepções dos professores podem ser considerados dados de base (*baseline data*). Dados de base são entendidos como informações

sobre o estado atual de um determinado evento ou situação. Segundo Allwright e Bailey (1991, p. 74, tradução nossa), eles "fornecem a base contra a qual podemos fazer alegações comparativas sobre o quão diferente ou igual o fenômeno que observamos pode estar".[51] Dessa forma, o fato de o EPLIS ter sido introduzido na EEAR, possivelmente, pela necessidade de aumentar sua relevância entre alunos e professores, faz com que as informações coletadas no estudo preliminar possam ser consideradas de base, podendo, então, ser confrontadas com os dados gerados posteriormente, a fim de averiguar se as mudanças pretendidas foram alcançadas.

O questionário para a coleta de dados de base foi elaborado em três partes. Na primeira parte, intitulada "Informações Gerais", os participantes foram orientados a escolher a alternativa que melhor representava seu perfil e o perfil de seus alunos. Foram coletados dados sobre sexo, idade, qualificação, experiência no ensino de inglês, experiência no ensino de inglês aeronáutico, séries para as quais costuma lecionar, número de alunos por sala, língua falada em sala de aula e material didático utilizado. Na segunda parte do questionário, foram feitas 10 perguntas, em escala *likert*, sobre o grau de influência do EPLIS em diversos aspectos do ensino e aprendizagem de inglês. Também foi perguntado como os professores conheceram o EPLIS e se o exame exercia algum impacto em suas aulas. A terceira parte, composta por três itens, focou nas crenças dos professores sobre ensino e aprendizagem de língua e em sua prática pedagógica. A primeira pergunta era aberta e contrapunha o ensino de inglês geral e o ensino para fins específicos. Nas demais, os professores tinham que selecionar as habilidades e os tipos de atividades mais trabalhadas em sala de aula.

Esse questionário foi redigido em inglês a fim de facilitar a publicação e divulgação dos resultados em âmbito internacional. No entanto, embora não houvesse solicitação de identificação no preenchimento, os professores se sentiram incomodados em ter que responder em inglês às perguntas abertas. Para evitar que isso os impedisse de expressar sua opinião sobre as questões apresentadas, optei imediatamente por permitir respostas em português.

Entretanto, não foi possível obter dados de base com os alunos da EEAR, porque, naquele momento, não havia ainda autorização oficial do Departamento de Ensino da Aeronáutica (DEPENS) para a coleta de dados diretos com esse grupo de participantes. Consegui aplicar um breve questionário (Apêndice B) a alguns alunos da 4ª série minutos após terem realizados a fase II do exame. Bastante similar ao do professor, esse questionário também possuía três partes. Na primeira, foram obtidos dados relativos a sexo, idade, anos de estudo de inglês e nível de proficiência na língua. Na segunda parte, os alunos foram orientados,

[51] No original: "It provides the basis against which we make comparative claims about how different or usual the phenomena we have seen may be".

primeiramente, a selecionar como conheceram o exame e, em seguida, a responder perguntas, em uma escala *likert*, sobre o grau de influência do exame em sala de aula. Na terceira parte do questionário, eles enumeraram os tipos de atividades mais trabalhadas por seus professores. O questionário foi respondido por oito alunos, cerca de 20% do número total de alunos da 4ª série daquele semestre. Esses dados não podem ser considerados representativos das percepções dos alunos dessa série sobre o EPLIS pelo número bastante reduzido de respostas. Também não podem ser considerados dados de base por terem sido coletados após a implementação do exame.

Quanto às visitas técnicas, elas foram realizadas entre os meses de janeiro a março de 2016, a fim de compreender a natureza do ensino de inglês na EEAR e sua relação com o exame. Embora eu tenha trabalhado na EEAR por um período de aproximadamente 2 anos, entre 2008 e 2009, esse fato não excluiu a necessidade de maior aprofundamento sobre o cenário de investigação, já que muitas mudanças poderiam ter ocorrido e afetado o ambiente escolar de alguma forma nos últimos anos. Além de conversas informais com professores, coordenadora pedagógica, chefe e adjunto da Subseção de Idiomas, essas visitas contemplaram observações de aulas não estruturadas nas 2ª, 3ª e 4ª séries do curso BCT, ou seja, sem a existência de hipóteses preestabelecidas a serem confirmadas ou refutadas, o que me permitiu observar o desenrolar dos eventos no contexto sem a necessidade julgar a relevância deles para o estudo (COHEN et al, 2007).

Os registros foram feitos por meio de notas de campo e abrangeram diversos aspectos sobre o ensino e aprendizagem de inglês no cenário investigado: número de alunos por turma, nível de proficiência dos alunos, tipos de atividades, conteúdo das atividades, instruções dadas pelo professor, padrão de interação para a realização das atividades (professor-classe, aluno-aluno, professor-aluno), tempo de fala do aluno, tempo de fala de professor. Também foram tomadas notas de conversas entre professor e alunos sempre que um ou outro mencionava o exame. Não utilizei gravações em áudio ou vídeo nesse momento, pois a intenção era tornar minha presença em sala de aula a menos intrusiva possível e, assim, garantir dados de interações mais espontâneas e naturais entre professor e alunos.

Os dados coletados contribuíram para um melhor entendimento sobre o cenário da pesquisa e sua relação com o exame e possibilitaram delimitar o escopo do estudo e refinar as perguntas, além de terem sido fundamentais para a escolha e construção dos instrumentos de coleta de dados.

3.7 A fase quantitativa

O objetivo da fase quantitativa foi identificar padrões e tendências nas percepções e atitudes dos participantes em relação ao efeito retroativo do EPLIS em diversos aspectos do ensino e aprendizagem de inglês na EEAR. Para isso, foram construídos e aplicados aos alunos e professores dois questionários. Os dados numéricos de ambos os questionários foram analisados a partir da estatística descritiva. Os dados dos alunos foram também submetidos à análise de componentes principais para a extração de seus construtos fundamentais e a testes de hipóteses a fim de identificar relações de (in)significância entre os componentes extraídos e três fatores: a proficiência linguística do aluno, seu conhecimento sobre o EPLIS e a proximidade do exame. Esses três fatores foram escolhidos por terem sido considerados muito salientes na determinação do efeito retroativo no aluno em outros estudos (WATANABE, 1996; STONEMAN, 2006; XIE; ANDREWS, 2012; STENLUND et al, 2017).

3.7.1 Questionários

O uso de questionários tem se mostrado bastante útil na coleta de informações sobre percepções, atitudes, crenças e comportamento de um determinado grupo de participantes, com o objetivo de descrever a natureza desses eventos, identificar padrões com que eles podem ser comparados e determinar as relações existentes entre eventos específicos (COHEN et al, 2007). Cohen et al (2007) apontam que o pesquisador que utiliza questionários em seus estudos pretende:

> coletar dados em grande escala de uma amostra representativa de uma população, a fim de afirmar, com uma medida de confiança estatística, que certas características observadas ocorrem com regularidade, ou que certos fatores se agrupam, ou que eles se correlacionam, ou que eles mudam em função do tempo e do espaço[52] (COHEN et al, 2007, p. 207, tradução nossa).

Esses propósitos coincidem com o objetivo desta pesquisa de investigar o efeito retroativo do EPLIS nas percepções e atitudes de professores e alunos em sala de aula e estabelecer padrões, tendências e relações entre eles. Somado a isso, o fato de haver cerca de

[52] No original: "[…] to gather large-scale data from as representative a sample population as possible in order to say with a measure of statistical confidence that certain observed characteristics occur with a degree of regularity, or that certain factors cluster together or that they correlate with each other, or that they change over time and location".

400 alunos favoreceu o uso de questionários fechados. Outra opção poderia ser custosa e demandar muito tempo de aplicação e análise.

Além da vantagem de se obter dados de um número expressivo de participantes de uma só vez, Cohen et al (2007) destacam outros benefícios: a) a geração de dados numéricos que contenham informações descritivas, inferenciais e explicativas; b) a elaboração de frequências e análise de correlações existentes entre os fatores principais e variáveis presentes nos questionários; c) observação de padrões de resposta; e d) generalizações entre esses fatores e variáveis, sendo todas essas propriedades de interesse para este estudo.

Embora sejam bastante práticos, bons questionários são difíceis de serem elaborados e necessitam de pilotagem para serem validados (CHENG, 2005; STONEMAN, 2006). Cohen et al (2007) apresentam três pré-requisitos para a elaboração de qualquer questionário: 1) especificação do propósito da investigação, 2) especificação da população foco do estudo, e 3) descrição dos recursos disponíveis.

Para definir exatamente o que eu pretendia com a aplicação do questionário, foi necessário, em primeiro lugar, estudar exaustivamente o referencial teórico já existente sobre efeito retroativo: sua natureza, características, os fatores que interagem com o fenômeno e o determinavam. Em relação à construção do questionário dos professores, as pesquisas com foco em suas percepções e atitudes são bastante numerosas. Ou seja, já existiam diversos questionários validados sobre a relação entre efeito retroativo e professor (CHENG, 1999, 2004, 2005; TSAGARI, 2009; QI, 2004). O mesmo, porém, não pode ser dito em relação às pesquisas que envolvem alunos, tornando essa tarefa mais árdua. De qualquer forma, por envolver contextos muito específicos e distintos de pesquisa, esses questionários não poderiam ser simplesmente replicados, havendo, portanto, a necessidade de adaptação para o cenário pesquisado.

Em ambos os casos, para a construção de um bom instrumento de coleta de dados, torna-se imprescindível conhecer a fundo o objeto de pesquisa, o contexto a ser investigado, assim como os participantes e suas relações. Nesse sentido, o estudo preliminar, conforme argumentado, foi de suma importância. As informações extraídas das respostas aos questionários iniciais, os dados coletados durante as visitas técnicas e observações de aulas, bem como o conhecimento adquirido sobre o efeito retroativo no processo de revisão da literatura contribuíram para gerar os itens utilizados nos questionários e garantir, assim, a validade do instrumento.

3.7.1.1 Questionário do professor (QP)

O questionário do professor (Apêndice C) contém 65 itens de múltipla escolha e está dividido em cinco partes. Com exceção das seis primeiras perguntas que estão em escala nominal, na qual o número relacionado à resposta do professor não tem significado quantitativo, representando uma classe ou categoria, todos os itens do questionário foram elaborados em escala *likert*, de 1 a 5, para analisar as percepções e atitudes dos professores em termos de nível de concordância, frequência e avaliação. O Quadro 6 apresenta os temas explorados no QP, o tipo de informações coletadas e o número de itens.

Quadro 6 − A estrutura do questionário do professor

Temas	Conteúdo	Itens
Parte 1 – Informações gerais	Informações sobre sexo, idade, qualificação profissional, formação em BCT, tempo de experiência no ensino de inglês e tempo de experiência no ensino de inglês aeronáutico. Nível de conhecimento sobre a área de Ensino para Fins Específicos e sobre inglês aeronáutico, séries para as quais lecionam.	9
Parte 2 – Conhecimento sobre o EPLIS	Nível de conhecimento sobre o exame em geral, objetivo do exame, Doc 9835 AN/453 da OACI, conteúdo, tarefas das fases 1 e 2, tabela de níveis de proficiência, desempenho do nível 4 de proficiência e manual do candidato, frequência de realização de palestras sobre o exame no ambiente escolar e frequência de consultas no *site* do EPLIS.	10
Parte 3 – Impacto do EPLIS no ensino	Nível de influência do exame na metodologia em sala de aula, no conteúdo, nas tarefas, no material didático, nas avaliações de rendimento, nos projetos; frequência com que o EPLIS é mencionado em sala de aula; discussão de aspectos do EPLIS e tarefas do exame em sala de aula; frequência com que o professor é questionado sobre o exame pelos alunos; frequência com que as atividades em sala de aula estão relacionadas com aviação e controle de tráfego aéreo.	15
Parte 4 − Preparação dos alunos para o EPLIS	Frequência com que certas atividades são trabalhadas em sala de aula; avaliação sobre a preparação do aluno para o exame.	16
Parte 5 – Visões e percepções sobre o EPLIS	Influência do exame na motivação dos alunos e na elevação do nível de proficiência; pressão do exame no professor; aumento na carga de trabalho e no nível de *stress*; importância do exame; relação entre exame e preparação.	15

A primeira parte do questionário, intitulada “Informações gerais”, tem como objetivo obter informações sobre o perfil dos professores e instrutores (sexo, idade, qualificação, formação em controle de tráfego aéreo, séries para as quais lecionam) e sobre sua experiência no ensino de inglês geral e no ensino de inglês aeronáutico. Outras duas perguntas

dessa parte abrangem o nível de conhecimento dos professores sobre ensino para fins específicos e sobre inglês aeronáutico. A parte II do questionário "Conhecimento sobre o EPLIS" apresenta nove itens e busca identificar como os professores avaliam os seus níveis de conhecimentos sobre o exame, sua base teórica, suas características, seu conteúdo, tipos de tarefas, bem como sobre a tabela de níveis de proficiência da OACI. A parte III, "Impacto do EPLIS no ensino" é composta por 15 perguntas, cujo objetivo é analisar o grau do efeito retroativo do exame na prática pedagógica do professor. Na parte IV, "Preparação dos alunos para o EPLIS", a intenção é identificar a frequência com que os professores trabalham diversos tipos de atividades em sala de aula e contém 16 itens. Na última parte, "Visões e percepções sobre o exame", são apresentadas ao professor 15 afirmações para as quais ele deve avaliar o seu nível de concordância. Elas abrangem questões relacionadas a impacto positivo do teste, impacto negativo, *status* e relevância.

3.7.1.2 Questionário do aluno (QA)

O questionário do aluno (Apêndice D) contém 64 itens de múltipla escolha e está dividido em quatro partes. A primeira parte é chamada de "Informações Gerais" e tem o objetivo de coletar informações sobre o perfil do aluno e o seu nível de proficiência geral na língua inglesa e nas quatro habilidades. Todas as nove questões dessa parte estão em escala nominal, no qual os números representam categorias. Com exceção desses itens, as demais questões do questionário estão em escala *likert* de concordância, avaliação ou frequência, e as respostas variam de 1 a 5. A segunda parte "Conhecimento sobre o EPLIS" é idêntica à do questionário do professor e busca dados sobre os conhecimentos do aluno a respeito das características do exame, sua base teórica, conteúdo, tarefas e tabela de níveis de proficiência. A terceira parte é chamada de "Preparação para o EPLIS" e tem como foco os tipos de tarefas realizadas pelo aluno dentro e fora da sala de aula e a dedicação do aluno na preparação para o exame. A maioria dos itens pertencentes a essa parte aparece no questionário do professor, porém aqui tomam a perspectiva do aluno. A última parte, intitulada "Visões e Percepções sobre o EPLIS", dedica-se a levantar informações sobre como o aluno percebe o exame e o seu impacto na vida estudantil. Como pode ser observado, o questionário do aluno é intencionalmente semelhante ao do professor para que possam ser feitas comparações entre as percepções e atitudes dos dois grupos de participantes em relação ao exame. O Quadro 7 apresenta a estrutura do QA, juntamente com o conteúdo coberto e o número de itens presentes em cada parte.

Quadro 7 − A estrutura do questionário do aluno

Construtos	Conteúdo	Itens
Parte 1 – Informações gerais	Informações sobre sexo, série, turma de inglês, tempo de estudo na língua inglesa; nível de proficiência geral na língua e nível de proficiência nas quatro habilidades.	9
Parte 2 – Conhecimento sobre o EPLIS	Nível de conhecimento sobre o exame em geral, objetivo do exame, Doc 9835 AN/453 da OACI, conteúdo, tarefas da fase I e fase II, tabela de níveis de proficiência, desempenho do nível 4 de proficiência e manual do candidato; frequência com que os professores mencionam o EPLIS e discutem aspectos específicos do exame.	10
Parte 3 - Preparação dos alunos para o EPLIS	Frequência com que são trabalhadas atividades semelhantes às do exame em sala de aula; frequência com que as atividades em sala de aula estão relacionadas à aviação e controle de tráfego aéreo; Frequência com que certas atividades são trabalhadas em sala de aula; preparação para o exame fora da sala de aula; tempo destinado à preparação para o exame.	32
Parte 4 – Visões e percepções sobre o EPLIS	Influência do exame na motivação do aluno e na elevação do nível de proficiência; pressão do exame no aluno; aumento na carga de estudos e no nível de *stress*; importância do exame; relação entre exame e preparação.	13

3.7.1.3 Validação e pilotagem dos questionários

A primeira versão dos questionários foi submetida à análise de três especialistas da área de Linguística Aplicada, escolhidos por pertencerem à equipe responsável pelo desenvolvimento e aplicação do EPLIS e por terem conhecimento em avaliação de proficiência. Esses profissionais também são professores de inglês aeronáutico, com mais de sete anos de experiência na área, atuando na pós-formação do controlador de tráfego aéreo. Além desses três especialistas, o questionário do aluno também passou pelo parecer de três professores de inglês da EEAR com Mestrado em Linguística Aplicada. A contribuição desses professores foi indispensável pelo alto conhecimento que eles possuíam sobre o cenário investigado. Ambos os questionários foram analisados em termos de conteúdo, redação e formato. As falhas encontradas foram corrigidas (questões dúbias e questões de difícil compreensão) e as sugestões e críticas apresentadas foram aceitas. O questionário do aluno foi aplicado a 35 alunos da 4ª série do curso BCT do 1º semestre de 2016 para fins de validação. Houve um retorno de 100% dos questionários. A aplicação foi conduzida por mim para que os alunos pudessem fazer perguntas e sanar dúvidas. Essa etapa foi considerada essencial para a implementação de melhorias e realização de ajustes ao instrumento.

Na primeira versão dos questionários, existia a possibilidade de os participantes acrescentarem outra opção de resposta às alternativas. Por exemplo, na parte do questionário em que era solicitado a eles avaliarem a frequência dos tipos de atividades em sala de aula, eles poderiam acrescentar à lista pré-estabelecida outras atividades que considerassem importantes. As sugestões oferecidas foram incluídas no questionário final. Isso foi feito para que variáveis vistas como relevantes pelos alunos e professores sobre o impacto do EPLIS estivessem contempladas no questionário. Na versão final dos questionários, essas perguntas abertas foram retiradas e ficaram apenas itens fechados.

3.7.2 Procedimentos de coleta

Para a distribuição e preenchimento do QP foi utilizada a ferramenta *Google Forms*, serviço disponibilizado gratuitamente pela empresa Google para a criação e análise de pesquisas online. A interface para a elaboração dos itens é bastante amigável e conta com sete opções de formatos para perguntas: texto (para respostas abertas curtas), texto do parágrafo (para respostas abertas longas), múltipla escolha, caixa de seleção (possibilidade de marcar mais de uma resposta), escolha de uma lista (variação da múltipla escolha), escala (usada para atribuir uma nota a determinada pergunta, por exemplo, frequência) e grade (possibilidade de criar uma lista de perguntas ou afirmações para serem classificadas). Esses formatos atenderam plenamente os tipos de perguntas existentes no QA e no QP.

O *link* para preenchimento do questionário foi enviado, por *email*, aos 21 professores e instrutores de inglês aeronáutico que atuaram no curso BCT no ano de 2016, juntamente com uma breve mensagem explicativa sobre a pesquisa reforçando o caráter voluntário da participação e garantindo o anonimato na divulgação dos resultados. O questionário ficou disponível por um período de 17 dias, de 12 a 28 de fevereiro de 2017. Foram recebidas 16 respostas, contabilizando um retorno de 76%. A Tabela 1 mostra a taxa de retorno obtida em relação ao corpo docente permanente e volante.

Tabela 1 − Retorno de resposta do QP

	N° professor/instrutor	**Respostas recebidas**	**Taxa de retorno**
Corpo Docente Permanente	17	12	70%
Corpo Docente Volante	4	4	100%
TOTAL	**21**	**16**	**76,19%**

Quanto aos alunos, optou-se por disponibilizar o questionário dentro da rede interna da Escola para evitar que a dificuldade de acesso à internet prejudicasse a taxa de retorno de respostas. Para isso, foi necessário substituir o *Google Forms* pelo software de gerenciamento de questionários *Lime Survey* que, diferentemente do serviço da empresa Google, possibilita o preenchimento do questionário no modo *off-line*. Em termos de criação de formulários, ambos possuem os mesmos modelos de perguntas, além de permitirem inserir o Termo de Consentimento Livre e Esclarecido (TCLE) antes de o participante iniciar o seu preenchimento (Apêndices G e H).

Para otimizar a aplicação do questionário aos alunos, a escola disponibilizou um laboratório com 60 computadores e foram agendados dias e horários para cada série do curso. No total, foram quatro dias de aplicação distribuídos da seguinte forma: 4ª série (15 de julho de 2016), 3ª série (14 de setembro de 2016), 2ª série (19 de outubro de 2016) e 1ª série (03 de novembro de 2016). Essas datas foram escolhidas levando em consideração a programação das aulas de inglês. Assim, nenhum aluno respondeu ao questionário sem ter tido, pelo menos, quatro semanas de aulas na série correspondente. O número de alunos total e por série e a taxa de retorno de respostas estão dispostos na Tabela 2.

Tabela 2 − Retorno de Resposta do QA

	N° alunos	**Respostas recebidas**	**Taxa de retorno**
1ª série	129	116	89,92%
2ª série	117	116	99,14%
3ª série	113	109	96,46%
4ª série	48	46	95,83%
TOTAL	**407**	**387**	**95,08%**

Devido ao formato e à necessidade de serem concisos e práticos de serem respondidos, os questionários não oferecem aos respondentes muita liberdade para expandir ou justificar suas respostas. Ou seja, o potencial dos questionários de produzir dados explicativos e refinados é limitado (COHEN et al, 2007). Por esse motivo, no final de ambos, foi incluído um espaço livre para comentários gerais dos participantes em relação à pesquisa e/ou aos assuntos abordados. Recebi um número bastante expressivo de comentários tornando-os uma fonte rica de dados, capaz de explicar os resultados numéricos. Esses textos foram analisados na fase qualitativa, juntamente com os dados produzidos no grupo focal, nas entrevistas e nas observações de aulas.

3.7.3 As análises estatísticas

Os dados coletados por meio dos questionários foram analisados, primeiramente, por meio da estatística descritiva que permite sintetizar um conjunto de dados, utilizando medidas de posição (média, mediana e moda), medidas de distribuição (valores mínimos e máximos, desvio padrão, variância) e percentis. Em relação aos alunos, também foram aplicados a análise fatorial de componentes principais para a redução dos dados e testes de hipóteses paramétricos que possibilitam tecer afirmações, com mais segurança, sobre tendências verificadas na amostra. Os dados referentes aos professores, em razão do pequeno número de respondentes, 16 participantes, não apresentavam as condições necessárias para serem submetidos à análise fatorial e aos testes de hipóteses. A análise fatorial raramente é aplicada a amostras menores que 50 observações, além de serem necessárias, pelo menos, cinco vezes mais observações do que o número de variáveis analisadas. Os testes de hipóteses paramétricos, por sua vez, exigem que a amostra apresente uma distribuição normal e os dois grupos em análise devem apresentar igualdade de variâncias, o que também não é possível assumir com um número inferior a 30 casos. Por essas razões, os dados dos professores foram somente analisados pela estatística descritiva. A seguir, detalho a análise fatorial e os testes de hipóteses empregados aos dados dos alunos por serem técnicas menos conhecidas, ao contrário da estatística descritiva que é amplamente divulgada.

3.7.3.1 A análise fatorial e a análise de componentes principais

A análise fatorial é uma técnica estatística de análise multivariada de dados que permite medir, explicar e prever o grau de relação entre múltiplas variáveis. Esse tipo de análise mostra-se bastante apropriado em termos estatísticos para a investigação do efeito retroativo, uma vez que esse fenômeno é considerado multifacetado e mediado por diversos fatores que podem interferir na sua determinação como, por exemplo, formação dos professores, experiência de professores e alunos, nível de conhecimento sobre o exame, nível de relevância do exame, entre outros.

Em uma análise fatorial, assume-se que as variáveis em estudo são aleatórias[53] e inter-relacionadas de tal maneira que seus diferentes efeitos não podem ser interpretados isoladamente (HAIR et al, 2009). Em outras palavras, em um estudo no qual múltiplas variáveis

[53] Uma variável aleatória pode ser interpretada como uma variável quantitativa, cujo resultado depende de fatores aleatórios como, por exemplo, o resultado do lançamento de um dado.

estão sendo investigadas, aumenta-se a probabilidade dessas variáveis estarem sobrepostas (correlacionadas). A análise fatorial é capaz de agrupar variáveis altamente correlacionadas e descrever os dados em um número menor de conceitos do que as variáveis originais, com perda mínima de informação. A esses conjuntos distintos de variáveis fortemente inter-relacionadas dá-se o nome de fatores ou componentes. Os fatores representam as dimensões mais amplas e latentes dentro dos dados (construtos) que resumem ou explicam o conjunto de variáveis observadas. Essas dimensões mais amplas e latentes são compostas por itens específicos altamente correlacionados que representam uma faceta da dimensão em análise e fornecem uma visão sobre a sua constituição. De acordo com Hair et al (2009), o nosso conhecimento sobre determinado conceito nos permite interpretar e descrever as relações entre esses itens, fazendo com que as dimensões tenham significado para aquilo que elas representam conjuntamente.

Para entender melhor o funcionamento da análise fatorial, Hair et al (2009) oferecem um exemplo bastante simples. Em uma pesquisa sobre uma lanchonete, os clientes são convidados a avaliar o estabelecimento em relação a seis diferentes critérios: sabor, temperatura e frescor da comida, tempo de espera, limpeza e atendimento. A partir da análise fatorial das respostas dos clientes, é possível reduzir essas seis variáveis em dois fatores. As variáveis sabor, temperatura e frescor estão altamente correlacionadas, formando um único fator denominado de "qualidade da comida". Por sua vez, as variáveis tempo de espera, limpeza e atendimento se inter-relacionam e formam o fator "qualidade de serviço".

Existem dois tipos de análise fatorial: de natureza exploratória ou confirmatória. A análise exploratória é utilizada para compreender a estrutura em um conjunto de variáveis ou para reduzir o número original de variáveis. A análise confirmatória possibilita avaliar o grau com que os dados satisfazem uma estrutura estabelecida a priori. Para atender aos propósitos deste estudo de identificar as dimensões latentes do efeito retroativo do EPLIS em alunos e professores e levantar hipóteses sobre padrões e tendências nas percepções e atitudes desses dois participantes, foi realizada uma análise fatorial exploratória no emaranhado de variáveis que se relacionam com o fenômeno. O objetivo principal foi o de redução de dados, ou seja, foram identificadas as variáveis representativas a partir de um grupo maior, obtendo, dessa forma, um conjunto mais parcimonioso de variáveis. As mais representativas guiaram a fase posterior da pesquisa.

Hair et al (2009) advertem sobre a importância de se selecionar variáveis teoricamente fundamentadas. Segundo eles, a escolha indiscriminada de um grande número de variáveis, na expectativa de que a análise fatorial as organize, pode resultar em resultados bastante pobres para a pesquisa. Para evitar esse tipo de óbice, é necessário conhecer

profundamente o referencial teórico sobre o objeto de estudo e suas premissas conceituais, uma vez que a "qualidade e o significado dos fatores obtidos refletem as bases conceituais das variáveis incluídas na análise" (HAIR et al, 2009, p. 106). O pesquisador deve, então, pressupor uma estrutura subjacente para o conjunto de variáveis escolhido, teoricamente válida e fundamentada, a fim de garantir a relevância dos fatores a serem obtidos na compreensão do objeto de estudo. Esses autores argumentam ser de responsabilidade do pesquisador a tarefa de incluir variáveis suficientemente correlacionadas para que fatores verdadeiramente representativos sejam gerados.

Portanto, para o planejamento de uma análise fatorial, deve-se selecionar, minuciosamente, as variáveis que irão compor o estudo. Sobre essa questão, Hair et al (2009) apontam que, idealmente, devem ser incluídas, pelo menos, cinco variáveis para representar cada fator proposto. Dentro desse grupo, é interessante identificar as variáveis consideradas chave na determinação do fator latente proposto. Em termos de tamanho de amostra, o número de observações deve ser, preferencialmente, maior que 100. A análise fatorial é dificilmente aplicada em amostras menores, com menos de 50 observações. Quando isso acontece, qualquer resultado deve ser interpretado com precaução. A regra geral é ter cinco vezes mais observações do que o número de variáveis propostas. No caso desta pesquisa, havia 387 para um grupo de 64 variáveis, ou seja, uma razão de aproximadamente seis para um.

Foi utilizado, neste estudo, o método de análise dos componentes principais (ACP, doravante). A ACP é considerada a técnica mais típica nos estudos envolvendo análise fatorial com o objetivo de redução de dados. Neste método, os fatores são chamados de componentes. Para identificar o número ótimo de componentes a serem extraídos, é necessário, primeiramente, entender o conceito de variância.

A variância é a medida da dispersão dos valores de uma variável em relação a sua média (o quadrado do desvio padrão), ou seja, o quão longe os seus valores estão em relação ao valor esperado (média). Hair et al (2009) definem variância entre variáveis como sendo a quantia de partilha entre elas. "Se duas variáveis têm uma correlação de 0,50, cada variável compartilha 25% ($0{,}50^2$) de sua variância com a outra" (HAIR et al, 2009, p. 112). A variância total de uma variável é composta por três diferentes tipos de variância: comum, específica e de erro. A primeira é conhecida como "comunalidade" e refere-se à variância compartilhada de uma variável nas correlações com as demais. A variância específica ou única está associada a apenas uma variável e não apresenta correlações com outras. A variância de erro está associada a questões de não confiabilidade na coleta de dados.

Na ACP, os primeiros componentes a serem extraídos explicam uma porção substancial da variância total entre os dados. O processo de extração continua até que toda a variância seja explicada, podendo serem extraídos "n" componentes, no qual "n" representa o número de variáveis existentes. Porém, existe um ponto em que a quantia de variância comum começa a ser dominada pela variância específica ou única. Nesta pesquisa, foram utilizados três critérios para determinar o número ótimo de componentes a serem extraídos: o critério da raiz latente (ou critério Kaiser), o critério do teste *scree plot* e a análise paralela.

O critério kaiser parte do princípio que qualquer componente ou fator individual deve explicar a variância de, pelo menos, uma variável. Na análise dos componentes principais, é atribuído a cada variável o valor um (1) do autovalor total. Assim, componentes com autovalores menores que 1 são considerados insignificantes e descartados. No critério do teste *scree plot*, popularmente conhecido como gráfico de cotovelo, é elaborado um gráfico com as raízes latentes dos componentes, e o ponto em que a curvatura da linha começa a se tornar horizontal determinará onde será feito o corte. O método da análise paralela de simulação Monte-Carlo, por sua vez, tem sido cada vez mais utilizado na retenção de componentes pela tendência que os dois critérios anteriores possuem em superestimar esse número. Esse método consiste na simulação aleatória de um conjunto hipotético de matrizes de correlação de variáveis, utilizando o mesmo número de variáveis e de sujeitos do conjunto real de dados. É gerada uma lista de autovalores para cada componente ou fator extraído. Compara-se, então, o autovalor hipotético do primeiro componente ao autovalor real do mesmo componente e assim por diante. O número de componentes a serem retidos refere-se àqueles que apresentam autovalor acima daquele obtido por meio dos dados aleatórios e maior que 1 (DAMÁSIO, 2012). Neste estudo, utilizei uma matriz fatorada em 100 vezes. A combinação desses três critérios definiu o número de componentes a serem extraídos como as dimensões mais representativas dos dados.

Para auxiliar na interpretação de uma matriz fatorial, Hair et al (2009) propõem cinco etapas: 1) examinar a matriz fatorial de cargas rotacionada; 2) identificar a carga significante para cada variável nos diversos componentes ou fatores; 3) avaliar a comunalidade das variáveis; 4) reespecificar o modelo fatorial, se necessário; e 5) rotular os componentes. A rotação de componentes é uma ferramenta importante na obtenção de um padrão fatorial mais simples e teoricamente mais significativo. Essa técnica consiste em rotacionar os eixos de referência dos componentes, ou seja, redistribuir a variância dos primeiros para os últimos. A rotação mais comum e difundida é a ortogonal, na qual os eixos são mantidos a 90 graus. Nesta pesquisa, foi utilizado o critério VARIMAX para a rotação dos componentes. Esse critério

concentra-se na simplificação das colunas na matriz fatorial (componentes). O objetivo é fazer com que valores em cada coluna se mantenham entre +1 e –1 a fim de facilitar a interpretação. Correlações variável-componente mais próximas de +1 ou –1 indicam uma clara associação positiva ou negativa, enquanto correlações próximas de 0 apontam para falta de associação.

A significância de carga fatorial refere-se ao grau de correlação de cada variável com o componente. De maneira simplificada, quanto maior a carga fatorial de determinada variável, maior a sua representação dentro daquele componente. Em termos de significância prática, podemos dizer que: a) cargas entre ± 0,30 e ± 0,40 atendem, minimamente, para a interpretação dos dados, uma vez que o valor de 0,30, por exemplo, representaria apenas 9% da explicação da variância naquele componente (valor da carga ao quadrado); b) cargas iguais ou maiores que ± 0,50 são consideradas significantes; c) cargas acima de ± 0,70 são consideradas ideais para uma estrutura fatorial bem definida; e d) cargas acima de ± 0,80 são raras. A significância estatística, diferentemente, relaciona o tamanho da amostra e a carga fatorial significante. A Tabela 3 apresenta os valores de carga significante para diferentes tamanhos de amostra. Essa abordagem, porém, é considerada bastante conservadora e deve ser tomada apenas como ponto de partida.

Tabela 3 − Diretrizes para determinação de significância fatorial

Carga Fatorial	Tamanho da amostra para significância
0,30	350
0,40	200
0,50	120
0,60	85
0,70	60

Fonte: Adaptado de Hair et al, 2009, p. 120

Hair et al (2009) sugerem que outras considerações sejam incluídas no processo de determinação da significância das cargas fatoriais. Por exemplo, uma carga um pouco menor pode ser considerada significante em uma amostra grande ou com um número alto de variáveis. Por outro lado, uma carga maior faz-se necessária especialmente em componentes posteriores, uma vez que, após a obtenção do primeiro componente, as variâncias única e de erro começam a ocorrer. Neste estudo, foram consideradas significantes cargas fatoriais iguais ou maiores que ± 0,50. No entanto, é possível aceitar cargas levemente abaixo desse valor na extração dos primeiros componentes, do mesmo modo que o valor de significância pode ser aumentado em componentes posteriores, considerando a comunalidade das variáveis, ou seja, sua variância comum.

Em relação à significância das variáveis, casos de carga cruzada (quando uma determinada variável apresenta carga significante em diversos componentes) podem apresentar problemas ao pesquisador, e uma ação corretiva deve ser tomada. É possível, primeiramente, empregar um novo critério de rotação na tentativa de desfazer a carga cruzada. Se a carga cruzada persistir, o pesquisador pode ignorar ou eliminar essa variável.

O sucesso na interpretação dos componentes irá depender do quão teoricamente fundamentado está o pesquisador. Essa fundamentação conceitual advém do paradigma teórico adotado na pesquisa, do conhecimento da literatura sobre o objeto de estudo, do contexto investigado e de pesquisas anteriores. Na etapa de rotulação, o pesquisador deve examinar cuidadosamente a natureza das variáveis significantes dentro de um componente, buscando um significado comum entre elas. Variáveis com cargas mais altas possuem maior influência na escolha do nome ou rótulo por serem as mais importantes na explicação daquele componente. Quanto mais bem fundamentado teoricamente o pesquisador estiver, mais adequados serão os rótulos criados por ele.

O uso de variáveis múltiplas em técnicas multivariadas como a análise fatorial, por exemplo, traz à discussão uma questão bastante importante – o erro de medida inerente às variáveis. O erro de medida é definido por Hair et al (2009, p. 26) como "o grau em que os valores observados não são representativos dos valores 'verdadeiros'". Isso acontece por diversas razões, entre elas estão: erros cometidos na entrada de dados ou falta de precisão dos indivíduos em suas respostas. Para reduzir o erro de medida, é preciso considerar a validade e a confiabilidade da medida em questão. A fim de analisar a consistência entre as variáveis de uma escala (confiabilidade), o coeficiente Alfa de Cronbach tem sido amplamente utilizado. Essa medida identifica o grau de covariância entre os itens (variáveis) e varia de 0 a 1. Valores próximos de 1 mostram um grau alto de confiabilidade dos itens. Para este estudo, itens com valores inferiores a 0,3 foram desprezados. Em termos de validade (se a medida representa precisamente o que se espera), as variáveis, representadas pelos itens no questionário, foram escolhidas minuciosamente para que houvesse correspondência entre elas e sua definição conceitual.

No entanto, antes de aplicar a análise fatorial a um conjunto de dados, é necessário, primeiramente, se certificar que a matriz de dados é passível de fatoração. Dois métodos são aplicados para esse fim: o critério de Kaiser-Meyer-Olkin (KMO) e o teste de esfericidade de Bartlett. O KMO indica a proporção de variância nos dados que pode estar sendo causada por fatores subjacentes. Esse índice é calculado por meio do quadrado das correlações totais dividido pelo quadrado das correlações parciais. Os resultados variam de 0 a 1. Valores

próximos a 1 indicam que o conjunto de dados é passível de ser submetido à análise fatorial. O teste de Bartlett, por sua vez, compara a matriz de correlação dos dados a uma matriz identidade e checa se existe redundância entre as variáveis que poderia ser sintetizada em fatores. Valores de Bartlett com nível de significância (*p*-valor ou sig) menores ou iguais a 0.05 indicam que os dados são passíveis de fatoração. Neste estudo, os valores de KMO ficaram acima de 0.85, indicando uma boa adequação, e o teste de Bartlett apresentou valor de significância em 0.000, o que demonstrou possibilidade de fatoração dos dados.

3.7.3.2 Os testes de hipóteses

Testes de hipótese são procedimentos estatísticos que testam uma afirmação sobre as características de uma amostra, aceitando-a ou rejeitando-a, e fazem parte da chamada estatística inferencial. Embora a estatística descritiva possa ser útil em descrever padrões gerais, somente por meio da estatística inferencial é possível tecer generalizações genuínas sobre tendências observadas nos dados da amostra para a população (AL-HOORIE; VITTA, 2018). Em outras palavras, a estatística inferencial nos permite tirar conclusões indo do particular (amostra) para o geral (população).

Em um teste de hipótese, chamamos de hipótese nula (H_0) aquela que consideramos como verdadeira para a construção de um teste. A hipótese alternativa (H_1) é assumida quando a hipótese nula não apresenta evidência estatística suficiente. Ou seja, o parâmetro considerado é diferente daquele indicado na hipótese nula (H_0). Esses testes estão sujeitos a dois tipos de erros: o erro tipo I (quando a hipótese nula é rejeitada apesar de ser verdadeira) e o erro tipo II (quando a hipótese nula não é rejeitada apesar de ser falsa). O símbolo α (Alfa) é utilizado para representar a probabilidade de ocorrer o erro tipo I e o símbolo β (Beta) indica a probabilidade de ocorrência do erro tipo II.

Denomina-se região crítica o conjunto de valores que nos leva a rejeitar a hipótese nula. Geralmente, utiliza-se o método do *p*-valor (*p-value*), no qual a hipótese nula é rejeitada se o *p*-valor for inferior ou igual a α (nível de significância estatística). Da mesma forma, não se rejeita a hipótese nula se o *p*-valor for maior que α. Escolhas habituais para o nível de significância (α) são 0.05, 0.01 e 0.10. Isso quer dizer que quando estabelecemos α a 0.05, por exemplo, aceitamos uma margem de até 5% de erro.

Neste estudo, utilizei como testes de hipótese o teste *t* de *student* para amostras independentes e a análise de variância Anova. Para a aplicação de ambos os testes, assumi que

os dados apresentavam distribuição normal com base no teorema do limite central, que atesta que as médias de amostras grandes e aleatórias são aproximadamente normais. Outro pré-requisito para a aplicação desses testes é o de equivalência de variância. O teste de Levene (teste de igualdade de variâncias) foi empregado aos dados e demonstrou que as variâncias eram homogêneas.

O teste *t* de *student* para amostras independentes compara a diferença entre as médias de duas populações independentes. A hipótese nula, nesse caso, é representada pela igualdade das médias (H_0: $M_1=M_2$) a hipótese alternativa atesta que as médias são diferentes (H_1: $M_1 \neq M_2$). Os alunos foram, então, divididos em dois grupos por nível de proficiência e nível de conhecimento sobre o EPLIS, e dois testes *t* de *student* foram aplicados respectivamente, com o objetivo de comparar as médias dos grupos em cada um dos componentes principais extraídos.

Para identificar diferenças por séries (proximidade do exame), utilizei a Anova, pois permite comparar as médias de dois ou mais grupos independentes. Na Anova, a hipótese nula afirma que todas as médias da população são iguais e a hipótese alternativa afirma que *pelo menos* uma média é diferente. No entanto, a Anova não especifica quais grupos possuem médias diferentes, havendo a necessidade de aplicação de um teste *post-hoc*. Para isso, utilizei dois testes *post-doc*, Scheffe e Hotchberg, pois permitem fazer comparações duas a duas em grupos de tamanhos distintos. Assim, é possível distinguir quais grupos apresentam médias consideradas estatisticamente diferentes. Somente os dados relativos ao teste de Scheffe são reportados na análise, uma vez que os resultados de ambos não mostraram discrepâncias.

O uso demasiado e a grande dependência do *p*-valor para atestar significância nas pesquisas em segunda língua têm recebido diversas críticas nas últimas duas décadas. Essas críticas residem no fato do *p*-valor, ou seja, o nível de significância estatística poder ser afetado por diversos fatores, sendo o tamanho da amostra o mais determinante. Assim, em tamanhos grandes de amostra é mais provável obter um *p*-valor significativo do que em amostras pequenas. Em outras palavras, o α pode ser afetado pelo tamanho da amostra, evidenciando que a hipótese nula tende a ser sempre falsa se a amostra for suficientemente grande. Como solução, tem sido proposta a complementação dos resultados de *p*-valor (significância estatística) com a apresentação dos cálculos de tamanho de efeito da amostra (*effect size*), denominada significância prática (PLONSKY; OSWALD, 2014; ESPÍRITO-SANTO; DANIEL, 2015; AL-HOORIE; VITTA, 2018).

O tamanho de efeito (TDE) descreve a medida de magnitude do efeito observado, ou seja, o grau em que a hipótese nula é falsa. A medida mais amplamente conhecida para testes

envolvendo dois grupos independentes é o *d* de Cohen (*Cohen's d*). Essa medida pode ser calculada considerando a soma das médias dos grupos dividida pelo desvio-padrão. Também é possível empregar o *d* de Cohen com mais de dois grupos, por exemplo, no caso de Anova, porém ele deve ser calculado a cada par de comparação. Na equação abaixo, M_1 refere-se à média do grupo 1, M_2 representa a média do grupo 2, e DP é o desvio-padrão:

$$d = \frac{M1 + M2}{DP}$$

No caso deste estudo, por termos grupos de tamanhos distintos, utiliza-se o desvio-padrão combinado, dado pela seguinte equação, na qual N1 representa o número de sujeitos no grupo 1 e N2 representa o número de sujeitos no grupo 2:

$$DP_{combinado} = \sqrt{\frac{(n1-1).DP1^2 + (n2-1).DP2^2}{n1+n2-2}}$$

Quanto maior o valor do tamanho de efeito, maior a probabilidade de manifestação do fenômeno observado. Cohen (1988) classificou os efeitos em pequeno, moderado e grande, porém sem a pretensão de rigidez, estando, portanto, esses índices sujeitos à variação dependendo da área científica em que o estudo se insere, do contexto estudado e de resultados de pesquisas anteriores (ESPÍRITO-SANTO; DANIEL, 2015). No caso deste estudo, as pesquisas sobre efeito retroativo, que empregaram análises estatísticas e com as quais eu tive contato, apoiavam-se exclusivamente no método do *p*-valor para atestar significância estatística às relações encontradas. Então, por não haver valores de referência anteriores para o TDE, adotei a escala de Cohen (1988), conforme disposto no Quadro 8, para interpretar os resultados. Calculei o TDE para cada relação entre os fatores analisados e os componentes principais, tendo ela alcançado ou não significância estatística.

Quadro 8 − Valores de referência para TDE

TDE pequeno	**TDE moderado**	**TDE grande**
0,20 < d < 0,49	0,50 < d < 0.79	d > 0.80

Fonte: Baseado em Cohen, 1988, p. 24.

Coe (2002), baseando-se em Cohen (1988), explica que um tamanho de efeito médio de 0.50 pode ser notado a olho nu por um observador meticuloso. Um efeito pequeno de

0.20 é notadamente menor que um efeito médio, porém não tão pequeno para ser considerado trivial. Já um efeito grande de 0.80 é grosseiramente perceptível.

3.8 A fase qualitativa

Pretendeu-se, com a fase qualitativa, a expansão, o refinamento e a corroboração dos dados quantitativos sobre o efeito retroativo do EPLIS, a fim de se obter uma compreensão mais rica e detalhada do fenômeno no cenário investigado, atendendo, dessa forma, ao propósito de complementaridade e triangulação da pesquisa de métodos mistos. Os dados foram coletados por meio de comentários escritos deixados pelos participantes no preenchimento do questionário, de observações de aulas, do grupo focal com alunos e das entrevistas individuais com a chefia da Subseção de Idiomas, coordenação pedagógica, professores e instrutor. Além da triangulação entre métodos qualitativos e quantitativos, foi realizada, dentro da fase qualitativa, a triangulação de dados obtidos por diferentes fontes (comentários, grupos focais, observações e entrevistas). De acordo com Ollaik e Ziller (2012, p. 234), esse procedimento "visa controlar vieses e enriquecer constatações, bem como confirmar e reafirmar validade e confiabilidade" dos dados qualitativos, propiciando uma compreensão multidimensional do fenômeno em estudo. A seguir, discuto as razões que levaram à escolha dos instrumentos de geração de dados e detalho os procedimentos adotados para a coleta e análise dos dados qualitativos.

3.8.1 Observações

Uma das críticas de Alderson e Wall (1993) às primeiras pesquisas sobre efeito retroativo residia, exatamente, na carência de dados obtidos por meio de observações, com o objetivo de checar se os relatos de professores e alunos refletiam de fato o que ocorria em sala de aula. Nesse sentido, Watanabe (2004a) argumenta que as observações em sala de aula são essenciais para que o pesquisador não se apoie somente no que dizem professores e alunos, uma vez que essas percepções podem não refletir o que eles realmente fazem e, por essa razão, são parte dos instrumentos de coleta de dados utilizados nesta pesquisa.

Diferentemente da entrevista e do questionário, a observação foca em fatos, eventos ou comportamentos e, por isso, tende a gerar dados mais autênticos e válidos. Segundo Cheng (2005), esse instrumento é adequado a este tipo de estudo, porque, por meio dele, é possível presenciar as mudanças relacionadas ao teste no seu cenário natural. Em outras palavras, sua

maior vantagem encontra-se na oportunidade de o pesquisador olhar diretamente as relações que se desenvolvem naturalmente no ambiente investigado (Cohen et al, 2007).

As observações são classificadas em três tipos: estruturada, semiestruturada e não estruturada. A observação estruturada apoia-se na pré-seleção de categorias a serem observadas. Os eventos são sistematicamente anotados de acordo com as categorias escolhidas, o que facilita a coleta de dados numéricos. Em uma observação semiestruturada, existem algumas questões que direcionam o olhar do pesquisar, porém de uma maneira menos predeterminada. Na observação não estruturada, o pesquisador observa os eventos que ocorrem sem julgar a importância deles para o estudo.

O papel do pesquisador em uma observação pode tomar diversas formas que variam de acordo com o grau de proximidade entre ele e os participantes. O observador total distancia-se completamente. O observador participante busca um pouco mais de envolvimento ao participar de algumas atividades em sala de aula como, por exemplo, ajudar o professor a monitorar os alunos durante a realização de uma tarefa. O participante observador desenvolve um relacionamento maior com o grupo, porém o grupo ainda o vê como alguém de fora. Já o participante total age como se fosse membro do grupo e, na maioria das vezes, o grupo não está ciente do seu papel de observador.

Para este estudo, foi realizada uma observação semiestruturada, guiada por uma ficha de observação de aula com diversas categorias a serem preenchidas para direcionar a observação de eventos específicos e sua relação com o exame. Esses critérios contemplaram categorias de baixa inferência (eventos que podem ser contados ou codificados sem muita inferência do pesquisador) e categorias de alta inferência (requerem do pesquisador um julgamento sobre o que está sendo observado). Apesar de conter campos a serem preenchidos, essa ficha não tinha a intenção de ser exaustiva e possibilitava a inserção de outros eventos que, porventura, viessem a ocorrer e fossem considerados relevantes para o estudo. Essa escolha me permitiu participar de pequenas atividades em sala de aula ao mesmo tempo em que registrava os eventos como observadora participante. A intenção foi desenvolver um relacionamento com o grupo e coletar impressões, conversas e comentários de professores e alunos relacionados ao exame e sua influência.

3.8.1.1 Ficha de observação de aula

A fim de observar o efeito retroativo de um exame nas ações dos participantes envolvidos, é necessário, primeiramente, conhecer quais aspectos no ensino e na aprendizagem

podem ser afetados. Wall e Horak (2007) citam algumas áreas que devem estar contidas em uma ficha de observação para a verificação do efeito retroativo: a natureza da linguagem, as habilidades e subhabilidades trabalhadas em sala de aula, a integração entre as habilidades, os tópicos e temas desenvolvidos, os tipos de materiais utilizados, os métodos de instrução aplicados, as avaliações em sala de aula, as atividades realizadas dentro e fora do ambiente escolar e o tempo despendido nas tarefas.

Para a construção da ficha de observação de aulas utilizada neste estudo (Apêndice E), me baseei na proposta de Allen et al (1984), denominada de COLT (*Communicative Orientation of Language Teaching*). Esse instrumento de observação foi criado para descrever e analisar as interações ocorridas em salas de aulas de línguas, tendo como perspectiva o Ensino Comunicativo de Línguas (CANALE; SWAIN, 1980). A vantagem em relação às propostas anteriores está na inclusão de categorias que possibilitam avaliar o grau de proximidade entre os padrões de interação em sala de aula e o uso da língua em situações comunicativas reais.

O esquema de observação proposto por Allen et al (1984) mostrou-se bastante apropriado para o contexto da pesquisa por estar alinhado com a visão de língua presente no Doc 9835 AN/ 453 da OACI, na qual a competência comunicativa é multidimensional, composta pelas competências linguística, sociolinguística e pragmática. A OACI recomenda fortemente aos seus países membros que desenvolvam programas de capacitação voltados para o ensino comunicativo da língua e que incluam os seguintes elementos: a) foco na comunicação bem-sucedida em detrimento da correção puramente gramatical; b) aulas centradas no aluno; c) aumento do tempo de fala do aluno e diminuição do tempo de fala do professor; d) correção mínima de erros; e) materiais atraentes ao aluno.

Essa não é a primeira vez em que o COLT é empregado em estudos sobre efeito retroativo. Cheng (2005) utilizou esse instrumento para verificar o impacto de uma proposta de exame nas salas de aulas de inglês de escolas de nível médio em Hong Kong. O novo exame teve como pilar para o seu desenvolvimento o Ensino Comunicativo de Línguas, e o esquema de observação de Allen et al (1984) mostrou-se eficaz e adequado. Burrows (2001) também fez uso do COLT para observar mudanças nas práticas de quatro professores, relacionadas à implementação de um exame para avaliar a competência dos alunos na produção oral e escrita em inglês,, em um programa de Inglês para Imigrantes Adultos na Austrália.

O COLT tem como principal unidade de análise uma atividade realizada em sala de aula e está dividido em duas partes. A primeira parte (Parte A) descreve os eventos ocorridos na sala de aula no nível da atividade. A segunda parte (Parte B) analisa mais profundamente as características comunicativas nas interações entre professores e alunos. Neste estudo, foi

utilizada a parte A do COLT. Nessa parte, a atividade é analisada a partir de cinco parâmetros: tipo de atividade (*activity type*), organização participante (*participant organization*), conteúdo (*content*), habilidade (*student modality*) e material (*material*).

1. O parâmetro "tipo de atividade" é deixado em aberto para que possa acomodar uma grande variedade de atividades em sala de aula. Allen et al (1984) acredita, porém, que pode haver a predominância de um conjunto de atividades específicas, dependendo do programa.
2. A "organização participante" compreende três padrões de interação: toda a classe (*whole class*), em grupo e trabalho individual/grupo. No padrão "toda a classe", o professor pode interagir com um grupo de alunos ou toda a classe e vice-versa; a interação pode ocorrer entre aluno e toda a sala; e, também, pode haver trabalho em coro, no qual toda a classe ou um grupo de alunos participa em coro de uma atividade, normalmente, repetindo um modelo dado previamente. No "trabalho em grupo", os alunos podem se engajar todos na mesma atividade ou em atividades diferentes. Nesse caso, costuma-se detalhar o número de grupos, o número de alunos em cada grupo, os procedimentos e o tipo de monitoria oferecida. No padrão "individual/grupo", os alunos podem trabalhar individualmente em uma mesma tarefa ou enquanto alguns trabalham em grupos outros trabalham sozinhos.
3. O "conteúdo" das atividades em sala de aula pode compreender questões de gerenciamento, foco explícito em aspectos da linguagem (forma, função, discurso e características sociolinguísticas), tópicos diversos com escopo restrito (restritos à rotina da sala de aula), limitado (relacionados à vida pessoal ou social do aluno) ou amplo (questões mais abstratas e subjetivas), além dos tipos de controle do tópico (controlado pelo professor, compartilhado entre professor e alunos, controlado pelo aluno).
4. A "modalidade estudante" refere-se às diversas habilidades ou integração de habilidades trabalhadas em sala de aula. Por esse motivo, será utilizado na ficha de observação de aulas o termo "habilidade".
5. Os tipos de materiais são divididos em escrito, áudio ou visual. Podem ser classificados em relação ao tamanho (mínimo ou estendido), fonte/propósito (pedagógico, semipedagógico ou não-pedagógico) e uso (totalmente controlado, semicontrolado ou minimamente controlado).

Para atender ao propósito de investigar o efeito retroativo do EPLIS nas ações em sala de aula, fiz adaptações ao COLT. Em relação aos tipos de atividade, com base nas visitas técnicas, elaborei uma lista prévia de atividades prováveis de ocorrer em sala de aula. Essas atividades foram confrontadas com as tarefas presentes no exame para verificar se elas eram correspondentes ou não. O critério "conteúdo" das atividades foi desmembrado em tópicos, funções e aspectos linguísticos. O Apêndice B do Doc 9835 AN/453 traz um conjunto de funções e tópicos próprios das comunicações entre pilotos e controle de tráfego aéreo, fonte para a elaboração dos itens de prova do EPLIS e para o desenvolvimento de programas de capacitação. Foi verificada se existia consonância entre as funções e tópicos trabalhos em sala de aula e aqueles descritos nesse documento. Em termos de aspectos linguísticos, a tabela de níveis de proficiência da OACI contempla seis áreas avaliadas: pronúncia, estrutura, vocabulário, compreensão, fluência e interação. Analisei, então, quais desses aspectos eram enfatizados em cada atividade. Em relação ao critério habilidade, avaliei se a atividade contemplava o uso integrado ou isolado de habilidades. Pela natureza da comunicação entre piloto e controle de tráfego aéreo, esperava que a integração entre as habilidades de compreensão e produção oral seria a mais explorada no contexto escolar.

Em termos de interação, também assumi que o trabalho em grupo ou em pares dominaria as interações em sala de aula por propiciar o desenvolvimento da fluência e interação, favorecendo o tempo de fala do aluno e reduzindo o tempo de fala do professor. Em relação à duração dos turnos de fala, nas partes 1, 3 e 4 da fase II do EPLIS, é esperado do candidato que ele produza turnos de falas um pouco mais longos, além de coesos e coerentes. Por outro lado, na parte 2 da fase II, os turnos de fala do candidato tendem a ser mais curtos, devido à natureza da tarefa: explicar uma situação-problema e apresentar uma sugestão. Foram, então, observadas, também, se as atividades em salas de aulas refletiam essas características do exame.

Foi, ainda, incluída a categoria "Tipo de preparação" para analisar, com base na proposta de Popham (1991) discutida no Capítulo II, como o aluno está sendo preparado para o exame. Popham (1991) discorre sobre cinco tipos de preparação: com provas anteriores, em estratégias gerais para realização de testes, em mesmo formato, em formato variado e com prova atual. A categoria "preparação com prova atual" está ligada à posse ilegal de uma cópia do exame causada por vazamento ou roubo. Pelo fato de o exame ser aplicado por outra instituição (ICEA), essas situações foram desconsideradas e essa categoria foi excluída da ficha de observação de aula.

3.8.1.2 Procedimentos de coleta

Foram convidados a participar dessa etapa todos os professores e instrutores de inglês aeronáutico que atuaram no curso BCT no ano de 2016 e que haviam respondido ao questionário. Porém, apenas três professores e um instrutor demonstraram interesse em ter suas aulas assistidas e disponibilidade para participar de entrevistas individuais que sucederiam às observações. O Quadro 9 descreve o perfil desses participantes em termos de qualificação profissional, experiência no ensino de inglês geral e inglês aeronáutico e séries para as quais lecionaram no período da observação. Os nomes atribuídos são pseudônimos para preservar a identidade dos participantes.

Quadro 9 − Perfil dos participantes das observações

Participante	**Qualificação Profissional**	**Experiência no ensino de inglês**	**Experiência em inglês aeronáutico**	**Série em que atuou no período**
Pilar	Letras – Inglês e Mestrado em LA	19 anos	10 anos	2ª BCT 3ª BCT
Eduardo	Letras – Inglês e Mestrado em LA em andamento	22 anos	7 anos	2ª BCT 3ª BCT
João	Controlador de Tráfego Aéreo e Instrutor de Inglês Aeronáutico	10 anos	5 anos	2ª BCT 3ª BCT
Ana	Letras – Inglês Mestrado em LA em andamento	15 anos	4 anos	1ª BCT 4ª BCT

Apesar da disponibilidade e do interesse demonstrados por esses participantes, sabemos que ser observado, no entanto, pode ocasionar imenso desconforto e afetar, indesejavelmente, os acontecimentos em sala de aula. A respeito disso, Allwright e Bailey (1991) alertam para o "efeito da reatividade" − alteração no comportamento dos observados provocada pela observação em si − e oferecem algumas sugestões para lidar com o problema: a) garantir a confidencialidade e anonimato de todos os participantes; b) explicar aos professores e alunos, a priori, o propósito da pesquisa e os procedimentos a serem utilizados; c) familiarizá-los com o uso de aparelhos eletrônicos (gravador e câmera); e d) se colocar à disposição para eventuais dúvidas ou esclarecimentos.

Três dos quatro profissionais observados tiveram suas aulas assistidas durante a fase de estudo preliminar e, por essa razão, acredito que minha presença em suas salas de aula geraria menos desconforto. O outro profissional já havia participado de um curso, no qual eu era responsável por observar suas aulas e oferecer *feedback*. De qualquer forma, antes do início

das observações, entrei em contato com os quatro participantes via e-mail para explicar, em linhas gerais, os propósitos gerais da pesquisa e apresentar-lhes o TCLE referente às observações e entrevistas. Orientei-os a lerem o termo cuidadosamente e a se manifestarem caso houvesse qualquer dúvida. Dois professores solicitaram que, nas observações de suas aulas, fosse usado apenas o gravador de áudio e esse pedido foi atendido prontamente, sem questionamentos. Os outros dois concordaram em ter suas aulas filmadas. Pedi, então, aos participantes que preparassem seus alunos para as observações, expondo os propósitos do estudo e os procedimentos de coleta de dados.

As aulas das 2ª e 3ª série aconteceram, simultaneamente, no período de 31 de janeiro a 04 de maio de 2017. As observações ficaram concentradas entre os dias 22 de março e 12 de abril. Três dos quatro participantes estavam atuando nessas duas séries. As aulas da 4ª série ocorreram em bloco no período de 03 de julho a 07 de agosto. Apenas um dos participantes ministrou aulas para essa série e as observações ocorreram nas duas primeiras semanas da disciplina. Não foram observadas aulas na 1ª série do curso, porque, na época, o seu currículo estava sendo reestruturado a fim de se adequar melhor às necessidades e intereses dos alunos. O Quadro 10 traz mais informações sobre as observações realizadas.

Quadro 10 − Cronograma de observação de aulas

Participante	Série	Data	Turma	Tempos de aula assistidos	Total de tempos de aulas por série
Pilar	2 BCT	22 de março	G	2	8
		22 e 23 de março	A	2	
		29 de março	A	3	
		10 de abril	B	2	
	3 BCT	27 de março	B	3	7
		29 de março	G	2	
		10 de abril	B	1	
		10 de abril	G	1	
Eduardo	2 BCT	28 de março	D	3	8
		31 de março	C	3	
		11 de abril	C	2	
	3 BCT	22 de março	F	2	7
		28 de março	D	1	
		31 de março	D	2	
		11 de abril	E	2	
João	2 BCT	23 de março	G	2	4
		30 de março	E	2	
	3 BCT	27 e 28 de março	G	3	7
		30 de março	E	2	
		12 de abril	F	2	
Ana	4 BCT	06 de julho	E	3	11
		11 de julho	C	3	
		25 de julho	F	2	
		27 de julho	A	3	

Retornando a Allwright e Bailey (1991, p. 69, tradução nossa), se "os professores estão preocupados com o que o observador está fazendo, corre-se o risco de a aula ser menos efetiva".[54] Por essa razão, no primeiro dia de observação, me reuni com cada um dos participantes e apresentei um panorama geral sobre a pesquisa. Isso foi feito para diminuir o nível de ansiedade e tranquilizá-los em relação ao meu interesse em sala de aula. Aproveitei para enfatizar que estava aberta caso eles quisessem expor qualquer dificuldade ou dúvida. Fui apresentada aos alunos como a pesquisadora do ICEA que estava ali para coletar dados sobre o EPLIS. Por eu ter aplicado o questionário a todos os alunos na fase quantitativa do estudo, eles já me conheciam e sabiam, resumidamente, a respeito do meu interesse de pesquisa. Me posicionei no fundo da sala, juntamente com a câmera ou gravador, para evitar qualquer constrangimento ou inibição.

3.8.2 Entrevistas individuais

As observações em sala de aula foram seguidas de entrevistas individuais. Essas entrevistas funcionaram como espaços nos quais os participantes puderam expor suas reações, intenções e razões perante os dados coletados. A decisão por esse tipo de instrumento foi baseada na afirmação de Dushku (2000, p. 763, tradução nossa) de que o uso de entrevistas pode "fornecer informações sobre comportamento, atitudes e crenças e contribuir para uma compreensão mais profunda sobre as perspectivas e experiências dos participantes",[55] o que condiz com o propósito desta pesquisa de investigar as razões que levam os participantes a pensarem e agirem de determinada forma (WALL; HORAK, 2007). Outro motivo que me levou a optar por conduzir entrevistas individuais com os participantes foi o fato de esse instrumento aumentar a confiabilidade dos resultados, uma vez que os dados originados são produtos da interação entre entrevistador e entrevistado (BLOCK, 2000). Além do exposto, as entrevistas também são capazes de promover a reflexão crítica do professor sobre sua prática pedagógica e, assim, estreitar os laços entre pesquisa e ensino.

Existem três tipos de entrevistas que se distinguem pelo grau de controle exercido pelo pesquisador sobre as questões discutidas: entrevista estruturada, semiestruturada e não estruturada. Lankshear e Knobel (2004) pontuam que os modelos "semiestruturado" e "não estruturado" permitem ao pesquisador explorar mais as respostas dos participantes, enquanto o

[54] No original: If they are worried about what you are doing, then their lesson risks being less effective.
[55] […] interviewing can provide information on reported behavior, attitudes, and beliefs, and contribute to an in-depth understanding of research participants' perspectives or experiences.

modelo “estruturado” o prende a uma agenda fixa e inflexível. Para registro dos dados, o pesquisador pode se ater a notas de campo ou utilizar gravadores de áudio e até mesmo de imagens. Sabe-se, porém, que a coleta de dados por meio de entrevistas é sempre parcial e incompleta. Os dados coletados são dependentes do contexto, do momento da coleta, da seleção de participantes e das perguntas que são feitas pelo pesquisador (LANKSHEAR; KNOBEL, 2004).

Foram realizadas, ao todo, três entrevistas semiestruturadas com cada participante professor e instrutor observados, e cobriram diversos aspectos, tais como a experiência de ensino do participante, seus conhecimentos sobre o EPLIS, suas percepções e atitudes para com o exame e sua compreensão sobre a preparação do aluno (conteúdo trabalhado em sala de aula, métodos utilizados e avaliações de rendimento). Informações sobre quando as entrevistas ocorreram e o tempo de duração de cada uma delas são encontradas no Quadro 11.

Quadro 11 − Cronograma de entrevistas

Participante	**Entrevista**	**Data**	**Duração**
Pilar	Entrevista 1	28 de março de 2017	1h42min
	Entrevista 2	31 de março de 2017	1h03min
	Entrevista 3	12 de abril de 2017	32min
Eduardo	Entrevista 1	30 de março de 2017	1h05min
	Entrevista 2	10 de abril de 2017	39min
	Entrevista 3	12 de abril de 2017	33min
João	Entrevista 1	29 de março de 2017	54min
	Entrevista 2	11 de abril de 2017	18min
	Entrevista 3	12 de abril de 2017	23min
Ana	Entrevista 1	11 de julho de 2017	46min
	Entrevista 2	25 de julho de 2017	20min
	Entrevista 3	27 de julho de 2017	21min
Isabel	Entrevista 1	27 de julho de 2017	15min

A primeira entrevista foi dividida em dois momentos: primeiramente, coloquei perguntas para acessar as percepções e atitudes do participante em relação ao exame. No segundo momento, apresentei ao participante a ficha de observação de aulas preenchida com base nas observações realizadas até aquele momento. Foi dada ao participante a oportunidade de comentar sobre os dados coletados, posicionando-se a favor ou contra a minha interpretação. As entrevistas que sucederam foram iniciadas com a análise da ficha de observação de aulas e, em seguida, algumas questões, oriundas das reflexões feitas por mim durante a observação, foram debatidas com o participante, com o objetivo de contradizer ou confirmar meu ponto de vista. Ressalto que Pilar, professora que participou da observação de aulas, era também, na

época da coleta de dados, a coordenadora pedagógica das 2ª, 3ª e 4ª séries. Por essa razão, a primeira entrevista com ela abordou uma parte inicial dedicada ao seu papel como coordenadora. Apesar de não ter tido suas aulas assistidas, optei também por entrevistar a professora e chefe da Subseção de Idiomas Isabel para obter informações mais detalhadas acerca das decisões tomadas no contexto escolar quanto ao currículo, material didático, processo avaliativo e capacitação dos professores.

3.8.3 Grupo focal

Para a coleta de dados qualitativos sobre as percepções e atitudes dos alunos em relação ao EPLIS, realizei um grupo focal com a participação de 12 estudantes. O grupo focal é uma técnica para obtenção de dados qualitativos na forma de uma entrevista em grupo. Segundo Bonfim (2009), seu principal objetivo é reunir informações para uma compreensão mais aprofundada sobre as percepções, crenças, e atitudes dos participantes sobre um determinado assunto em um ambiente de interação. Sua maior vantagem é a possibilidade de produzir dados mais detalhados e ricos sobre determinado assunto, uma vez que, na presença de outros participantes, pontos de vista diferentes tendem a ser justificados em pormenores. Barbour e Schostak (2005, p. 43, tradução nossa) pontuam que os grupos focais se assemelham "ao mundo real onde as pessoas discutem, formulam e modificam suas visões e dão sentido às suas experiências".[56] Somado a isso, o grupo focal propicia a coleta de dados de diversos participantes de uma só vez, em um espaço curto de tempo e sem custos elevados. Em relação às desvantagens dessa técnica, os participantes podem se sentir constrangidos ao ter que expressar sua opinião na frente de outros. Questões muito pessoais também podem causar desconfortos e devem ser evitadas.

Por se tratar de um tipo de entrevista, o grupo focal também apresenta os modelos estruturado, semiestruturado e não estruturado. O papel do pesquisador em um grupo focal restringe-se à moderação e à condução. As principais tarefas são: estabelecer os objetivos, apresentar as regras, introduzir os assuntos a serem discutidos, encorajar a participação de todos, manter o foco nos interesses do estudo e explorar as respostas que possam ser mais relevantes para a pesquisa. Cabe também ao pesquisador criar um clima favorável ao debate, permitindo aos participantes expressarem seus pontos de vista sem, entretanto, induzir o grupo.

[56] No original: "Focus groups get as close as possible to the real-life situations where people discuss, formulate and modify their views and make sense of their experience".

Em suma, Barbour e Schostak (2005, p. 45, tradução nossa) definem o pesquisador como "a principal ferramenta para a geração de dados".[57]

Em relação à formação de um grupo focal, é necessário, primeiramente, selecionar os participantes considerando os propósitos da pesquisa e definir local e horário mais adequado a todos. O número de participantes deve variar entre 6 a 15. Grupos muito grandes podem dificultar a manter o foco das discussões, além de prejudicar a participação de todos os presentes. A duração irá depender do número de participantes e da complexidade dos assuntos abordados, porém não deve ultrapassar 2 horas para não se tornar exaustivo e repetitivo. Também é ideal que os participantes estejam reunidos ao redor de uma mesa ou que seus assentos formem um círculo.

Neste estudo, o grupo focal contou com a participação de 12 alunos, sendo 3 alunos de cada série do curso. Esses alunos foram selecionados, primeiramente, pelo interesse em participar e contribuir com a pesquisa. Essa seleção buscou, também, contemplar alunos com diferentes níveis de proficiência. Assim, dos 3 alunos selecionados de cada série, um demonstrava alto nível de proficiência, outro, nível de proficiência mediano, e o terceiro, nível de proficiência baixo. Foi agendado um único encontro no dia 11 de abril de 2017, em uma sala de aula nas dependências da escola, em horário noturno, para não prejudicar a rotina escolar.

O Quadro 12 retrata os participantes do grupo focal, a série do curso e a turma de inglês. Ressalto que, com exceção da 1ª série, os demais alunos são divididos em turmas de acordo com o nível de proficiência. Dessa forma, na turma A, estão os alunos com maior proficiência na língua inglesa. Esse nível cai gradualmente nas turmas B, C, D, E, F e G. Novamente, isso não se aplica aos alunos da 1ª série.

Quadro 12 − Participantes do Grupo Focal

Participante	**Série**	**Turma de inglês**
Aluno 1	1ª BCT	E
Aluno 2	1ª BCT	F
Aluno 3	1ª BCT	A
Aluno 4	2ª BCT	A
Aluno 5	2ª BCT	D
Aluno 6	2ª BCT	F
Aluno 7	3ª BCT	A
Aluno 8	3ª BCT	D
Aluno 9	3ª BCT	G
Aluno 10	4ª BCT	B
Aluno 11	4ª BCT	C
Aluno 12	4ª BCT	D

[57] No original: "Researcher as the principal data generating tool".

Antes do início do grupo focal, foi apresentado aos participantes o TCLE, no qual constava o propósito da pesquisa em termos gerais, benefícios e riscos. Foi garantido, aos alunos, o anonimato na divulgação dos resultados da pesquisa. Todos assinaram o termo e concordaram com o uso do gravador durante a sessão. Em seguida, foi explicado o objetivo do grupo focal e seu funcionamento. A opção pelo gravador, em substituição às notas de campo, foi feita com o objetivo de promover mais interação com os participantes, encorajando-os a elaborarem suas respostas e a fornecerem informações mais detalhadas. A filmagem, entretanto, apesar de ajudar a identificar o que cada participante disse e em que momento, causa maior inibição e, por isso, foi descartada.

Para direcionar as discussões, foi elaborada uma lista de 10 questões em um modelo semiestruturado (Apêndice F). Nesse modelo, existe um plano geral dos temas que serão abordados, mas é dada aos participantes certa liberdade para falar sobre outros assuntos que considerem relevantes. As perguntas abrangeram os seguintes tópicos: conhecimentos sobre o exame, relevância do exame, influência do exame no ensino e aprendizagem, preparação para o exame, incentivo, desafios e dificuldades. Esses temas foram extraídos da análise fatorial conduzida na fase quantitativa da pesquisa. O grupo focal teve duração de 1 hora e 28 min e alcançou o objetivo proposto.

3.8.4 A análise de conteúdo

Os dados coletados por meio de comentários escritos, entrevistas individuais e grupo focal foram analisados sob a luz da análise de conteúdo (AC doravante). A AC é um conjunto de técnicas para análise de dados qualitativos com o objetivo de reduzir grandes quantidades de texto em uma porção maleável e compreensível por meio de um processo sistemático (COHEN et al, 2007). Bardin (1977) apresenta duas funções da AC: levantar hipóteses ou confirmar hipóteses estabelecidas previamente. No caso deste estudo, hipóteses já haviam sido suscitadas na fase quantitativa, o que fez com que a AC fosse empregada com uma função confirmatória. No entanto, como o propósito da fase qualitativa não é apenas corroborar os dados iniciais, mas também explorá-los, ampliá-los e refiná-los, a possibilidade de novas hipóteses serem estabelecidas durante essa fase foi considerada.

Segundo Bardin (1977), a AC trata as informações presentes nas mensagens, principalmente, por meio da análise dos significados (temas), embora possa também contemplar um estudo dos significantes (escolha lexical ou estrutura sintática, por exemplo). Busca-se uma correspondência entre as estruturas semânticas ou linguísticas e as condutas,

crenças, percepções e atitudes do falante, em um processo inferencial. Bardin (1977) define a AC como um:

> conjunto de técnicas de análise das comunicações, visando obter, por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção destas mensagens (BARDIN, 1977, p. 42).

Para a condução de uma AC, é necessário, primeiramente, decidir sobre o tipo de unidade de registro a ser utilizado. A unidade de registro corresponde ao recorte do conteúdo para servir de unidade base. A unidade de registro pode ser uma palavra ou frase ou tema e, por isso, seu tamanho é variável. Para Bardin (1977, p. 105), o tema corresponde a uma afirmação acerca de um assunto, "geralmente utilizado como unidade de registro para estudar motivações de opiniões, de atitudes, de valores, de crenças, de tendências etc".

Bardin (1977) propõe dividir o processo da AC em três fases: pré-análise, análise e interpretação dos dados. Na pré-análise, são tomadas diversas decisões: o material é coletado e organizado, as hipóteses são formuladas, as categorias de análise são selecionadas, e as unidades de análise são definidas. A fase de análise corresponde à aplicação dessas decisões. Na fase de interpretação de dados, opta-se pelo emprego de um método de investigação quantitativo ou qualitativo ou misto para o processo de inferência. A abordagem quantitativa refere-se à frequência de aparição de elementos considerados relevantes para o estudo, sejam eles palavras, expressões ou temas. A análise qualitativa, por sua vez, apoia-se nas inter-relações entre as categorias e as unidades de registro, sem contagem de ocorrências. Como já havia feito uma análise estatística na primeira fase da pesquisa, optei pela abordagem qualitativa para análise do material.

Em relação à categorização dos dados, Laville e Dionne (1999) definem as categorias analíticas como dimensões latentes ao redor das quais se organiza o conteúdo das mensagens. Para a definição das categorias de análise, três modelos são propostos: o modelo aberto, no qual as categorias tomam forma no decorrer da análise; o modelo fechado, no qual as categorias são selecionadas a priori com base em um quadro teórico; e o modelo misto que, apesar das categorias estarem preestabelecidas, existe flexibilidade para novas interpretações.

Para este estudo, a AC foi empregada na análise dos dados qualitativos obtidos nos comentários escritos, entrevistas e grupo focal. O material em áudio foi transcrito utilizando-se os símbolos de transcrição propostos por Marcuschi (2008). Os textos foram, então, organizados em unidades de análise temáticas. Para a categorização dos temas, adotei o modelo

misto, tendo como categorias de análise os componentes principais extraídos na fase quantitativa do estudo. Os elementos discrepantes, que não se encaixaram nas categorias selecionadas, foram analisados com foco nos propósitos do estudo e, quando julgados relevantes, agruparam-se formando novas categorias como, por exemplo, a cultura de avaliar na EEAR e a relevância das provas de rendimento.

3.9 Resumo do capítulo

Para investigar o efeito retroativo do EPLIS nas percepções, atitudes e ações de professores e alunos em um contexto de formação de controladores de tráfego aéreo, optei por uma pesquisa de métodos mistos em um modelo sequencial explicativo. Nesse tipo de abordagem, são coletados, primeiramente, dados quantitativos. Os resultados iniciais direcionam a coleta de dados qualitativos. O objetivo dos dados qualitativos são a expansão, o detalhamento e a corroboração dos resultados a fim de compreender melhor a complexidade do fenômeno em estudo. Para a fase quantitativa da pesquisa, foram desenvolvidos dois questionários: um para o aluno e outro para o professor. Para a análise dos dados, empreguei a estatística descritiva, a análise de componentes principais e testes de hipótese para identificar relações de (in)significância entre os componentes extraídos, o nível de conhecimento dos alunos sobre o exame, seus níveis de proficiência e a proximidade do teste. Os componentes principais identificados na ACP guiaram a coleta e a análise dos dados qualitativos, obtidos por meio de comentários ao questionário, observações de aulas semiestruturadas, entrevistas individuais com os professores e instrutor observados, entrevistas com a coordenadora pedagógica e a chefe da Subseção de Idiomas e grupo focal com alunos. Para a análise dos dados produzidos nos comentários, nas entrevistas e no grupo focal foi empregada a análise de conteúdo, tendo como categorias analíticas os componentes principais identificados na fase quantitativa. Para as observações de aulas, foi utilizada uma ficha com áreas possíveis de serem afetadas pelo teste. Os dados qualitativos foram triangulados para fins de validade e confiabilidade dessa etapa da pesquisa.

CAPÍTULO IV

AS PERCEPÇÕES E ATITUDES DE PROFESSORES, INSTRUTORES E ALUNOS: UM ESTUDO QUANTITATIVO

4.1 Considerações iniciais

Neste capítulo, apresento a análise quantitativa dos dados relativos às percepções e atitudes dos professores, instrutores e alunos acerca da influência do EPLIS. Essa fase da pesquisa teve como propósito identificar padrões e tendências entre os participantes. Para a geração de dados, foram aplicados questionários. Primeiramente, exponho a análise dos dados referentes aos professores, coletados, em 2014, durante o estudo preliminar. Discuto, em seguida, os resultados da análise estatística descritiva dos dados dos professores e instrutores, obtidos em 2016, comparando-os àqueles obtidos em 2014, a fim de verificar um possível aumento na intensidade do efeito retroativo após implementação do exame na EEAR. Passo, então, a tratar do efeito retroativo do EPLIS nos alunos. Para a análise dos dados foram empregadas, além da estatística descritiva, técnicas[58] mais sofisticadas como a análise fatorial de componentes principais e os testes de hipóteses. Descrevo, primeiramente, o perfil dos alunos. Em seguida, abordo como foram feitas a extração e a interpretação dos componentes principais. Analiso, então, diferenças estatisticamente significantes entre os componentes extraídos e três fatores: o nível de conhecimento geral sobre o exame, o nível de proficiência declarado pelos alunos e a proximidade à realização do teste. Por fim, faço uma comparação entre as respostas de alunos e professores em relação à influência do EPLIS com o intuito de verificar possíveis discrepâncias.

4.2 O efeito retroativo nos professores antes da implementação do EPLIS

Em 2014, como parte do estudo preliminar conduzido, foi aplicado aos professores de inglês do curso BCT um questionário inicial sobre suas percepções e atitudes a respeito da influência do EPLIS no cenário escolar. Esse questionário foi elaborado em três partes com o

[58] Os testes estatísticos foram conduzidos pela Empresa Estat Júnior de Estatística da Unicamp, com o emprego do software livre RStudio, e verificados por esta pesquisadora por meio do software SPSS (*Statistical Package for the Social Sciences*) da empresa IBM.

objetivo de coletar informações sobre o perfil dos professores, bem como suas percepções e atitudes em relação à influência do exame em diversos aspectos do ensino e aprendizagem de inglês na EEAR. Como explicado no capítulo anterior, por terem sido obtidos antes da implementação do exame para os alunos da 4ª série do curso, esses dados foram considerados de base. Os resultados da análise estatística descritiva dos dados preliminares estão organizados em Perfil dos professores, Intensidade do efeito retroativo e Padrões de respostas.

4.2.1 O perfil dos professores em 2014

Em 2014, a EEAR contava com 20 professores de inglês ministrando aulas para o curso BCT. Desses, 12 professores responderam ao questionário, o que contabilizou um retorno de 60%. Em relação ao sexo, 10 dos 12 respondentes eram mulheres. Em relação à faixa etária, seis professores, ou seja 50%, possuíam entre 31 e 40 anos de idade, cinco professores tinham idade superior a 40 anos e apenas um professor apresentava até 30 anos de idade. Em termos de qualificação profissional, mais da metade, ou seja, sete professores haviam realizado cursos de especialização ou mestrado na área de ensino de línguas. Os demais, cinco professores, possuíam o nível de graduação em Letras − Inglês. O grupo, como um todo, demonstrava extensa experiência no ensino de inglês: nove professores possuíam mais de 10 anos de magistério e três professores alegaram entre 7 e 10 anos de atuação na área de ensino de línguas. Todavia, quando perguntados sobre a experiência como professores de inglês no contexto aeronáutico, esse cenário se modificou bastante. 50%, ou seja, seis professores tinham 2 anos ou menos, quatro professores possuíam entre 3 e 4 anos, e somente dois professores reportaram lecionar inglês aeronáutico há mais de 5 anos.

4.2.2 A intensidade do efeito retroativo

Os professores avaliaram o grau de influência do EPLIS em diversos aspectos do ensino e aprendizagem: currículo, material didático, atividades, metodologia da escola, prática de sala de aula, avaliações de rendimento e estratégias de aprendizagem utilizadas pelos alunos. No Quadro 13, é possível observar as médias, os valores de desvio padrão, bem como os coeficientes de variação das respostas dos professores para cada aspecto analisado. Os níveis de intensidade a serem selecionados foram: 1 (muito baixo), 2 (baixo), 3 (razoável), 4 (bom) e 5 (alto).

Quadro 13 − Efeito retroativo do EPLIS nas percepções dos professores

Aspectos no ensino e aprendizagem de inglês	Média	Desvio Padrão	Coef. Variação	Erro padrão	Erro amostral	Intervalo de Confiança de 95%
Influência do EPLIS no currículo escolar	2,583	1,505	58,20%	0,434	0,85	1,733 a 3,433
Influência do EPLIS no material didático	2,667	1,614	60,50%	0,465	0,911	1,756 a 3,578
Influência do EPLIS na metodologia da escola	2,250	1,215	54%	0,35	0,686	1,981 a 2,936
Influência do EPLIS na prática do professor	2,167	0,937	43%	0,27	0,529	1,638 a 2,696
Influência do EPLIS nas tarefas em sala de aula	2,417	1,165	48,20%	0,336	0,658	1,759 a 3,075
Influência do EPLIS nas avaliações de rendimento	1,750	0,866	49,40%	0,249	0,127	1,623 a 1,877
Influência do EPLIS nas estratégias de aprendizagem dos alunos	1,667	0,778	46,60%	0,224	0,439	1,228 a 2,106
	2,214	**1,154**	**51,41%**	**0,333**	**0,6**	**1,674 a 2,814**

Os dados revelaram que, de forma geral, em 2014, o grau de influência do EPLIS sobre o contexto escolar era ainda relativamente baixo na perspectiva dos professores. A média de resposta desses participantes em todos os aspectos analisados variou entre 1.667 e 2.538. A média geral foi calculada em 2.214, o que corresponde ao critério considerado "baixo" dentro da escala adotada.

Se a intensidade do efeito retroativo é considerada baixa, podemos supor que as condições pontuadas por Hughes (1989,1993) e outros teóricos para a ocorrência dos efeitos não estão sendo atendidas. Uma dessas condições está relacionada à familiaridade daqueles envolvidos no processo para com o exame. Por essa razão, pedi aos professores que também avaliassem o nível de conhecimento geral que possuem sobre o EPLIS e o nível de conhecimento de seus alunos. Na Tabela 4, observamos as respostas dos professores a essas duas perguntas. Em relação ao seu próprio conhecimento sobre o EPLIS, a média (M) das respostas foi calculada em 2.083, considerada "baixa". Em relação ao conhecimento dos alunos sobre o exame, os professores o avaliaram com média ainda menor, 1.750, ou seja, entre os critérios "muito baixo" e "baixo". Essa carência no nível de conhecimento sobre o EPLIS também pode embasar a hipótese de que, em 2014, o efeito retroativo do exame era ainda bastante limitado.

Tabela 4 − Nível de conhecimento dos professores e alunos sobre o EPLIS

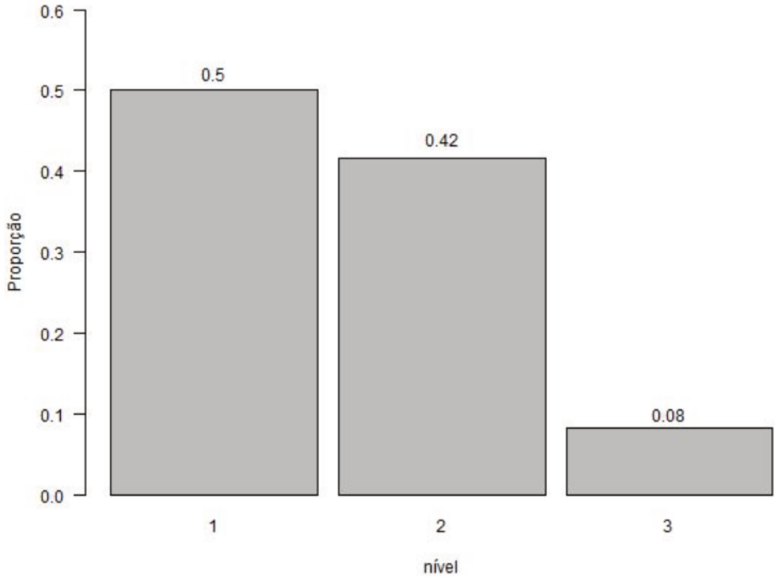
Conhecimento geral sobre o EPLIS	M	DP	CV	Intervalo de Confiança
dos professores	2,083	1,240	59,5%	1,384 a 2,787
dos alunos	1,750	0,965	55,1%	1,749 a 2.294

A relevância de um exame é apontada como outro fator essencial na determinação do efeito retroativo e de sua intensidade (HUGHES, 1989, 1993; CHENG, 2005; GREEN, 2007a). Alderson e Hamp-Lyons (1996) sustentam que quanto maior a relevância de um exame para os seus usuários, maior o seu potencial de exercer mudanças no sistema educacional. Contudo, o conceito de relevância depende mais da percepção dos usuários sobre as

consequências do teste do que das intenções de seus elaboradores ou daqueles envolvidos na elaboração das políticas linguísticas (STONEMAN, 2006). Apesar de o EPLIS ser considerado um exame de alta-relevância para o controle de tráfego aéreo brasileiro, até 2014 ele não era um teste obrigatório para os alunos na EEAR. Ou seja, os alunos só se submetiam ao exame depois de terem concluído o curso, quando já estavam atuavam profissionalmente nos órgãos operacionais. Essa situação nos parece ter causado a diminuição da percepção tanto dos professores quanto dos alunos sobre sua relevância, o que pode ter contribuído para a falta de expressividade de seu efeito retroativo.

A pressão exercida por um exame em seus usuários também pode ser levada em consideração para se avaliar o seu grau de relevância. Sobre isso, cerca de 92% dos professores utilizaram os critérios "muito pouco" (1) e "pouco" (2) para descrever o quão pressionados eles se sentiam em relação ao EPLIS e apenas 8% classificaram a pressão como "razoável" (3), como podemos perceber no Gráfico 1.

Gráfico 1 − Pressão exercida pelo EPLIS nos professores



Ademais, quando perguntados se seus alunos modificaram suas estratégias de aprendizagem por razão do EPLIS, as respostas dos professores alcançaram média 1.667 (Quadro 13). Nos comentários, alguns professores alegaram que os alunos priorizavam o curso de formação: *porque os alunos veem sempre a graduação e colocam-na como prioridade* (professor 3, questionário inicial). Outros apontaram a falta de conscientização dos alunos como um dos problemas: *Acredito que, em geral, os alunos precisam se conscientizar da importância do inglês em seu trabalho* (professor 7, questionário inicial). Um professor justificou que os

alunos não modificaram seu modo de aprender por falta de conhecimentos sobre o exame: *Não, porque pouquíssimos conhecem o exame* (professor 1, questionário inicial). Com base nesses dados, presumimos que a baixa intensidade do efeito retroativo do EPLIS, nessa época, pode estar ligada ao pouco conhecimento que professores e alunos possuíam, até aquele momento, sobre o exame, acrescida de sua baixa relevância no contexto investigado.

4.2.3 Padrões de resposta

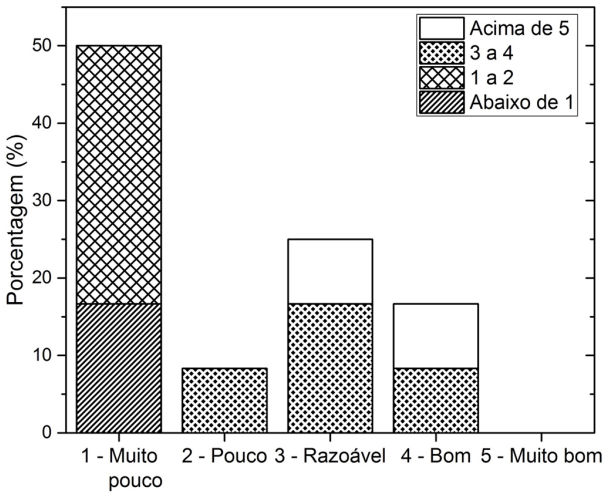
Ao voltarmos ao Quadro 13, percebemos que o coeficiente de variação[59] (CV) dos dados oscilou entre 43% e 60%. Valores de CV acima de 25% sugerem uma amostra bastante heterogênea. Esse índice sustenta a hipótese de Alderson e Wall (1993) de que o efeito retroativo de um exame pode ocorrer para alguns professores e não para outros, principalmente, em contextos de baixa-relevância. Sabemos, porém, que, apesar de os professores agirem diferentemente à influência de um exame, suas respostas podem revelar certos padrões e similaridades (BURROWS, 2004). Examinei, então, as respostas dos professores à procura de modelos e dois subgrupos distintos de participantes emergiram: aqueles com 2 anos ou menos de experiência no ensino de inglês aeronáutico, representados por 50% dos professores (subgrupo 1) e aqueles que afirmaram trabalhar com inglês aeronáutico há 3 anos ou mais (subgrupo 2). Esses dois subgrupos apresentaram respostas semelhantes entre si em relação ao nível de conhecimento sobre o EPLIS e à influência do exame na prática pedagógica, nas atividades realizadas em sala de aula e nas avaliações de rendimento.

4.2.3.1 Conhecimento sobre o EPLIS

Todos os professores, pertencentes ao subgrupo 1, com pouca experiência no ensino de inglês aeronáutico, avaliaram o nível de conhecimento geral sobre o exame em "muito baixo", ao passo que 84% dos professores mais experientes, subgrupo 2, avaliaram-no entre "razoável" e "bom", como podemos observar na Figura 6.

[59] O coeficiente de variação (CV), ou desvio padrão relativo, indica a dispersão dos dados em relação a sua média. Geralmente expresso em porcentagem, o CV é calculado ao se dividir o desvio padrão pela média e multiplicá-lo por 100. Essa medida nos auxilia a identificar se o desvio padrão pode ser considerado grande ou pequeno. Quanto menor for o valor de CV, mais homogêneos serão os dados. Valores iguais ou menores que 25% são considerados baixos. Valores acima indicam uma amostra de dados heterogênea.

Figura 6 – Experiência no ensino de inglês aeronáutico e conhecimento sobre o exame

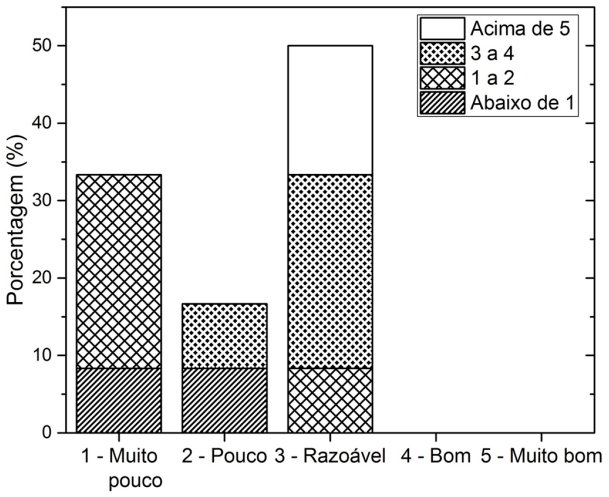


Esses resultados apontam uma possível relação direta entre a experiência no ensino de inglês aeronáutico e o conhecimento geral sobre o EPLIS. Em outras palavras, quanto maior a experiência do professor no ensino de inglês aeronáutico, maior o seu nível de conhecimento geral sobre o exame. Portanto, ao assumirmos que o subgrupo 2 possui mais conhecimento sobre o EPLIS e sobre a natureza da linguagem avaliada pelo teste – inglês aeronáutico, acreditamos que o efeito retroativo nas percepções e atitudes dos professores pertencentes a esse subgrupo seja relativamente mais intenso em comparação ao subgrupo 1, o que discuto em seguida.

4.2.3.2 Influência do exame em sala de aula

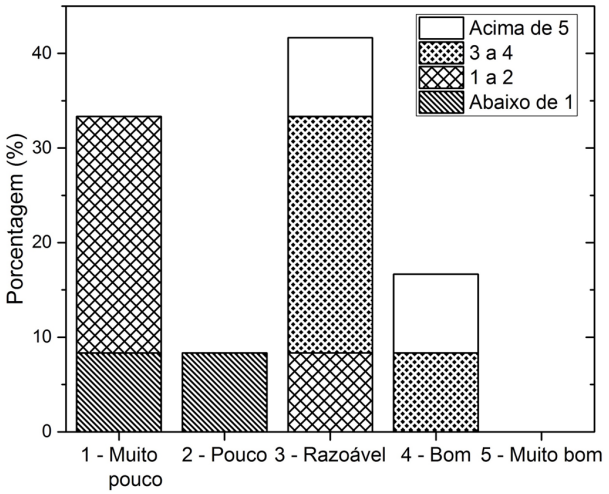
Em relação à influência do EPLIS na prática pedagógica, todos os professores avaliaram o impacto entre “muito baixo” e “razoável”, com um valor médio de resposta em 2.167 (Quadro 13). Embora esse resultado indique que o efeito retroativo era bastante limitado na prática do professor de forma geral, os dois subgrupos de professores divergiram em suas respostas. No subgrupo 1, 84% responderam “pouco” e “muito pouco”, ao passo que a mesma proporção do subgrupo 2 julgou o impacto como “razoável”, ou seja, um pouco mais intenso, como pode ser observado na Figura 7.

Figura 7 − Experiência no ensino de inglês aeronáutico e a prática pedagógica



Algo semelhante ocorreu em relação à influência do exame nas atividades realizadas em sala de aula. Podemos observar que 100% dos professores do subgrupo 2 consideraram o impacto entre "razoável" e "bom", enquanto, 84% do subgrupo 1 escolheram os critérios "pouco" ou "muito pouco" (Figura 8).

Figura 8 − Experiência no ensino de inglês aeronáutico e as atividades em sala de aula



Foi também pedido aos professores que descrevessem as mudanças ocorridas em sua prática pedagógica em decorrência do teste. O subgrupo 1, que avaliou o impacto nas atividades em sala de aula como "baixo", não respondeu a essa pergunta alegando não ter

conhecimentos suficientes sobre o EPLIS. Diferentemente, alguns professores do subgrupo 2 disseram ter percebido mudanças no modo como eles trabalhavam com os alunos:

> *Sim. Eu tenho dado mais tarefas para desenvolver a fluência, descrever figuras e narrar histórias* (Professor 1, questionário inicial).
>
> *Sim. Porque agora eu peço para os meus alunos descreverem figuras e contarem histórias com mais frequência do que eu fazia no passado, antes de eu saber que essas tarefas eram parte do EPLIS.* (Professor 2, questionário inicial).
>
> *Sim. Eu tento dar para os alunos tarefas como descrever figuras, recontar uma história ou áudio, solucionar problemas e dar opiniões.* (Professor 5, questionário inicial).

Os tipos de atividades mencionadas pelos professores do subgrupo 2 são similares às tarefas encontradas na fase II do EPLIS, reforçando a hipótese de que o efeito retroativo do EPLIS é um pouco mais intenso para os professores pertencentes a esse subgrupo, ou seja, para aqueles com mais experiência no ensino de inglês aeronáutico e maior nível de conhecimento geral sobre o exame.

4.2.3.3 Influência do exame nas avaliações de rendimento

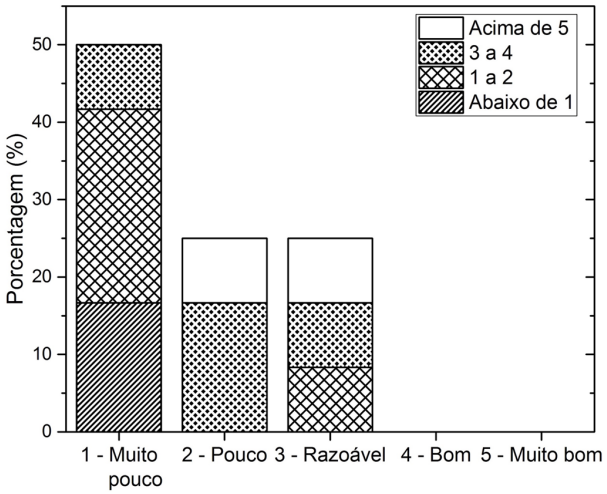
Os professores consideraram, em geral, baixo o impacto do exame no processo de avaliação em sala de aula, com média das respostas calculada em 1.750 (Quadro 13). Em relação aos dois subgrupos, 84% do subgrupo 1 optou pelo critério "muito pouco" e 84% do subgrupo 2 marcou entre "pouco" e "razoável" (Figura 9).

Nos comentários, o subgrupo 2 argumentou que, em termos de compreensão oral, os itens das avaliações em sala de aula eram similares àquelas encontradas na fase I do exame, mas não havia, até aquele momento, nenhuma prova que avaliasse a produção oral dos alunos. Esse posicionamento dos professores também nos faz supor que por conhecerem um pouco mais sobre o exame, eles teriam melhores condições de analisar a influência do exame no ensino e aprendizagem.

Em suma, os resultados alcançados ofereceram um panorama sobre as percepções dos professores em relação ao efeito retroativo do EPLIS em diversos aspectos do ensino e aprendizagem de inglês, antes da implementação do exame no contexto investigado e, por isso, são considerados, aqui, dados de base. Dois fatores pareceram mediar a intensidade do efeito retroativo: a experiência no ensino de inglês aeronáutico e o conhecimento geral sobre o teste. Professores com mais tempo de experiência no ensino de inglês aeronáutico apresentaram

maior nível de conhecimento sobre o EPLIS, afirmaram trabalhar tarefas similares às do exame com seus alunos e consideraram como razoável a influência do exame em sua maneira de ensinar. Apesar de o efeito retroativo ter ocorrido com mais intensidade para esse subgrupo de professores, podemos dizer que, de modo geral, antes da implementação do exame, o efeito era fraco nas percepções e atitudes desses professores.

Figura 9 − Experiência no ensino de inglês aeronáutico e as avaliações em sala de aula



4.3 O efeito retroativo nos professores após a implementação do EPLIS

A fim de investigar a influência do exame nas percepções e atitudes dos professores após a implementação do exame no contexto escolar, analisei estatisticamente as respostas dos participantes ao questionário principal, com o intuito de descrever tendências e padrões presentes nos dados. A parte I do questionário coletou informações sobre o perfil desses participantes. As partes II, III, IV e V formaram os quatro construtos (temas) analisados: Conhecimento sobre o EPLIS, Impacto do EPLIS no ensino, Preparação dos alunos para o exame e Visões e percepções sobre o exame. Além do grupo de professores que participou do estudo preliminar em 2014, foram também incluídos aqui os controladores de tráfego aéreo que passaram a atuar como instrutores de inglês aeronáutico na EEAR em 2016. Os resultados estão organizados de acordo com os temas abordados no questionário: Perfil dos participantes, Conhecimentos sobre o EPLIS, Impacto do EPLIS no ensino, Preparação dos alunos para o EPLIS e Percepções e visões sobre o EPLIS.

4.3.1 O perfil dos professores e instrutores em 2016

Ao longo de 2016, 21 professores de inglês aeronáutico (professores e instrutores) atuaram no curso BCT. 16 profissionais responderam ao questionário, totalizando uma taxa de retorno de, aproximadamente, 76%. Desses, 25% eram controladores de tráfego aéreo (quatro respondentes) e 75% (12 respondentes) eram professores com graduação em Letras − Inglês. 11 participantes declararam-se do sexo feminino e 5 do sexo masculino.

Em relação à faixa etária, sete respondentes (44%) apresentaram idade entre 31 e 40 anos, quatro (25%) declararam ter mais de 50 anos, três (19%) possuíam entre 20 e 30 anos e dois (12%) encontravam-se na faixa entre 41 e 50 anos de idade. Quando perguntados sobre a maior qualificação profissional na área de ensino de línguas, dos 12 professores, cinco responderam ter realizado curso de especialização ou pós-graduação lato sensu, quatro disseram possuir o título de mestre, e apenas três disseram possuir o título de graduação. Entre os controladores de tráfego aéreo, três disseram ter realizado cursos de preparação de professores do tipo *TEFL*[60] e um alegou não possuir nenhum tipo de capacitação na área.

Sobre o tempo de experiência no ensino de inglês, a grande maioria, 69% (11 participantes), declarou possuir mais de 10 anos e apenas um respondente (6%) afirmou ter menos de 1 ano de magistério. Porém, em relação à experiência no ensino de inglês aeronáutico, apenas 12% dos participantes declararam estar atuando nessa área há mais de 10 anos. A maioria deles, 38%, selecionou a resposta "entre 3 e 5 anos", como podemos observar no Gráfico 2.

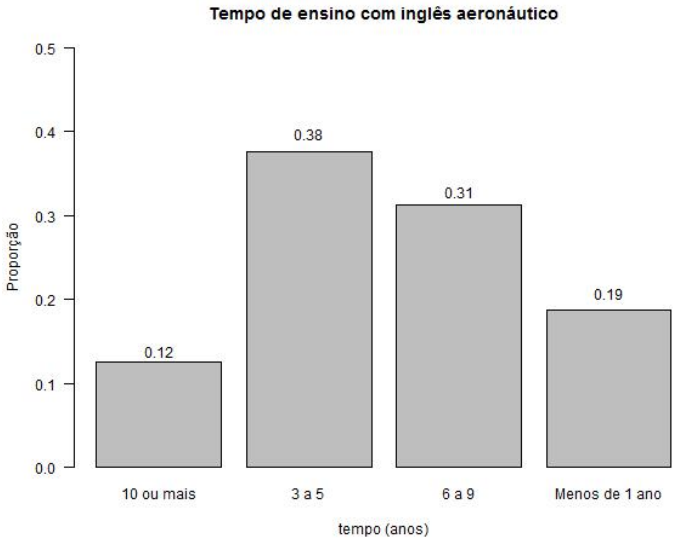
Ao serem perguntados sobre o nível de conhecimento em inglês aeronáutico, a imensa maioria, representada por 75%, respondeu "bom" ou "ótimo" e 19% disseram ainda possuir pouco conhecimento. Sobre as séries para as quais lecionam, pudemos notar que nenhum dos professores ou instrutores costuma atuar em todas as quatro séries do curso. A maior parte, 75%, trabalha em duas ou três séries, enquanto o restante, 25%, ministra aulas apenas para uma série. Dos que atuam em apenas uma das séries, três afirmaram dar aulas somente para a 4ª série e um professor disse trabalhar apenas com a 1ª série do curso BCT.

Como era de se esperar, em comparação ao grupo de 2014, é percebida diferença no tempo de experiência no ensino de inglês aeronáutico. Em 2014, 50% dos professores

[60] *TEFL* é uma sigla em inglês para designar um certificado de ensino de inglês como segunda língua. É, geralmente, um curso curto, com 3 a 4 semanas de duração, voltado para aqueles que querem iniciar uma carreira como professores de inglês.

possuíam 2 anos ou menos de atuação no ensino de inglês aeronáutico, ao passo que, no grupo de 2016, esse percentual foi reduzido para 19%.

Gráfico 2 − Experiência no ensino de inglês aeronáutico



4.3.2 Construto 1 – Conhecimentos sobre o EPLIS

O construto 1, denominado *Conhecimentos sobre o EPLIS*, é representado pela parte II do questionário do professor e foi elaborado com o objetivo de investigar as facetas que envolvem o conhecimento do professor sobre diversos aspectos do exame. É formado por 10 itens/variáveis. As questões de 10 a 17 abordam assuntos ligados aos conhecimentos dos professores e instrutores em relação aos requisitos de proficiência da OACI (Doc 9835, escala de níveis de proficiência, nível mínimo exigido), conhecimento geral sobre o exame (objetivo) e conhecimento específico (conteúdo, tipos de tarefas e manual do candidato). As questões 18 e 19 avaliam a frequência com que o profissional participa de eventos sobre o EPLIS e consulta o *site* do exame para obter informações mais específicas.

Na Tabela 5, podemos observar as médias de respostas dos professores e instrutores às questões de 10 a 17, na qual N representa o número total de participantes, M refere-se à média de respostas, com variação de 1 (nenhum) a 5 (ótimo), DP é a abreviação usada para desvio padrão, e EPM significa o erro padrão da média. Verificamos que os professores e instrutores avaliaram de forma bastante positiva os conhecimentos que possuem sobre o exame. Com exceção do item relacionado ao Doc 9835 AN/453, todos os demais itens atingiram média superior a 4, que representa o critério “bom” no questionário. Quanto ao conhecimento sobre o

Doc 9835 AN/453 que estabelece os requisitos de proficiência linguística da OACI, a média de respostas dos participantes ficou próxima a 3, considerada “razoável”.

Tabela 5 − Conhecimentos dos professores sobre o EPLIS

	N	M	DP	EPM
Geral	16	4.000	0.966	0.241
Doc 9835 AN/453	16	3.187	1.167	0.291
Objetivo do exame	16	4.187	0.834	0.208
Conteúdo do exame	16	4.065	0.853	0.213
Tarefas – Fase I	16	4.065	0.928	0.232
Tarefas – Fase II	16	4.065	0.997	0.249
Tabela de níveis	16	4.125	0.885	0.221
Desempenho nível 4	16	4.062	0.853	0.213

Em relação ao nível de conhecimento geral sobre o exame em porcentagem, as respostas dos professores e instrutores concentraram-se nos níveis “bom” (51%) e “ótimo” (31%), mostrando um cenário bastante distinto daquele de 2014, quando 58% dos professores afirmaram possuir pouco ou nenhum conhecimento sobre o teste. Isso sugere que houve um aumento no nível de familiaridade dos participantes com o EPLIS, o que podemos relacionar à implementação do exame no contexto escolar e à melhoria em sua divulgação.

Devido ao fato de que, no ano de 2016, a EEAR passou a contar com a participação de controladores de tráfego aéreo como instrutores de inglês aeronáutico para o curso BCT, decidi avaliar como professores e instrutores percebem a influência do exame. A Tabela 6 apresenta as diferenças entre esses dois grupos de participantes em relação aos conhecimentos que possuem sobre o EPLIS. O subgrupo 1 representa os professores com licenciatura em Letras e o subgrupo 2 representa os controladores de tráfego aéreo convidados a atuarem como instrutores de inglês aeronáutico.

Tabela 6 − Comparação entre professores e instrutores em relação aos conhecimentos sobre o EPLIS

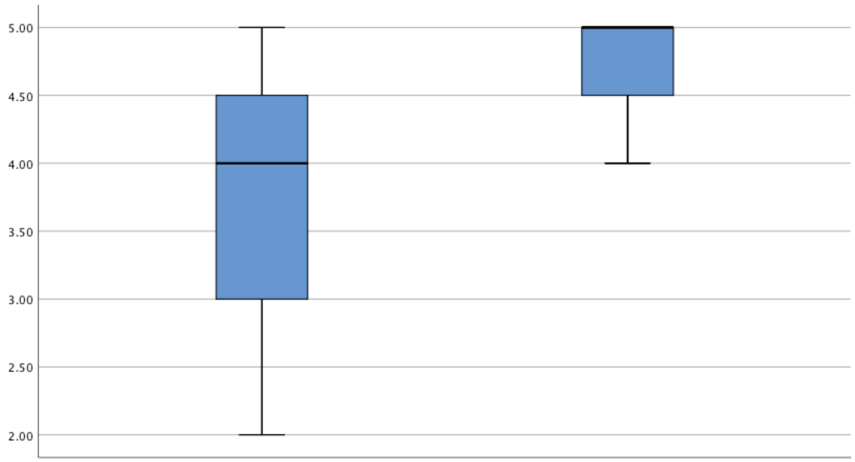
	Subgrupos	N	M	DP
Geral	1	12	3.833	1.029
	2	4	4.500	0.577
Doc 9835 AN/453	1	12	3.083	1.311
	2	4	3.500	0.577
Objetivo do exame	1	12	4.000	0.852
	2	4	4.750	0.500
Conteúdo do exame	1	12	3.916	0.900
	2	4	4.500	0.577
Tarefas – Fase I	1	12	3.833	0.937
	2	4	4.750	0.500
Tarefas – Fase II	1	12	3.833	1.029
	2	4	4.750	0.500
Tabela de níveis	1	12	3.916	0.900
	2	4	4.750	0.500
Desempenho nível 4	1	12	3.916	0.900
	2	4	4.500	0.577

Podemos constatar que as médias de respostas dos controladores de tráfego aéreo (subgrupo 2) são mais altas em todos os itens analisados, quando comparadas às médias dos professores. Enquanto as respostas dos professores variaram entre 3.083 e 4.000 (entre os critérios “razoável” e “bom”), as médias dos instrutores flutuaram entre 3.500 e 4.750, sendo que somente no que se refere ao Doc 9835 AN/453 essa média ficou abaixo de 3, indicando que o nível de conhecimentos desse subgrupo sobre o EPLIS varia de “bom” a “ótimo”. Já, o subgrupo dos professores apresenta valores de desvio padrão (medida de dispersão em torno da média) mais elevados, apontando maior dispersão e heterogeneidade nas respostas desses participantes.

Os gráficos 3 e 4, em box plot, exemplificam essa variação nas respostas do subgrupo 1 quanto ao conhecimento das tarefas da fase I e das tarefas da fase II do EPLIS, respectivamente. No eixo vertical, estão representados os níveis de conhecimento atribuídos pelos participantes de 1 (nenhum) a 5 (ótimo) e, no eixo horizontal à esquerda está o subgrupo 1 e à direita o subgrupo 2.

No Gráfico 3, é possível observar que os dados do subgrupo 1 (à esquerda) se dispersam mais, com respostas variando entre 2 e 5. No subgrupo 2 (à direita), os dados estão mais aglutinados e as respostas mostram pequena flutuação (entre 4 e 5). Os retângulos marcam a concentração das respostas e a linha horizontal preta no retângulo informa a simetria dos dados. Em outras palavras, no subgrupo 1 as respostas concentraram-se no nível 4, enquanto que, no subgrupo 2, as respostas concentraram-se no nível 5.

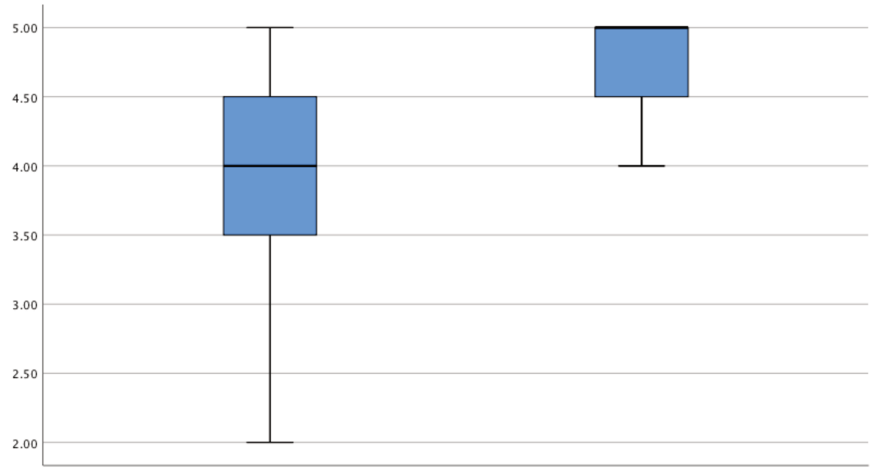
Gráfico 3 − Dispersão das respostas dos subgrupos 1 e 2 em relação às tarefas da fase I do EPLIS



Algo parecido ocorreu em relação às tarefas da fase II do exame, como pode ser observado no Gráfico 4. É possível notar, no subgrupo 1 (professores, à esquerda), maior

dispersão dos dados, com respostas entre 2 e 5. Essa dispersão é diminuída no subgrupo 2, com valor mínimo de 4 e máximo de 5, indicando um subgrupo mais homogêneo. No caso do subgrupo 1, enquanto a maioria das respostas dos professores ficou ao redor de 4, verificamos algumas respostas no nível 2. O mesmo não aconteceu dentro do subgrupo 2. Esses resultados sugerem que, embora tenha havido uma melhora expressiva em relação a 2014, no subgrupo de professores, ainda existem casos de profissionais que consideram possuir poucos conhecimentos sobre o EPLIS. Podemos presumir que os professores que afirmaram ter menos conhecimentos sobre o exame são aqueles com pouco tempo de experiência no ensino de inglês aeronáutico.

Gráfico 4 − Dispersão das respostas dos subgrupos 1 e 2 em relação às tarefas da fase II do EPLIS



Por essa razão, realizei a comparação entre o tempo de experiência no ensino de inglês aeronáutico e o nível de conhecimentos sobre o EPLIS. Recordo que, no estudo preliminar, constatamos evidências de que o tempo de experiência no ensino de inglês aeronáutico influenciou positivamente o grau de conhecimento geral sobre o exame e, consequentemente, a intensidade do efeito retroativo. Para isso, dividi os participantes, professores e instrutores, em dois subgrupos: subgrupo 1 (professores e instrutores com até 5 anos de experiência no ensino de inglês aeronáutico) e subgrupo 2 (professores e instrutores com mais de 6 anos de experiência). Essa divisão foi pautada nos parâmetros do estudo preliminar, cujo subgrupo considerado menos proficiente apresentava até 2 anos de experiência e o subgrupo mais proficiente possuía 3 anos ou mais no ensino de inglês aeronáutico. Considerando que, aproximadamente, 3 anos decorreram entre a aplicação do questionário preliminar e a aplicação do questionário principal, o primeiro grupo estaria, portanto, com até

5 anos de experiência no ensino de inglês aeronáutico, e o segundo com 6 anos ou mais. Os resultados dessa análise encontram-se dispostos na Tabela 7.

Tabela 7 − Comparação entre experiência no ensino de inglês aeronáutico e conhecimentos sobre o EPLIS

	Subgrupos	**N**	**M**	**DP**
Geral	1	9	3.777	1.201
	2	7	4.285	0.487
Doc 9835 AN/453	1	9	2.777	1.201
	2	7	3.714	0.951
Objetivo do exame	1	9	4.111	1.054
	2	7	4.285	0.487
Conteúdo do exame	1	9	4.000	1.118
	2	7	4.142	0.377
Tarefas – Fase I	1	9	3.888	1.666
	2	7	4.285	0.487
Tarefas – Fase II	1	9	3.888	1.269
	2	7	4.285	0.487
Tabela de níveis	1	9	4.000	1.118
	2	7	4.285	0.487
Desempenho nível 4	1	9	3.888	1.054
	2	7	4.062	0.487

Os valores das médias indicam que o subgrupo 2 (com mais de seis anos de experiência) possui níveis mais elevados de conhecimentos em todos os itens analisados. A diferença é mais expressiva em relação ao conhecimento geral do exame, do Doc 9835 AN/453 da OACI, bem como das tarefas das fases I e II. No entanto, percebemos novamente valores mais altos de desvio padrão no subgrupo 1, indicando grande dispersão nas respostas desses participantes. Esses resultados mostram que, embora a maioria dos participantes tenha avaliado o seu nível de conhecimentos sobre o exame positivamente, esses conhecimentos não se encontram totalmente consolidados e ainda existem lacunas a serem preenchidas, principalmente, no subgrupo com profissionais menos experientes.

Quanto à capacitação dos professores e instrutores em relação ao EPLIS, a Tabela 8 resume os resultados das respostas dos participantes às questões 18 e 19 do questionário. O item 18 aborda a participação em palestras sobre o exame, e o item 19 avalia a frequência com que eles acessam o *site* do EPLIS em busca de informações. Constatamos médias inferiores a 3 (“às vezes”) para os dois itens, com maior variação de respostas no item “Acesso ao *site* do exame” (DP=1.289).

Tabela 8 − Oportunidades de capacitação

	N	**M**	**DP**	**EPM**
Participação em palestras e oficinas	16	2.562	0.629	0.157
Acesso ao *site* do exame	16	2.937	1.289	0.322

A Tabela 9 ajuda-nos a visualizar melhor essa situação ao distribuir as respostas dos professores e instrutores em porcentagem. Vemos que 50% dos professores e instrutores afirmaram que raramente participam de seminários e oficinas sobre o exame. Em relação ao acesso ao *site*, as respostas variaram mais, com 57% entre "nunca" e "às vezes", e 43% entre "frequentemente" e "sempre".

Tabela 9 − Oportunidades de capacitação em porcentagem

	1	**2**	**3**	**4**	**5**
Participação em palestras e oficinas		0.50	0.43	0.07	
Acesso ao *site* do exame	0.19	0.19	0.19	0.37	0.06

Ao comparar professores (subgrupo 1) e controladores de tráfego aéreo (subgrupo 2) em relação a esses dois itens, vemos que a situação é praticamente a mesma para ambos os subgrupos, como mostra a Tabela 10. As médias são muito próximas, entre 2.500 e 3.000, porém notamos maior variação (desvio padrão) no item "Acesso ao **site**" tanto no subgrupo de professores quanto no subgrupo de instrutores.

Tabela 10 − Comparação entre professores e instrutores em relação às oportunidades de capacitação

	Subgrupos	**N**	**M**	**DP**
Participação em palestras e oficinas	1	12	2.583	0.668
	2	4	2.500	0.577
Acesso ao *site* do exame	1	12	2.916	1.443
	2	4	3.000	0.816

Verifiquei, então, se havia relação entre essa variação e a experiência no ensino de inglês aeronáutico. Relembro que o subgrupo 1 é formado por profissionais com até cinco anos de experiência e o subgrupo 2 com mais de seis anos (Tabela 11).

Tabela 11 − Comparação entre experiência no ensino de inglês aeronáutico e acesso ao *site* do exame

	Subgrupos	**N**	**M**	**DP**
Acesso ao site do exame	1	9	2.888	1.269
	2	7	3.000	1.414

Podemos observar que os dois subgrupos apresentaram diferença mínima nas médias de respostas em relação à frequência com que acessam o *site* do exame, o que nos leva a crer que o quê distingue esses dois subgrupos nesse item não é o tempo de experiência no ensino de inglês aeronáutico, mas outro fator que pode estar relacionado à disposição, dedicação, profissionalismo e comprometimento desses profissionais.

Temos, portanto, um contrassenso. Embora a grande maioria dos professores e instrutores tenha avaliado positivamente os conhecimentos que possuem sobre o exame, a frequência com que disseram participar de eventos e buscar informações é mais baixa. Essa situação levanta questões que precisam ser analisadas mais cuidadosamente, por exemplo, como esses profissionais constroem conhecimentos sobre o EPLIS, qual a contribuição da escola e do ICEA nesse processo, e por que o *site* do exame é pouco acessado.

As discrepâncias encontradas nos dados em relação aos conhecimentos que os participantes possuem sobre o EPLIS, principalmente, no grupo dos professores e dos profissionais menos experientes, juntamente com o fato de que a busca por informações e a frequência de eventos de capacitação terem sido baixas, podem interferir negativamente na intensidade do efeito retroativo do EPLIS. Isso porque esses óbices ferem as condições estabelecidas por Hughes (1989, 1993) e previstas no modelo de efeito retroativo de Green (2007a) de que a familiaridade dos participantes para com o exame, bem como a disponibilidade de recursos e assistência, são fundamentais para que os efeitos ocorram em plenitude.

4.3.3 Construto 2 – Impacto do EPLIS no ensino

O construto 2, *Impacto do EPLIS no ensino*, refere-se à parte III do questionário do professor formada pelas questões de 20 a 34, totalizando 15 variáveis observadas. O propósito foi analisar as percepções e atitudes dos professores e instrutores quanto ao efeito retroativo do exame no ensino de inglês no curso BCT. Os itens de 20 a 25 avaliaram o impacto do exame na metodologia, no conteúdo abordado em sala de aula, no material didático, nas tarefas realizadas, nos projetos desenvolvidos e nas avalições de rendimento. As questões de 26 a 29, por sua vez, abordaram a frequência com que professores e instrutores mencionam e discutem aspectos do exame e tarefas em sala de aula, assim como a frequência com que os alunos os procuram para tirar dúvidas sobre o teste. Os itens de 30 a 34 contemplaram o grau com que as atividades em sala de aula estão relacionadas a temas da aviação e do controle de tráfego aéreo.

As respostas aos itens de 20 a 25 podem ser verificadas na Tabela 12. Pelo fato de essas questões terem sido elaboradas em uma escala de concordância, a visualização por porcentagem torna os dados mais facilmente interpretáveis. O nível 1 representa a resposta "discordo totalmente", 2 "discordo", 3 "nem concordo nem discordo", 4 "concordo" e 5 "concordo totalmente". Percebemos que a maior parte dos professores e instrutores concordou ou concordou totalmente com as afirmações feitas acerca da influência do exame em todos os aspectos analisados. Esses resultados são, novamente, bastante distintos daqueles coletados em

2014, sugerindo uma mudança de posicionamento dos participantes em relação à influência do exame.

Tabela 12 − Influência do EPLIS na sala de aula em porcentagem

	1	**2**	**3**	**4**	**5**
Metodologia do professor	0.06	0.19	0.00	0.62	0.12
Conteúdo em sala de aula	0.00	0.19	0.06	0.62	0.12
Material didático	0.00	0.25	0.12	0.56	0.06
Tarefas	0.00	0.00	0.06	0.75	0.19
Projetos	0.00	0.06	0.06	0.69	0.19
Avaliações	0.06	0.12	0.25	0.56	0.00

Embora o questionário de 2014 tenha sido mais curto e mais simples do que o questionário de 2016, alguns itens se sobrepuseram, permitindo uma comparação mais específica entre esses dois momentos. Na Tabela 13, podemos contrastar os dados de 2014 e 2016 em relação à influência do exame na abordagem de ensino do professor, no material didático, nas tarefas e avaliações em sala de aula.

Tabela 13 − Comparação da influência do EPLIS na sala de aula em 2014 e 2016

	Ano	**1**	**2**	**3**	**4**	**5**
Metodologia do professor	2016	0.06	0.19	0.00	0.62	0.12
	2014	0.33	0.17	0.50	0.00	0.00
Material didático	2016	0.00	0.25	0.12	0.56	0.06
	2014	0.42	0.00	0.25	0.17	0.17
Tarefas	2016	0.00	0.00	0.06	0.75	0.19
	2014	0.33	0.08	0.42	0.17	0.00
Avaliações	2016	0.06	0.12	0.25	0.56	0.00
	2014	0.50	0.25	0.25	0.00	0.00

Em 2014, 33% dos professores afirmaram que o EPLIS influenciava pouco ou não influencia o seu modo de ensinar. Em 2016, apenas 6% discordaram do impacto do exame em suas aulas. Em relação ao material didático, em 2014, 42% percebiam pequena ou nenhuma influência do exame. Já em 2016, nenhum respondente discordou da afirmação de que o exame influencia o material didático utilizado. Também podemos perceber um impacto maior do exame nas tarefas em sala de aula: em 2014, 33% dos professores disseram que o EPLIS influenciava pouco ou não influencia os tipos de atividades desenvolvidas com os alunos, percentual esse que foi reduzido a zero em 2016. Quanto às avaliações de rendimentos, em 2014, 50% dos professores disseram perceber um pequeno ou nenhum efeito do exame. Em 2016, somente 18% discordaram que o EPLIS influencia as provas em sala de aula. Esses dados sugerem que o efeito retroativo do exame se intensificou após sua implementação no contexto escolar.

Contudo, embora seja visível esse aumento nas percepções e atitudes dos participantes quanto à intensidade do efeito em sala de aula, índices de discordância ainda foram reportados em 2016, sugerindo que o efeito retroativo do exame ainda é fraco para alguns profissionais. Os itens com maior porcentagem de discordância foram a metodologia do professor e o material didático, com 25% de respostas discordantes em cada um. As avaliações de rendimento, por sua vez, receberam 25% de respostas no nível "nem concordo nem discordo".

Comparei, novamente, as respostas dos professores (subgrupo 1) e dos instrutores de inglês aeronáutico (subgrupo 2). Porém, não foram encontradas diferenças expressivas entre as médias dos dois subgrupos. Já em relação aos profissionais com mais ou menos experiência no ensino de inglês aeronáutico, percebemos diferenças maiores nas médias de respostas aos itens "Metodologia do professor" e "Conteúdo em sala de aula". As médias nos demais aspectos são relativamente próximas, como podemos notar na Tabela 14.

Tabela 14 − Comparação entre experiência no ensino de inglês aeronáutico e influência do EPLIS em sala de aula

	Subgrupos	**N**	**M**	**DP**
Metodologia do professor	1	9	4.000	0.866
	2	7	3.000	1.290
Conteúdo em sala de aula	1	9	4.000	0.866
	2	7	3.287	0.951
Material didático	1	9	3.444	1.013
	2	7	3.428	0.957
Tarefas	1	9	4.111	0.600
	2	7	4.142	0.377
Projetos	1	9	4.111	0.600
	2	7	3.857	0.899
Avaliações	1	9	3.222	0.833
	2	7	3.428	1.133

Nos dois primeiros itens, enquanto o subgrupo mais experiente atingiu média próxima a 3, o subgrupo menos experiente apresentou média 4, sugerindo que os profissionais menos experientes são mais influenciados pelo exame em relação ao modo como ensinam e ao conteúdo abordado em sala de aula. Interessante notar que os participantes com mais experiência pertencem ao subgrupo que atingiu maior nível de conhecimentos sobre o exame e, ao mesmo tempo, veem uma influência um pouco menor do teste em sua prática pedagógica. Isso pode ter relação com o fato de os mais experientes estarem mais conscientes sobre as desvantagens de uma preparação voltada exclusivamente para o formato do exame (*teaching to the test*) e, de certa forma, tentam evitar que isso ocorra em suas aulas.

Investiguei também a maneira como o exame é abordado em sala de aula por meio das respostas às questões de 26 a 29. A Tabela 15 resume os resultados em relação à frequência

com que o EPLIS é mencionado, discutido e trabalhado por professores e instrutores e à frequência com que eles são procurados para tirar dúvidas a respeito do teste.

Tabela 15 − Abordagem do exame em sala de aula

	N	M	DP	EPM
Menção ao exame	16	3.937	0.853	0.213
Discussão de aspectos do exame	16	3.625	0.957	0.239
Trabalho envolvendo tarefas do exame	16	3.312	1.195	0.298
Dúvidas dos alunos	16	2.875	1.024	0.256

Verificamos que a "menção ao exame" é o item com a maior média de respostas, seguido do item "discussão de aspectos do exame", ambos com médias acima de 3. A frequência com que os alunos procuram os professores acerca do exame, no entanto, apresenta média inferior (2.875). Uma variação mais expressiva é encontrada nas respostas a esse item e ao relacionado às tarefas do exame em sala de aula.

A Tabela 16 apresenta a distribuição das respostas por porcentagem, na qual 1 significa "nunca", 2 "raramente", 3 "às vezes", 4 "frequentemente" e 5 "sempre". Notamos que 75% dos professores e instrutores avaliaram a frequência com que citam o exame em sala de aula entre "frequentemente" e "sempre". No entanto, essa proporção diminui em relação à discussão de aspectos do exame e tarefas com os alunos. Em relação às dúvidas dos alunos, as respostas concentram-se no nível "às vezes". Esses dados nos fazem supor que os alunos podem não estar tão interessados no exame quanto previámos.

Tabela 16 − Abordagem do exame em sala de aula em porcentagem

	1	2	3	4	5
Menção ao exame	0.00	0.06	0.19	0.50	0.25
Discussão de aspectos do exame	0.00	0.12	0.31	0.38	0.19
Trabalho envolvendo tarefas do exame	0.12	0.06	0.31	0.38	0.12
Dúvidas dos alunos	0.06	0.25	0.56	0.12	0.12

As médias de respostas dos professores (subgrupo 1) e instrutores (subgrupo 2) divergem, principalmente, em relação à discussão de aspectos do exame e, especialmente, no que se refere a dúvidas dos alunos, como podemos ver na Tabela 17. Esses resultados indicam que os professores costumam discutir com os alunos aspectos do exame com maior frequência do que os instrutores, porém os últimos são mais procurados para sanar dúvidas sobre o teste. Essa diferença pode estar associada à questão de os alunos assumirem que os instrutores, por já terem realizado o teste como candidato, o conhecem melhor que os demais e, portanto, estão em posição privilegiada para discuti-lo e prover esclarecimento.

Tabela 17 − Comparação entre professores e instrutores em relação à abordagem do exame em sala de aula

	Subgrupos	N	M	DP
Menção ao exame	1	12	4.000	0.852
	2	4	3.750	0.957
Discussão de aspectos do exame	1	12	3.750	0.965
	2	4	3.250	0.957
Trabalho envolvendo tarefas do exame	1	12	3.333	1.302
	2	4	3.250	0.957
Dúvidas dos alunos	1	12	2.583	0.668
	2	4	3.750	1.500

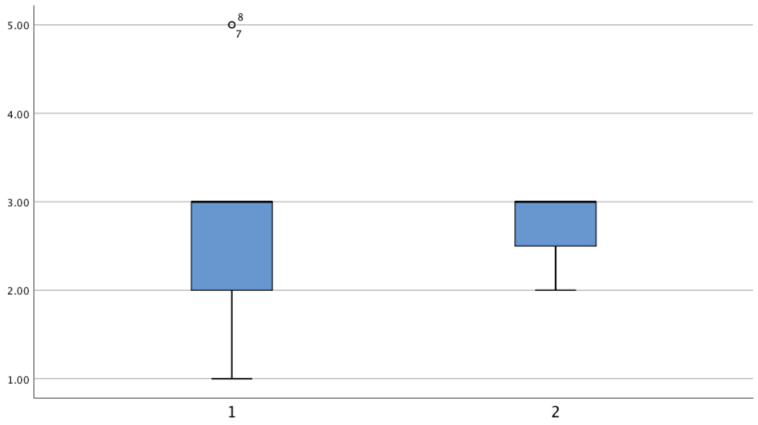
A relação entre a experiência no ensino de inglês aeronáutico e a abordagem ao exame em sala de aula está descrita na Tabela 18. Percebemos, novamente, que as médias de respostas diferem pouco entre os participantes menos e mais experientes. No item "Dúvidas dos alunos", notamos que o subgrupo menos experiente (subgrupo 1) alcança média um pouco mais elevada, porém existe maior variação em suas respostas. Isso pode ter ocorrido devido a três dos quatro controladores de tráfego aéreo pertencerem a esse subgrupo, dado que esses profissionais afirmaram ser procurados com mais frequência pelos alunos no que concerne ao teste.

Tabela 18 − Comparação entre experiência no ensino de inglês aeronáutico e abordagem do exame em sala de aula

	Subgrupos	N	M	DP
Menção ao exame	1	9	3.888	0.927
	2	7	4.000	0.816
Discussão de aspectos do exame	1	9	3.666	0.866
	2	7	3.571	1.133
Trabalho envolvendo tarefas do exame	1	9	3.333	1.000
	2	7	3.287	1.496
Dúvidas dos alunos	1	9	3.000	1.322
	2	7	2.714	0.487

O Gráfico 5, no formato *box plot*, indica a dispersão das médias dos subgrupos menos e mais experientes no ensino de inglês aeronáutico em relação às dúvidas dos alunos. Podemos observar a presença de um *outlier*, representado pelo ponto ou círculo fora da linha do subgrupo 1 (à esquerda), no nível 5 do eixo vertical. *Outliers* são considerados valores discrepantes que fogem à normalidade. No caso do subgrupo 1, enquanto a maior parte das respostas concentrou-se nos níveis 2 e 3, com uma resposta no nível 1, dois participantes assinalaram 5 e, por essa razão, esses valores são considerados atípicos. Os dois casos de nível 5 foram atribuídos, exatamente, por instrutores de inglês aeronáutico.

Gráfico 5 − Dispersão das respostas dos subgrupos 1 e 2 em relação às dúvidas dos alunos



O grau com que as habilidades e áreas linguísticas desenvolvidas em sala de aula estão relacionadas a tópicos do ambiente de trabalho do controlador de tráfego aéreo foi investigado por meio dos itens de 30 a 34. Pelo fato de as comunicações aeronáuticas utilizarem a radiotelefonia e serem limitadas a determinados assuntos, esperava que, como efeito retroativo do exame, os temas em sala de aula estivessem ligados à aviação e ao controle de tráfego aéreo e que houvesse mais ênfase na prática das habilidades de compreensão e produção orais. A Tabela 19 apresenta as respostas dos professores e instrutores a essas questões.

Tabela 19 − Influência do EPLIS no conteúdo

Influência do EPLIS no conteúdo	**N**	**M**	**DP**	**EPM**
Compreensão oral	16	4.750	0.447	0.111
Leitura	16	4.562	0.892	0.223
Produção oral	16	4.562	0.629	0.157
Vocabulário	16	4.687	0.478	0.119
Estrutura	16	4.062	0.997	0.249

Notamos, portanto, que as médias em todos os itens atingiram valores superiores a 4, tendo a compreensão oral apresentado a maior média entre elas. Esse resultado sugere que o efeito retroativo do exame em termos de conteúdo é bastante expressivo no ambiente escolar na perspectiva dos professores e instrutores. A Tabela 20 reporta esses resultados em porcentagem. Percebemos que mais de 80% dos professores e instrutores assinalaram os critérios “frequentemente” e “sempre” para expressar o grau com que as habilidades e áreas linguísticas desenvolvidas em sala de aula estão associadas a tópicos pertinentes ao ambiente profissional.

Embora haja uma maior concentração de respostas no nível “sempre”, constatamos respostas nos níveis inferiores, indicando que existem eventos em sala de aula que não mantêm

relação com o ambiente profissional. Ademais, ainda que esses itens não abordem diretamente a frequência com que essas áreas são desenvolvidas, os dados sugerem que a leitura pode ser mais privilegiada do que a produção oral em sala de aula, contrastando com os efeitos pretendidos, dado que é esperado que as habilidades orais sejam as mais enfatizadas.

Tabela 20 − Influência do EPLIS no conteúdo em porcentagem

	1	**2**	**3**	**4**	**5**
Compreensão oral	0.00	0.00	0.00	0.25	0.75
Leitura	0.00	0.06	0.06	0.12	0.75
Produção oral	0.00	0.00	0.06	0.31	0.62
Vocabulário	0.00	0.00	0.00	0.31	0.69
Estrutura	0.00	0.06	0.25	0.25	0.44

Diferenças mais expressivas entre as médias de respostas dos subgrupos (professor, instrutor e experiência) foram encontradas somente no item relacionado ao ensino de estrutura. O subgrupo formado por instrutores (controladores de tráfego aéreo) apresentou média 3.500, enquanto o subgrupo de professores marcou 4.250. Em relação à experiência no ensino de inglês aeronáutico, os mais experientes apresentaram média superior (4.571) a dos menos experientes (3.666). Essa diferença pode estar associada a visões distintas sobre o ensino de gramática e como a gramática é abordada por esses participantes em sala de aula.

Os dados apresentados aqui sugerem efeito retroativo do EPLIS em diversos aspectos do ensino e aprendizagem, tais como no conteúdo abordado em sala de aula em termos de tópicos e habilidades, na abordagem de ensino do professor, nas tarefas desenvolvidas e nas atividades avaliativas. Além disso, percebemos um aumento da intensidade do efeito ao compararmos esses dados com aqueles obtidos em 2014. Vimos também que a intensidade desse efeito parece sofrer variação, provavelmente, em função da experiência no ensino de inglês aeronáutico e dos níveis de conhecimentos sobre o exame que professores e instrutores alegam possuir, sustentando, assim, a hipótese de Alderson e Wall (1993) de que a intensidade do efeito retroativo varia de professor para professor. Outro dado importante que requer nossa atenção é a hipótese de que o desenvolvimento da leitura é bastante frequente em sala de aula, o que pode ser reflexo de outro fator, ainda desconhecido, atuando em conjunto com o exame.

4.3.4 Construto 3 – Preparação dos alunos para o EPLIS

O construto 3 buscou interpretar as facetas do efeito retroativo na preparação dos alunos para o EPLIS, embasado na afirmação de Smith (1991) de que, diante da alta-relevância de um exame, os participantes fazem o necessário para se prepararem ou prepararem aqueles

que irão se submeter ao teste. Essa dimensão é representada pela parte IV do questionário e contém 16 itens que questionam os participantes sobre a frequência com que trabalham atividades de produção e compreensão orais, leitura e escrita em sala de aula.

A Tabela 21 resume as respostas dos professores e instrutores à frequência com que trabalham atividades de produção e compreensão orais em sala de aula. CTA é a sigla comumente utilizada para designar controle de tráfego aéreo. Na perspectiva dos professores e instrutores, a atividade de responder oralmente perguntas colocadas sobre aviação e CTA é a prática mais frequente. Outras atividades que atingiram média superior a 4 foram: parafrasear palavras e expressões utilizadas no CTA, discutir um assunto relacionado à aviação e ao CTA, responder questões com base na escuta de conversas radiotelefônicas e descrever uma figura relacionada à aviação e ao CTA.

Tabela 21 − Atividades de compreensão e produção orais

	N	M	DP	EPM
Responder perguntas sobre aviação e CTA	16	4.500	0.730	0.182
Discutir assunto sobre aviação e CTA	16	4.312	0.704	0.176
Fazer uma apresentação oral sobre aviação e CTA	16	3.875	0.885	0.221
Descrever uma figura relacionada à aviação e CTA	16	4.187	0.910	0.227
Contar uma história com base nessa figura	16	3.937	0.997	0.249
Recontar uma história relacionada à aviação e ao CTA	16	3.750	1.000	0.250
Explicar uma situação problemática de CTA	16	3.875	0.885	0.221
Dar uma sugestão para a solução dessa situação	16	4.000	0.894	0.223
Simular uma conversa entre piloto e controlador	16	3.625	1.087	0.271
Explicar uma palavra relacionada à aviação e ao CTA	16	4.375	0.610	0.154
Responder questões sobre um áudio ou vídeo	16	4.187	0.910	0.227

Essas atividades apresentam certa semelhança às tarefas do EPLIS. Na fase I do exame, o candidato é exposto a 30 conversas piloto-controlador para atestar sua compreensão oral. Na fase II, o interlocutor lê situações de controle de tráfego aéreo para que o candidato possa explicá-las, coloca perguntas gerais sobre tópicos relacionados à área de atuação do profissional e apresenta uma figura para que o candidato a descreva. Isso sugere que existe sobreposição entre as atividades desenvolvidas em sala de aula e as tarefas do exame, evidenciando provável efeito retroativo. Essa suposição é reforçada ao constatar que a tarefa de simular com um colega uma conversa entre piloto e controlador – não contemplada no EPLIS, é a menos frequente em sala de aula em comparação com as demais, mesmo sendo considerada a mais representativa do uso da língua no ambiente profissional. Por outro lado, a baixa frequência desse tipo de tarefa pode indicar que as atividades em sala de aula estão sendo limitadas àquelas encontradas no teste.

Diferenças mais expressivas entre as respostas de professores e instrutores foram encontradas para os itens 38, 41 e 42, como podemos verificar na Tabela 22. O subgrupo dos

professores apresentou médias mais elevadas para esses itens, porém, dessa vez, vemos dispersão maior (desvio padrão) nos dados referentes aos instrutores. Esses itens fazem referência direta a tarefas presentes no exame: descrição de figuras (item 38), explicação de situações problemáticas de CTA (item 41) e oferta de sugestão para a solução dessas situações (item 42), o que pode indicar que os professores tendem a trabalhar esses tipos de atividades com maior frequência do que os instrutores.

Tabela 22 − Comparação entre professores e instrutores em relação às atividades orais em sala de aula

	Subgrupos	**N**	**M**	**DP**
Descrever uma figura relacionada à aviação e ao CTA	1	12	4.333	0.778
	2	4	3.750	1.258
Explicar uma situação problemática de CTA	1	12	4.000	0.738
	2	4	3.500	1.290
Dar uma sugestão para a solução dessa situação	1	12	4.166	0.717
	2	4	3.500	1.290

Entre os profissionais com menos ou mais experiência, as diferenças que ressaltaram estão nos itens 39 e 42 e podem ser observadas na Tabela 23. Nesse caso, os professores e instrutores menos experientes (subgrupo 1) apresentaram médias maiores em relação aos mais experientes (subgrupo 2), sugerindo que algumas atividades semelhantes ao exame são mais propensas a ocorrerem na prática pedagógica dos menos experientes, mesmo sendo eles menos conhecedores do exame.

Tabela 23 − Comparação entre experiência no ensino de inglês aeronáutico e atividades orais em sala de aula

	Subgrupos	**N**	**M**	**DP**
Descrever uma figura relacionada à aviação e ao CTA	1	9	4.222	0.666
	2	7	3.571	1.272
Dar uma sugestão para a solução dessa situação	1	9	4.222	0.833
	2	7	3.714	0.951

Em relação aos itens de 45 a 48, a Tabela 24 apresenta os resultados obtidos. O primeiro item aborda a prática de compreensão de textos, habilidade essa que não é exigida no exame, porém apontada como bastante recorrente em sala de aula com média acima de 4. Em porcentagem (Tabela 25), vemos que 76% dos professores e instrutores afirmaram que a leitura de textos sobre aviação e controle de tráfego aéreo é uma atividade que desenvolvem frequentemente ou sempre com os alunos. Esse resultado é interessante, pois indica a presença de outros elementos operando na determinação do efeito retroativo do EPLIS como, por

exemplo, as crenças e experiências dos professores, o que eles acreditam ser a maneira mais eficiente de ensinar, a influência do material didático, entre outros.

Tabela 24 − Atividades de leitura e escrita em sala de aula

	N	M	DP	EPM
Responder questões sobre um texto de aviação ou CTA	16	4.062	0.928	0.237
Escrever um diálogo entre piloto e controlador	16	3.062	1.388	0.347
Ler em voz alta um diálogo entre piloto e controlador	16	3.437	1.093	0.273
Escrever uma história relacionada ao CTA	16	2.500	1.264	0.316

Outro dado que nos chama a atenção refere-se à prática de leitura em voz alta em sala de aula. Essa atividade alcançou média aproximada de 3.5, com 56% dos participantes afirmando que a desenvolvem frequentemente ou sempre, porém não é uma tarefa exigida no exame. Nesse sentido, uma investigação deve ser feita a fim de averiguar se essa atividade é realmente recorrente, em que situações ela ocorre e seus propósitos.

Tabela 25 − Atividades de leitura e escrita em sala de aula em porcentagem

	1	2	3	4	5
Responder questões sobre um texto de aviação ou CTA	0.00	0.06	0.19	0.38	0.38
Escrever um diálogo entre piloto e controlador	0.19	0.12	0.31	0.19	0.19
Ler em voz alta um diálogo entre piloto e controlador	0.06	0.12	0.25	0.44	0.12
Escrever uma história relacionada ao CTA	0.25	0.31	0.19	0.19	0.06

A atividade de escrita de histórias relacionadas ao controle de tráfego aéreo apresentou a menor média. 56% dos participantes responderam que nunca ou raramente desenvolvem esse tipo de prática em sala de aula. Esse fato coincide com os efeitos pretendidos do exame, uma vez que a ênfase deve estar no desenvolvimento das habilidades orais dos alunos. As respostas em relação à escrita de diálogos apresentaram maior dispersão, o que pode indicar que alguns professores e instrutores costumam desenvolvê-la e outros não.

Ao comparar as médias de respostas entre professores e instrutores, foi possível constatar uma diferença maior em relação à leitura em voz alta. Enquanto a média dos professores atingiu 3.666, a média dos instrutores ficou em 2.750, indicando que esse tipo de atividade é menos recorrente na prática do instrutor. Diferença ainda maior foi vista em relação à experiência no ensino de inglês aeronáutico. Os mais experientes apresentaram média de 4.142 e os menos experientes 2.888 para esse item, indicando que os menos experientes tendem a recorrer menos a essa técnica.

Por fim, o item 50 do questionário pedia aos participantes que avaliassem a preparação para o exame oferecida na escola. A Tabela 26 distribui a porcentagem de respostas

nos seguintes níveis: 1 “extremamente insuficiente”, 2 “insuficiente”, 3 “razoável”, 4 “suficiente” e 5 “extremamente suficiente”.

Tabela 26 − Preparação para o exame em porcentagem

	1	2	3	4	5
Preparação para o exame	0.00	0.00	0.25	0.75	0.00

Verificamos que 75% consideraram suficiente a preparação oferecida pela escola em relação ao exame, ao passo que 25% afirmaram ser razoável. Nenhuma resposta foi marcada nos níveis mais baixos. A Tabela 27 nos mostra a diferença nas respostas entre professores e instrutores. A maioria dos professores considerou a preparação suficiente, ao passo que a maioria dos instrutores a julgou como razoável. Recordo que os instrutores assumiram possuir maior conhecimento sobre o EPLIS, além de terem realizado o exame como candidatos, o que os colocaria em uma melhor posição para avaliarem a adequação da preparação oferecida pela EEAR. Isso pode indicar que, na percepção dos instrutores, existem discrepâncias entre o que é ensinado e o que é cobrado no teste.

Tabela 27 − Comparação entre professores e instrutores em relação à preparação para o exame

	1	2	3	4	5
Professores	0.00	0.00	0.09	0.91	0.00
Instrutores	0.00	0.00	0.75	0.25	0.00

Em termos de experiência (Tabela 28), vemos que as respostas são bem próximas entre os dois subgrupos. A preparação para o exame é vista como suficiente por 78% dos menos experientes e 71% dos mais experientes, indicando que não há relação entre esses dois fatores.

Tabela 28 − Comparação entre experiência no ensino de inglês aeronáutico e preparação para o exame

	1	2	3	4	5
Menos experientes	0.00	0.00	0.22	0.78	0.00
Mais experientes	0.00	0.00	0.29	0.71	0.00

4.3.5 Construto 4 – Visões e percepções sobre o EPLIS

O construto 4 é representado pela parte V do questionário e contém 15 questões. Seus itens/variáveis abordaram questões referentes à influência do exame na motivação dos alunos, no nível de *stress* do professor e do aluno e na carga de trabalho e estudo desses participantes. Nessa parte, também foram tratados assuntos referentes à autenticidade das

tarefas do exame, a confiabilidade de seus resultados e sua relevância para o desempenho da função profissional. A importância de avaliarmos as percepções e atitudes dos participantes nessas questões reside no fato de que o modo como eles percebem e interpretam as demandas do exame, tradicionalmente denominado validade de face, pode influenciar suas atitudes em relação ao teste e, consequentemente, informar sobre os efeitos gerados. Green (2013) assegura que a perspectiva dos participantes sobre um exame é mais relevante na determinação de seu impacto do que a intenção de seus elaboradores e, por essa razão, essas questões devem ser consideradas em um estudo sobre efeito retroativo.

A Tabela 29 apresenta os resultados das respostas dos professores e instrutores às questões 51, 52, 53, 54, 55 e 57. Notamos que os itens mais bem avaliados correspondem à motivação causada pelo exame, a sua contribuição para a elevação da proficiência e à importância da aprovação. Os itens sobre autenticidade e confiabilidade do exame apresentaram médias menores, ao redor de 3.500. O último item apresentou média próxima a 2, indicando que professores e instrutores tendem a discordar da possibilidade de o exame se tornar opcional na escola.

Tabela 29 – Visões e percepções sobre o EPLIS

	N	M	DP	EPM
O exame motiva os alunos	16	4.062	0.853	0.213
O exame colabora para elevação do nível de proficiência	16	4.375	0.718	0.179
O exame representa bem as atividades do profissional CTA	16	3.562	0.963	0.240
O exame avalia precisamente o nível de proficiência dos alunos	16	3.437	1.093	0.273
O sucesso no exame é importante para o desempenho da função	16	4.000	0.730	0.182
A realização do exame na EEAR deve ser opcional	16	2.125	1.024	0.256

Em porcentagem (Tabela 30), verificamos que nenhum participante discordou da afirmação de que a aplicação do exame na EEAR contribui para a elevação do nível de proficiência dos alunos (níveis 1 e 2). No caso da motivação, apesar de haver discordância, mais de 80% dos professores e instrutores concordaram que o EPLIS é um fator motivador nos estudos de inglês na EEAR. Por outro lado, 81% se mostraram contrários à possibilidade de o exame se tornar opcional. Com base nesses resultados, assumo que o exame é bem aceito entre professores e instrutores. Eles o veem como um instrumento benéfico ao ambiente escolar, capaz de motivar os alunos e de impulsionar o nível de proficiência. Essa situação é, portanto, favorável à ocorrência de efeitos.

Quanto à autenticidade, definida por Green (2007a) como o grau de correspondência entre as tarefas presentes no exame e as tarefas reais no uso da língua alvo, 50% dos participantes demonstraram concordar, enquanto apenas 12% discordaram. Contudo,

podemos observar que um grupo expressivo de respondentes (38%) preferiu não emitir nenhum parecer sobre o assunto, escolhendo o critério 3 “nem concordo nem discordo”. A princípio, podemos associar essa falta de posicionamento à carência de entendimentos ainda existente por parte de alguns participantes envolvendo a natureza da linguagem da comunicação aeronáutica e o exame em si.

Tabela 30 – Visões e percepções sobre o EPLIS em porcentagem

	1	**2**	**3**	**4**	**5**
O exame motiva os alunos	0.00	0.06	0.12	0.50	0.31
O exame colabora para elevação do nível de proficiência	0.00	0.00	0.12	0.38	0.50
O exame representa bem as atividades do profissional CTA	0.00	0.12	0.38	0.31	0.19
O exame avalia precisamente o nível de proficiência dos alunos	0.00	0.25	0.25	0.31	0.19
O sucesso no exame é importante para o desempenho da função	0.00	0.00	0.25	0.50	0.25
A realização do exame na EEAR deve ser opcional	0.25	0.56	0.00	0.19	0.00

Sobre a confiabilidade dos resultados, embora, 50% dos professores e instrutores tenham concordado que o EPLIS avalia o nível de inglês dos alunos de forma precisa, 25% discordaram dessa afirmação, o que pode afetar a ocorrência de efeito retroativo nas práticas desses participantes. Por outro lado, sobre a relação entre o desempenho no exame e o desempenho na profissão, 75% a consideraram como positiva e o restante posicionou-se no centro. Verifiquei, então, diferenças mais expressivas nas respostas de professores comparadas àquelas dadas pelos instrutores. Os resultados dessa análise podem ser consultados na Tabela 31.

Tabela 31 − Comparação entre professores e instrutores em relação a visões e percepções sobre o EPLIS

	Subgrupos	**N**	**M**	**DP**
O exame motiva os alunos	1	12	4.166	0.717
	2	4	3.750	1.258
O exame colabora para elevação do nível de proficiência	1	12	4.500	0.674
	2	4	4.000	0.816
O exame representa bem as atividades do profissional CTA	1	12	3.750	0.965
	2	4	3.000	0.816
O exame avalia precisamente o nível de proficiência dos alunos	1	12	3.833	0.937
	2	4	2.250	0.500
O sucesso no exame é importante para o desempenho da função	1	12	4.166	0.717
	2	4	3.500	0.577
A realização do exame na EEAR deve ser opcional	1	12	1.833	0.834
	2	4	2.125	1.154

Percebemos que os instrutores apresentaram médias inferiores em todos os itens. No entanto, as diferenças nos itens autenticidade, confiabilidade e importância da aprovação se sobressaíram, sugerindo que os instrutores, por serem controladores de tráfego aéreo, são mais críticos ao julgar a verossimilhança do exame com a prática profissional e a relevância de sua

aprovação para o desempenho da função. O que nos chama a atenção, no entanto, é o fato de esses profissionais se mostrarem duvidosos quanto à capacidade do exame em avaliar com precisão o nível de inglês dos candidatos, o que também pode comprometer os efeitos pretendidos. Nesse item, as respostas dos instrutores alcançaram uma média considerada baixa (2.250). Tenho que ressaltar que essa percepção não pode, de forma alguma, ser considerada representativa dos controladores de tráfego brasileiros, uma vez que esses dados se referem à visão de, apenas, quatro participantes. Por fim, ambos os subgrupos mantêm médias baixas para a afirmação de que o exame deveria ser opcional para os alunos da EEAR, indicando que os benefícios superam os aspectos negativos.

Os itens autenticidade e confiabilidade também marcaram diferenças mais expressivas nos subgrupos com menos e mais experiência no ensino de inglês aeronáutico (Tabela 32). Diferentemente da comparação anterior, os dois itens alcançaram médias superiores a 3 nos dois subgrupos, com o subgrupo mais experiente marcando médias menores, porém com maior dispersão de respostas (DP=1.224). Dado que o grupo mais experiente atribuiu níveis mais altos de conhecimentos sobre o exame, esse resultado sugere que quanto maior o nível de conhecimentos sobre o exame, menor o nível de percepção em relação à autenticidade do teste e à confiabilidade de seus resultados.

Tabela 32 − Comparação entre experiência no ensino de inglês aeronáutico e visões e percepções sobre o EPLIS

	Subgrupos	**N**	**M**	**DP**
O exame representa bem as atividades do profissional CTA	1	9	3.777	0.971
	2	7	3.285	0.951
O exame avalia precisamente o nível de proficiência dos alunos	1	9	3.666	1.224
	2	7	3.142	0.899

O impacto negativo do exame pode ser depreendido nas respostas aos itens de 59 a 65. A Tabela 33 reporta os resultados dessa análise. Sabemos que efeitos negativos sempre irão existir por estarem associados à prática da avaliação como, por exemplo, o aumento da ansiedade, do medo e da pressão causados pela necessidade ou vontade de serem bem-sucedidos no exame ou de que seus alunos o sejam. Contudo, esses efeitos devem ser mantidos sob controle e minimizados para que a validade das interpretações e usos de seus resultados não seja comprometida. Ademais, ao sentirem muito pressionados pelo teste, professores e alunos podem se engajar em práticas de preparação consideradas indesejáveis por permitirem aumento dos resultados sem, necessariamente, promoverem o desenvolvimento da proficiência.

Tabela 33 − Impacto negativo do exame

	N	M	DP	EPM
O professor/instrutor sente-se confortável com o exame	16	3.687	0.873	0.218
O professor/instrutor sente-se pressionado pelo exame	16	2.625	1.147	0.286
A carga de trabalho do professor/instrutor aumenta com o exame	16	2.875	1.258	0.314
O nível de stress dos alunos aumenta com o exame	16	3.625	0.957	0.239
A carga de estudos dos alunos aumenta com o exame	16	3.000	1.032	0.258
O exame limita os tipos de atividades realizadas	16	1.875	0.885	0.221
O exame promove o uso de técnicas para sua realização	16	2.375	1.204	0.301

Os itens referentes aos níveis de ansiedade e *stress* dos alunos decorrentes do teste alcançaram média superior a 3. Os professores e instrutores reportaram níveis menores e tenderam a concordar com a afirmação de que se sentem confortáveis com o teste. Eles também demonstraram discordar sobre a possibilidade de estreitamento do currículo como efeito retroativo negativo do teste. Somado a isso, a possibilidade do exame de promover o uso de técnicas para a realização de provas em detrimento do desenvolvimento da proficiência também atingiu média relativamente baixa (2.375). Esses resultados sugerem que o impacto negativo do exame afeta um pouco mais os alunos, em termos de *stress*, ansiedade e carga de estudos, do que o professor/instrutor e sua prática pedagógica, porém não é considerado extremo.

Na Tabela 34, é possível observar a variação nas respostas dos participantes a esses itens por porcentagem. Em relação ao impacto negativo nos professores e instrutores, 56% afirmaram não sofrer qualquer pressão relacionada à presença do exame na escola e 68% se sentem confortáveis na preparação dos alunos para o teste. No estudo preliminar, a porcentagem de professores que alegou não se sentir pressionada pelo exame é bastante similar, 50%, sugerindo que, mesmo com a implementação do teste no contexto escolar, o nível de pressão sobre os professores não se alterou, o que nos faz questionar em relação à relevância atribuída ao teste por parte dos professores, instrutores e escola.

Tabela 34 − Impacto negativo do exame em porcentagem

	1	2	3	4	5
O professor/instrutor sente-se confortável com o exame	0.00	0.12	0.19	0.56	0.12
O professor/instrutor sente-se pressionado pelo exame	0.12	0.44	0.19	0.19	0.06
A carga de trabalho do professor/instrutor aumenta com o exame	0.12	0.31	0.25	0.19	0.12
O nível de *stress* dos alunos aumenta com o exame	0.00	0.12	0.31	0.38	0.19
A carga de estudos dos alunos aumenta com o exame	0.00	0.38	0.38	0.12	0.12
O exame limita os tipos de atividades realizadas	0.38	0.44	0.12	0.06	0.00
O exame promove o uso de técnicas para sua realização	0.31	0.25	0.19	0.25	0.00

Sobre o aumento na carga de trabalho, as respostas ao questionário de 2016 mostraram que 43% discordaram e 25% nem concordaram e nem discordaram. Em relação ao estreitamento do currículo, 82% foram categóricos em afirmar que o teste não propicia esse tipo de situação, e sobre o uso de estratégias de fácil solução para a realização do teste 56% se

mostraram contrários. No entanto, 25% se sentem pressionados, 12% desconfortáveis, 27% veem aumento na carga de trabalho, 25% concordaram sobre o uso de técnicas indesejáveis e 6% afirmaram que o exame reduz os tipos de atividades em sala de aula, indicando que uma parcela desses professores e instrutores apresentam visões negativas sobre o teste.

Diferenças mais salientes, na comparação entre os subgrupos com mais e menos experiência no ensino de inglês aeronáutico, foram encontradas nos itens referentes à pressão exercida pelo exame no professor/instrutor e ao aumento na carga de trabalho. O resultado pode ser visto na Tabela 35.

Tabela 35 − Comparação entre experiência no ensino de inglês aeronáutico e impacto negativo do exame

	Subgrupos	**N**	**M**	**DP**
O professor/instrutor sente-se pressionado pelo exame	1	9	2.333	1.000
	2	7	3.000	1.290
A carga de trabalho do professor/instrutor aumenta com o exame	1	9	3.222	1.394
	2	7	2.428	0.975

Os professores e instrutores mais experientes reportaram se sentirem mais pressionados pelo exame, porém não julgaram que sua carga de trabalho tenha sofrido considerável aumento. Já os menos experientes expressaram sentir uma pressão menor, porém alegaram maior aumento na carga de trabalho. Ou seja, apesar de não ser grande a pressão exercida pelo exame, os professores com menos experiência no ensino de inglês aeronáutico sentem-se mais sobrecarregados, provavelmente, em razão de carências nos conhecimentos sobre o teste.

Por fim, em relação ao resultado no EPLIS ser considerado na nota final da disciplina de inglês da 4ª série do curso, 69% dos respondentes declararam-se contrários, frente a 18% que concordaram, conforme podemos verificar na Tabela 36. Nesse caso, todos os instrutores de inglês aeronáutico posicionaram-se contra e somente 17% dos professores foram favoráveis. Podemos assumir que a maioria parece estar consciente dos impactos que essa decisão pode acarretar para a escola e para o sistema de controle de tráfego aéreo como, por exemplo, maior índice de retenção escolar e, consequentemente, diminuição no número de profissionais habilitados.

Tabela 36 − EPLIS como exame de saída na perspectiva dos professores e instrutores

	1	2	3	4	5
EPLIS como exame de saída	0.19	0.50	0.12	0.12	0.06

Retornando a Qi (2004), para a promoção de efeitos benéficos, é necessário manter a relevância de um teste em um nível ideal: sem ser demasiadamente relevante, o que acarretaria práticas preparatórias indesejáveis, e sem ser pouco relevante, o que não contribuiria para inovações. A chave está, portanto, em saber em que ponto desse contínuo se situa o EPLIS no contexto investigado. Com base nos resultados aqui apresentados, se, por um lado, as respostas dos professores e instrutores sugerem que houve um provável aumento da intensidade do efeito retroativo, principalmente, em relação aos conteúdos, tipos de atividades e avaliações, por outro lado, não nos parece que o exame já tenha alcançado um nível ótimo em termos de relevância.

4.4 O efeito retroativo nos alunos

A fim de investigar a influência do exame nas percepções e atitudes dos alunos após sua implementação no contexto escolar, analisei as respostas desses participantes ao questionário principal utilizando a análise fatorial, a estatística descritiva e os testes de hipóteses. A parte I do questionário coletou dados sobre o perfil dos alunos. As partes II, III e IV formaram os três construtos estudados: *Conhecimentos sobre o EPLIS*, *Preparação para o EPLIS* e *Visões e percepções sobre o EPLIS*. O primeiro construto, representado pela parte II do questionário, contém 10 itens/variáveis que avaliaram o grau de conhecimento do aluno em diversos aspectos do exame. O construto 2, relacionado à parte III do questionário, contém 32 itens/variáveis que exploraram questões ligadas à preparação do aluno dentro e fora da sala de aula. O último construto é composto por 13 itens/variáveis presentes na parte IV do questionário que analisaram as percepções sobre o exame e sua influência no contexto escolar.

Todas as variáveis foram submetidas, primeiramente, ao critério KMO e ao teste de Bartlett que confirmaram a possibilidade de serem fatoradas. Em seguida, três critérios foram utilizados para se obter o número ótimo de fatores a serem extraídos: o teste *Scree plot*, o critério Kaiser e a análise paralela. A rotação ortogonal varimax possibilitou minimizar o número de variáveis em cada fator e maximizar a carga fatorial, facilitando, assim, a interpretação da matriz de correlações. O coeficiente de Alfa Cronbach, por sua vez, garantiu a consistência dos dados coletados. Foram consideradas significantes cargas fatoriais superiores a 0.5 (em módulo) ou próximas a esse valor. Os resultados estão organizados em *Perfil dos participantes*, C*onhecimentos sobre o EPLIS*, *Preparação para o EPLIS* e *Percepções e visões sobre o exame*.

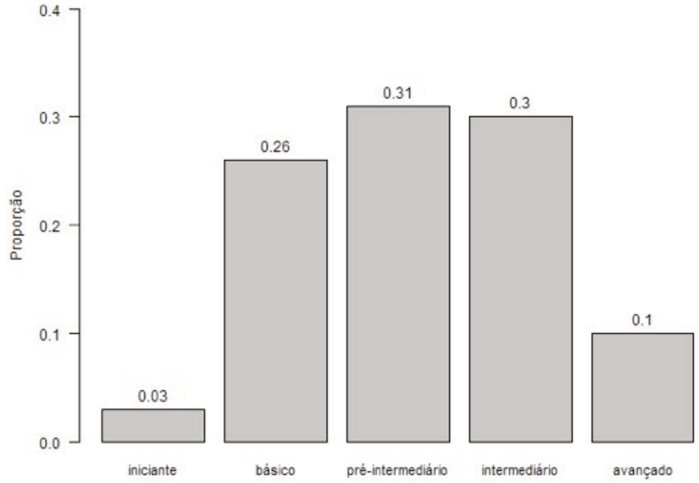
4.4.1 O perfil dos alunos

O questionário foi distribuído aos alunos das quatro séries do curso BCT. No total, foram recebidas respostas de 387 alunos, o que correspondeu a uma taxa de retorno de 95%. Dos respondentes, 56% eram do sexo feminino e 44% do sexo masculino. Por existir limite de idade para ingresso na EEAR, todos os alunos possuíam entre 17 e 24 anos de idade.

Em relação ao tempo de aprendizagem de inglês, 37% dos alunos declararam ter estudado inglês por pouco tempo (0 a 1 ano) antes de entrarem na EEAR e 39% responderam possuir entre 2 e 4 anos de aprendizagem na língua. 13% dos alunos disseram ter entre 5 e 6 anos de estudo, 10% entre 7 e 10 anos e somente 1% estudou inglês por mais de 10 anos.

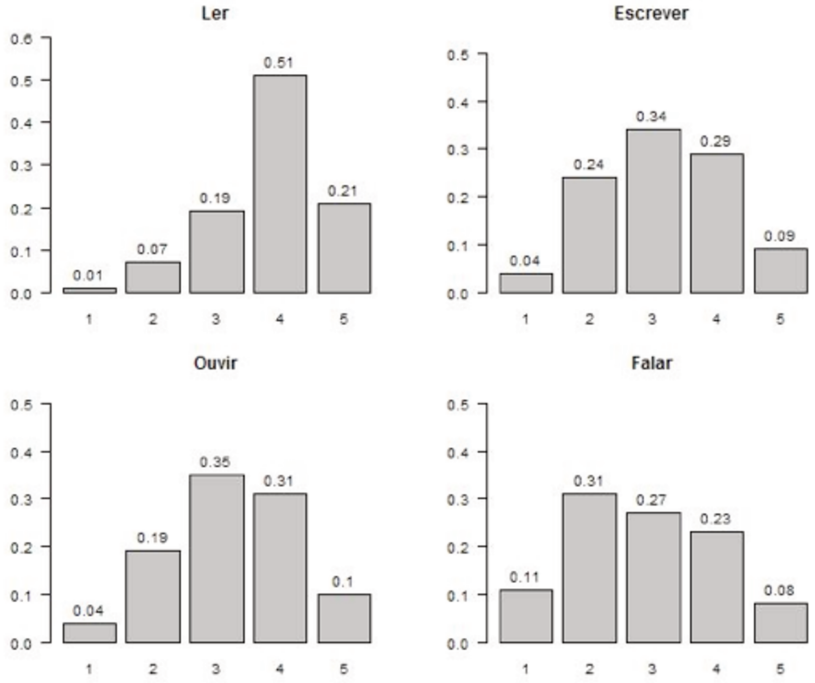
Quanto à autoavaliação do nível de proficiência, as respostas dos alunos apresentaram maior concentração entre os níveis pré-intermediário e intermediário, conforme Gráfico 6. Apesar de o processo seletivo para entrada na EEAR exigir o nível intermediário na língua inglesa, percebemos que 60% alegam possuir nível inferior ao exigido: 31% dos alunos se consideraram no nível pré-intermediário, 26% se autoavaliaram no nível básico e 3% se julgaram iniciantes.

Gráfico 6 − Nível de proficiência linguística autodeclarado



Foi também pedido aos alunos que avaliassem o nível de proficiência nas quatro habilidades linguísticas: ler, ouvir, falar e escrever. Os resultados estão dispostos no Gráfico 7, no qual 1 representa o nível “iniciante”, 2 “básico”, 3 “pré-intermediário”, 4 “intermediário” e 5 “avançado”.

Gráfico 7 − Proficiência autodeclarada por habilidade linguística



Em relação à compreensão de textos, a maior porcentagem de respostas encontra-se no nível intermediário, com 51%. Para a habilidade de compreensão oral, as respostas concentraram-se no nível pré-intermediário com 35%, seguidas pelo nível intermediário com 31%. As respostas dos alunos para a habilidade de escrita foram bastante semelhantes às da compreensão oral, com 34% dos alunos alegando possuir nível pré-intermediário nessa habilidade e 29% no nível intermediário. A habilidade de produção oral recebeu menor índice de respostas nos níveis mais elevados: 27% dos alunos se autoavaliaram no nível pré-intermediário e 23% no nível intermediário. A maior concentração de respostas se deu no nível básico, com 31%, o que indica que essa habilidade é considerada, pelos alunos, a menos desenvolvida.

4.4.2 Construto 1 − Conhecimentos sobre o EPLIS

Como discutido anteriormente, a familiaridade com a natureza e as características do teste é um elemento importante a ser considerado na determinação do efeito retroativo (HUGHES, 1989, 1993; GREEN, 2007a). Os conhecimentos dos alunos sobre o exame foram avaliados em aspectos referentes ao seu objetivo, ao conteúdo abordado, aos tipos de tarefas exigidas, ao desempenho mínimo desejável, aos níveis e critérios avaliados, ao manual do candidato, e à frequência com que o professor menciona o exame em sala de aula e discute seus

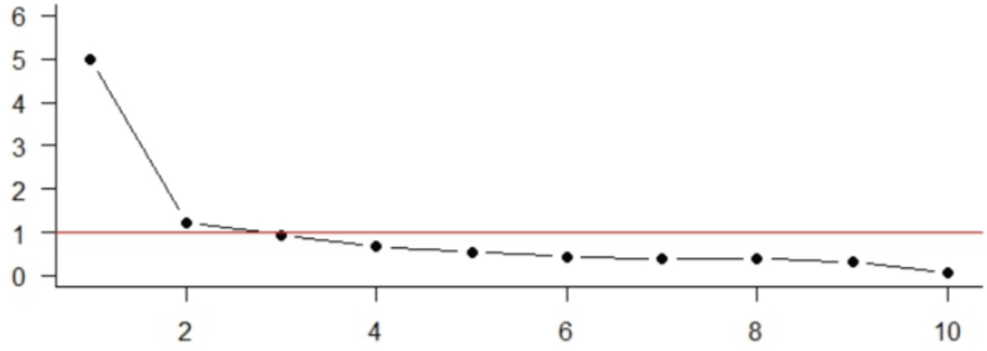
aspectos. As 10 variáveis foram submetidas ao critério KMO e ao teste de Bartlett para checar a adequação à análise de componentes principais. O critério KMO apresentou valor acima de 0.80, considerado bom. O teste de Bartlett apresentou significância (Sig) inferior a 0.05, possibilitando a fatoração dos dados. Esses resultados podem ser observados na Tabela 37.

Tabela 37 − KMO e Teste de Bartlett no construto 1

Medida de Adequação de Amostra Kaiser-Meyser-Olkin		0.861
Teste de Esfericidade de Bartlett	Aprox. Chi-quadrado	2198.206
	df	45
	Sig.	0.000

Para esse construto, foram extraídos 2 componentes principais que explicam a maior parte de variância presente nos dados. O Gráfico 8 demonstra o critério kaiser combinado ao teste *scree plot*. No eixo x, está representado o número de componentes e no eixo y encontramos seus respectivos autovalores. A linha horizontal indica onde foi feito o corte para a definição dos fatores, utilizando esses dois critérios.

Gráfico 8 − Extração dos componentes no construto 1



Na Tabela 38, apresento os resultados da análise paralela comparando os autovalores reais aos autovalores simulados. Nota-se que os autovalores reais para os dois primeiros componentes superam os valores gerados aleatoriamente, apontando que apenas dois componentes devem ser retidos. Os autovalores do 3º ao 10º componente estão abaixo de 1 e, conforme o critério kaiser, foram desprezados.

Tabela 38 − Análise paralela para retenção de componentes no construto 1

componentes	**Autovalor aleatório (análise paralela)**	**Autovalor real (dados originais)**
1	1.2655	4.986
2	1.1827	1.229
3	1.1188	0.947
...	**...**	**...**
10	0.7614	0.067

Os itens que se correlacionaram em cada um dos componentes extraídos podem ser verificados na Tabela 39, juntamente com suas respectivas cargas fatoriais. Nesta pesquisa, foram consideradas significantes cargas fatoriais iguais ou maiores que 0.5 (em módulo) ou próximas a esse valor. O coeficiente de Alfa Cronbach, que mede a consistência interna dos itens, foi calculado em 0.88 para esse grupo de variáveis, indicando alta confiabilidade dos dados. Os dois componentes combinados explicam 62% da variância total dos dados.

Tabela 39 − Itens e cargas fatoriais no construto 1 com rotação varimax

Itens	**componente 1**	**componente 2**	**Comunalidades**
Q10		0.51	0.41
Q11	0.64		0.46
Q12	0.70		0.59
Q13	0.84		0.74
Q14	0.85		0.75
Q15	0.69		0.50
Q16	0.75		0.61
Q17	0.76		0.61
Q18		0.86	0.75
Q19		0.82	0.74

	componente 1	**componente 2**
Autovalores	4.99	1.23
Variância Explicada (%)	49.85	12.29
Variância Acumulada (%)	49.85	62.14
Alfa Cronbach	0.89	0.68

O primeiro componente é composto por sete itens altamente correlacionados, com cargas fatoriais iguais ou superiores a 0.50. Possui autovalor calculado em 4.99, o que explica aproximadamente 50% da variância total dos dados. A confiabilidade nesse fator é considerada alta, com o coeficiente de Alfa Cronbach sendo calculado em 0.89. O componente 2 apresenta três itens com associação significativa entre eles. Seu autovalor é de 1.23 e é responsável por 12% da variância total dos dados. Sua consistência interna também é considerada adequada (α=0.68) e dois de seus itens apresentaram carga fatorial bastante elevada, acima de 0.80.

4.4.2.1 Interpretação dos componentes no construto 1

Os componentes extraídos no construto 1 foram interpretados como *Conhecimentos específicos sobre o exame* (componente 1) e *O exame em sala de aula* (componente 2). O Quadro 14 retrata a estrutura fatorial nesse construto, com seus componentes, a descrição de seus itens e suas cargas fatoriais.

Quadro 14 − Estrutura fatorial do construto 1

Construto 1 – Conhecimentos sobre o EPLIS (α=0.88) – 2 componentes	
1. Conhecimentos específicos sobre o exame α=0.89	**Ainda sobre o EPLIS, como você avalia o seu nível de conhecimento sobre os seguintes itens?**
	Q11 o objetivo do EPLIS 0.64
	Q12 o conteúdo do EPLIS 0.70
	Q13 tarefas da fase I 0.84
	Q14 tarefas da fase II 0.85
	Q15 a tabela de níveis de proficiência da OACI 0.69
	Q16 o desempenho necessário para obter o nível mínimo (nível 4) 0.75
	Q17 manual do candidato 0.76
2. O exame em sala de aula α=0.68	Q10 De forma geral, como você avalia o seu nível de conhecimento sobre o EPLIS? 0.51
	Q18 Com que frequência os professores da série atual mencionam o EPLIS em sala de aula? 0.86
	Q19 Com que frequência os professores da série atual discutem aspectos específicos do exame em sala de aula? 0.82

O componente 1, *Conhecimentos específicos sobre o exame*, como o próprio nome sugere, correlaciona itens específicos sobre o EPLIS, tais como, tarefas, objetivo, conteúdo, escala de níveis de proficiência e desempenho mínimo desejável, além do manual do candidato. Destaco que o manual do candidato, além de trazer informações sobre a estrutura do EPLIS, é a principal fonte de pesquisa para exemplares de tarefas que são contempladas no exame. As cargas fatoriais dos itens demonstraram variação entre 0,85 a 0,64. Essa dimensão é a mais representativa por carregar a maior parte de explicação dos dados, cerca de 50%.

A segunda dimensão é composta por três itens relacionados ao conhecimento geral do exame e sua abordagem em sala de aula. Dois itens, Q18 e Q19, apresentaram cargas fatoriais bastante elevadas e, por isso, tiveram mais representatividade na categorização do componente. Esse componente, por abordar questões relacionadas à frequência com que o exame é mencionado e discutido em sala de aula, foi interpretado como *O exame em sala de aula*. A Tabela 40 apresenta a porcentagem de respostas aos itens pertencentes ao componente 1, na qual 1 significa “nenhum”, 2 “pouco”, 3 “razoável”, 4 “bom” e 5 “ótimo”.

Tabela 40 − Nível de conhecimentos específicos sobre o EPLIS

	1	**2**	**3**	**4**	**5**
Q11	0.03	0.18	0.36	0.29	0.13
Q12	0.03	0.30	0.37	0.25	0.04
Q13	0.26	0.33	0.26	0.12	0.04
Q14	0.27	0.34	0.25	0.11	0.03
Q15	0.20	0.26	0.26	0.18	0.10
Q16	0.13	0.30	0.30	0.19	0.08
Q17	0.49	0.25	0.16	0.10	0.02

Ao analisar as respostas dos alunos aos itens do componente 1 separadamente, *Conhecimentos específicos sobre o exame*, descobrimos que, em relação às questões com maior

carga fatorial (Q13, Q14 e Q17), as respostas dos alunos concentraram-se entre os níveis de conhecimento "nenhum" e "pouco", indicando que eles possuem pouco ou nenhum conhecimento sobre as tarefas do teste e o manual do candidato. Nos demais itens, a maior parte dos alunos marcou "pouco" e "razoável, com exceção da Q11 (objetivo do exame) que recebeu níveis mais altos de respostas, "razoável" e "bom".

Em relação aos itens do componente 2, *O exame em sala de aula*, as respostas dos alunos concentraram-se nos níveis "pouco" e "razoável" para caracterizar o conhecimento geral que possuem sobre o EPLIS. Por outro lado, os alunos alegaram que o exame é bastante mencionado em sala de aula, com 54% das respostas no nível 4 (frequentemente). Além disso, 43% dos alunos disseram que "às vezes" aspectos do exame são discutidos em sala de aula, enquanto 30% responderam "frequentemente". Esses dados parecem contraditórios uma vez que, apesar de os alunos afirmarem que o exame é mencionado frequentemente em sala de aula e seus aspectos são discutidos razoavelmente, eles avaliaram como baixo os níveis de conhecimentos que possuem sobre o teste. As porcentagens podem ser consultadas na Tabela 41, cuja legenda é a mesma da tabela anterior para a Q10. Para os itens Q18 e Q19, lê-se: 1 "nunca", 2 "raramente", 3 "às vezes", 4 "frequentemente" e 5 "sempre".

Tabela 41 − Nível de conhecimento dos alunos sobre o formato do exame

	1	**2**	**3**	**4**	**5**
Q10	0.02	0.41	0.41	0.15	
Q18		0.03	0.22	0.53	0.22
Q19	0.02	0.12	0.43	0.30	0.12

Esses resultados revelam que, na percepção dos alunos, ainda existem falhas relacionadas aos conhecimentos que eles possuem sobre o EPLIS, especialmente em termos específicos (tarefas e manual do candidato). Poucos alunos afirmaram possuir bom nível de conhecimentos, e uma pequena minoria alegou possuir nível ótimo. Lembro que não só os professores, mas também os alunos, precisam se familiarizar com especificidades do exame que irão realizar. Como discuti anteriormente, a carência de entendimentos sobre o teste pode ocasionar prejuízo para a promoção de efeitos desejados.

4.4.3 Construto 2 – Preparação para o EPLIS

De acordo com Hughes (1993) e Bailey (1996), um exame poderá influenciar as atitudes dos participantes que, por sua vez, irão engajar-se em novas práticas (processos) para a geração de novos produtos. No construto 2, busquei verificar as facetas do efeito retroativo

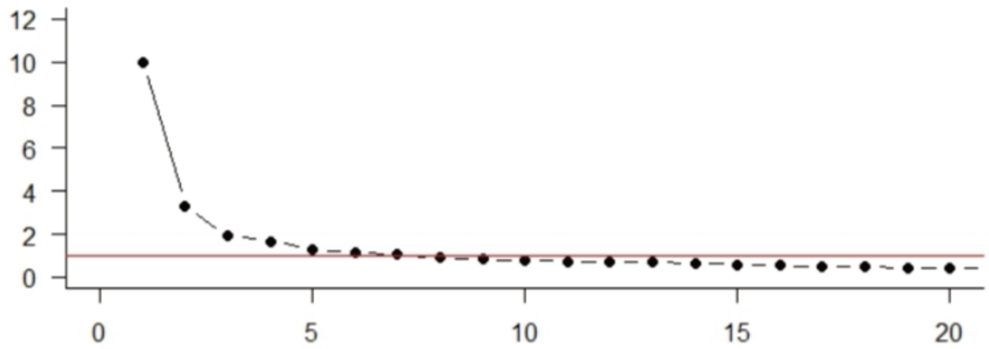
do EPLIS nos processos dentro e fora da sala de aula sob a perspectiva dos alunos. Esse construto é formado por 32 itens/variáveis que exploram a influência do exame nos tópicos, habilidades e atividades trabalhadas em sala de aula pelos professores e na preparação do aluno extraclasse. Essas variáveis foram submetidas ao critério KMO e ao teste de Bartlett para checar a adequação à análise de componentes principais. Os resultados foram considerados satisfatórios e podem ser verificados na Tabela 42.

Tabela 42 − KMO e Teste de Bartlett no construto 2

Medida de Adequação de Amostra Kaiser-Meyser-Olkin		0.907
Teste de Esfericidade de Bartlett	Aprox. Chi-quadrado	6372.195
	df	496
	Sig.	0.000

Em relação à extração dos componentes principais, o critério Kaiser e o teste *Scree plot* apresentaram, primeiramente, a possibilidade de extração de sete componentes, conforme pode ser observado no Gráfico 9. A linha horizontal marca o ponto onde foi feito o corte para a obtenção dos valores.

Gráfico 9 − Extração dos componentes no construto 2



No entanto, a análise paralela de Monte-Carlo indicou a retenção de apenas de quatro componentes, uma vez que o autovalor do quinto componente (1.277) foi menor do que aquele obtido na matriz aleatória (1.3537), como mostra a Tabela 43. Para chegar a uma conclusão quanto ao número de componentes a serem retidos, verifiquei o número de itens com carga fatorial acima ou próximas de 0.50 em cada componente. O componente 5 apresentou quatro itens correlacionados, enquanto que, no componente 6, foram correlacionados apenas dois itens, e no componente 7 apenas um item. Em relação à consistência interna, o quinto componente apresentou α=0.70. Porém, os componentes 6 e 7 tiveram confiabilidade abaixo de 0.50. Considerando esses aspectos, optei por excluir os componentes 6 e 7, pelo número

reduzido de itens correlacionados e baixa consistência interna, e manter o componente 5, mesmo tendo a análise paralela indicado que esse componente possui autovalor abaixo do autovalor aleatório.

Tabela 43 − Análise paralela para retenção de componentes no construto 2

componentes	Autovalor aleatório (análise paralela)	Autovalor real (dados originais)
1	1.5847	10.005
2	1.5073	3.327
3	1.4495	1.930
4	1.3986	1.675
5	1.3537	1.277
6	1.3113	1.139
7	1.2715	1.048
8	1.2339	0.900
...	...	...
32	0.5384	0.169

Portanto, foram retidos cinco componentes principais nesse construto que explicam aproximadamente 57% da variância total dos dados. A consistência interna dos dados é alta, com Alfa Cronbach calculado em 0.92. Os itens que se correlacionaram dentro de cada um dos cinco componentes extraídos e suas respectivas cargas fatoriais estão descritos na Tabela 44.

O primeiro componente extraído é formado por sete itens correlacionados que explicam 31% da variância total dos dados. As cargas fatoriais dos itens variam de 0.47 a 0.81 e são consideradas significantes. A confiabilidade dos itens nesse componente também é expressiva (α=0.88). O componente 2 detém 10% da variância total dos dados, é composto por oito itens altamente correlacionados e apresenta consistência interna elevada (α=0.89). No componente 3, cinco itens são capazes de explicar 6% da variância dos dados originais. Todos apresentam carga fatorial significante e bom nível de confiabilidade (α=0.85). O quarto componente é responsável por 5% da variância e três itens apresentam alta correlação com cargas fatoriais significantes. O coeficiente de Alfa Cronbach para esse componente foi calculado em 0.73. No quinto componente, quatro itens apresentaram alta correlação, com cargas fatoriais elevadas, e são responsáveis pela variância explicada de 4%. A confiabilidade desses itens foi calculada em 0.70.

Houve um caso de carga cruzada no item Q32, com 0.52 no componente 1 e 0.56 no componente 2. Por fazer referência a uma tarefa da fase II do exame, esse item se correlacionou melhor com as variáveis do segundo grupo e foi, portanto, retirado do primeiro.

Tabela 44 − Itens e cargas fatoriais no construto 2 com rotação varimax

Itens	Comp1	Comp2	Comp3	Comp4	Comp5	Comunalidades
Q21			0.65			0.61
Q22			0.78			0.65
Q23			0.74			0.66
Q24			0.82			0.73
Q25			0.68			0.62
Q26		0.55				0.61
Q27		0.58				0.65
Q28		0.60				0.60
Q29		0.74				0.66
Q30		0.77				0.70
Q31		0.67				0.65
Q32	~~0.52~~	0.56				0.62
Q33		0.61				0.65
Q34	0.77					0.71
Q35	0.75					0.75
Q36	0.47					0.52
Q37	0.70					0.71
Q38	0.81					0.75
Q39	0.65					0.68
Q41					0.56	0.53
Q42					0.71	0.57
Q43					0.71	0.64
Q44					0.66	0.61
Q45				0.61		0.59
Q49				0.87		0.82
Q50				0.86		0.81

	Comp1	Comp2	Comp3	Comp4	Comp5
Autovalores	10.01	3.33	1.93	1.67	1.27
Variância Explicada (%)	31.26	10.39	6.03	5.23	3.99
Variância Acumulada (%)	31.26	41.66	47.69	52.93	56.92
Alfa Cronbach	0.88	0.89	0.85	0.73	0.70

4.4.3.1 Interpretação dos componentes no construto 2

Os componentes do construto 2 foram interpretados como *Atividades de compreensão oral e escrita e produção oral e escrita* (componente 1), *Atividades de produção oral relacionadas ao exame* (componente 2), *Impacto do exame no conteúdo – tópicos e habilidades* (componente 3), *Preparação individual focada no exame fora da sala de aula* (componente 4) e *Preparação em grupo de natureza variada fora da sala de aula* (componente 5). É possível observar a estrutura fatorial do construto 2 no Quadro 15.

O primeiro componente é composto por itens que abordam atividades ligadas à leitura de textos sobre controle de tráfego aéreo e aviação, compreensão oral de interações entre controlador e piloto, produção oral de diálogos entre esses dois profissionais e atividades de escrita. Denominado *Atividades de compreensão oral e escrita e produção oral e escrita*, esse componente carrega 31% da variância total dos dados. As respostas dos alunos a seus itens podem ser consultadas na Tabela 45.

Quadro 15 − Estrutura fatorial do construto 2

Construto 2 – Preparação para o EPLIS (α=0.92) – 5 componentes	
1. Atividades de compreensão oral e escrita e produção oral e escrita (α=0.88)	**Considerando a série atual, com que frequência você realiza as seguintes atividades em sala de aula?**
	Q34 Simular, com um colega, um diálogo entre piloto e controlador. 0.77
	Q35 Responder questões sobre um áudio ou vídeo envolvendo uma conversa entre piloto e controlador. 0.75
	Q36 Responder questões sobre um texto cujo tema está relacionado à aviação e ao controle de tráfego aéreo. 0.47
	Q37 Escrever um diálogo entre piloto e controlador. 0.70
	Q38 Ler em voz alta um diálogo entre piloto e controlador.0.81
	Q39 Escrever uma história envolvendo uma situação complicadora de tráfego aéreo. 0.65
2. Atividades de produção oral relacionadas ao exame (α=0.89)	Q26 Responder, oralmente, perguntas colocadas pelo(a) professor (a) sobre aviação e controle de tráfego aéreo. 0.55
	Q27 Discutir com os colegas um assunto específico sobre aviação e controle de tráfego aéreo. 0.58
	Q28 Fazer uma apresentação oral sobre um assunto relacionado à aviação e ao controle de tráfego aéreo. 0.60
	Q29 Descrever uma figura relacionada à aviação e ao controle de tráfego aéreo. 0.74
	Q30 Contar uma história com base em uma figura relacionada à aviação e ao controle de tráfego aéreo. 0.77
	Q31 Recontar uma história com base em um áudio ou texto relacionado à aviação e ao controle de tráfego aéreo. 0.67
	Q32 Explicar, oralmente, uma situação-problema de tráfego aéreo. 0.56
	Q33 Dar uma sugestão, oralmente, para um problema de tráfego aéreo. 0.61
3. Impacto do exame no conteúdo – tópicos e habilidades (α=0.85)	**Considerando a série atual, com que frequência as seguintes atividades realizadas em sala de aula estão relacionadas à AVIAÇÃO E AO CONTROLE DE TRÁFEGO AÉREO?**
	Q21 compreensão de áudios e vídeos 0.65
	Q22 compreensão de textos escritos 0.78
	Q23 prática de produção oral 0.74
	Q24 prática de vocabulário 0.82
	Q25 prática de itens gramaticais 0.68
4. Preparação individual focada no exame fora da sala de aula (α=0.73)	**Com que frequência você se envolve nas seguintes atividades EXTRA-CLASSE?**
	Q45 Faço os exercícios de áudio disponíveis no *site* do exame.0.61
	Q49 Assisto aos vídeos da fase II do EPLIS disponíveis no *site* do exame. 0.87
	Q50 Faço uso do *sample* da fase II do EPLIS disponível no *site* do exame. 0.86
5. Preparação em grupo de natureza variada fora da sala de aula (α=0.70)	Q41 Participo de atividades extras oferecidas pelo Galpão de inglês. 0.56
	Q42 Converso com os colegas em inglês para praticar para o exame. 0.71
	Q43 Tiro dúvidas sobre o exame e seu conteúdo com os professores de inglês. 0.71
	Q44 Participo de grupos de estudos voltados para o exame. 0.66

Os itens Q35, Q36 e Q38 foram os que mais concentram respostas nos níveis "frequentemente" e "sempre", demonstrando que as atividades de compreensão oral, de compreensão de textos e leitura em voz alta são bastante recorrentes em sala de aula. A compreensão de textos, apesar de não estar contemplada no EPLIS, parece fazer parte do

repertório em sala de aula, o que diminui a chance de estreitamento do currículo como efeito negativo do exame, mas também indica a presença de um outro fator interagente.

Tabela 45 − Atividades de leitura, compreensão oral, produção oral e escrita

	1	2	3	4	5
Q34	0.19	0.15	0.22	0.25	0.19
Q35	0.12	0.10	0.17	0.28	0.33
Q36	0.04	0.05	0.17	0.33	0.41
Q37	0.29	0.23	0.21	0.17	0.10
Q38	0.18	0.14	0.18	0.28	0.21
Q39	0.30	0.24	0.24	0.16	0.06

As questões 37 e 39 referem-se à prática escrita e foram as que menos receberam respostas nos níveis "frequentemente" e "sempre", o que faz sentido em termos de efeito retroativo, uma vez que o exame avalia somente as habilidades orais. Porém, é necessário investigar o quê participantes têm a dizer sobre essa situação, se eles a consideram positiva ou negativa para o desenvolvimento da proficiência.

Em relação à simulação de diálogos (Q34), as respostas dos alunos variaram bastante em todos os níveis, com maior concentração nos níveis "às vezes" e "frequentemente". Assim como a compreensão de textos, esse tipo de tarefa também não é cobrado no exame, o que contraria a possibilidade de estreitamento do currículo. Somado a isso, esse fato pode confirmar que os valores e crenças do professor ou outros fatores interagem no efeito retroativo de um exame. No estudo qualitativo, explorarei as razões que fazem com que essas atividades sejam enfatizadas em sala de aula.

O segundo componente a emergir, *Atividades de produção oral relacionadas ao exame***,** apresenta itens relacionados à fase II do EPLIS e explica 10% da variância presente nos dados. Os itens Q29, Q30 e Q31 possuem cargas fatoriais mais altas e fazem referência a duas tarefas presentes no teste: relato de histórias e descrição de figuras. O item Q33 recebeu carga fatorial acima de 0.60 e representa outra tarefa presente no exame: oferta de sugestão para solucionar uma situação problemática de tráfego aéreo. As respostas dos alunos a esses itens concentraram-se nos níveis "às vezes" e "frequentemente". O item 26 alcançou maior proporção de respostas no nível "sempre" (25%), como pode ser observado na Tabela 46. O tipo de tarefa representado nesse item, responder oralmente perguntas sobre aviação e controle de tráfego aéreo, também faz parte da fase II do exame, na qual o interlocutor coloca ao candidato quatro questões a serem respondidas. Essa dimensão sugere que os professores estão engajados em familiarizar e preparar seus alunos para o teste. Por outro lado, sabemos que a prática excessiva com itens semelhantes aos do teste pode reduzir os métodos de ensino em sala

de aula e apresentar um efeito retroativo negativo (SMITH, 1991; GREEN, 2013), o que deve ser investigado mais a fundo.

Tabela 46 − Atividades de produção oral relacionadas ao exame

	1	2	3	4	5
Q26	0.01	0.08	0.24	0.42	0.25
Q27	0.02	0.12	0.33	0.29	0.24
Q28	0.07	0.28	0.29	0.26	0.10
Q29	0.04	0.16	0.32	0.27	0.21
Q30	0.10	0.25	0.31	0.25	0.10
Q31	0.10	0.22	0.35	0.23	0.09
Q32	0.07	0.17	0.26	0.36	0.14
Q33	0.11	0.19	0.34	0.24	0.13

O terceiro componente, *Impacto do exame no conteúdo – tópicos e habilidades*, englobou questões altamente correlacionadas aos tópicos e habilidades das atividades desenvolvidas em sala de aula. As respostas dos alunos a todos os itens desse componente concentraram-se nos níveis "sempre" e "frequentemente", sugerindo que as atividades realizadas em sala de aula abordam tópicos relativos ao controle de tráfego aéreo e à aviação. A habilidade de compreensão oral foi a que mais recebeu respostas no nível "sempre" (56%), seguida da prática de vocabulário (48%) e da compreensão de textos (47%). Ao considerarmos que a primeira fase do EPLIS é totalmente voltada à avaliação da compreensão oral do candidato em situações específicas de tráfego aéreo, é compreensível que o professor priorize o desenvolvimento da compreensão oral do aluno e do vocabulário em sala de aula. Porém, o EPLIS não avalia a compreensão de textos, e o fato de essa habilidade ser vista como bastante frequente também pelos alunos necessita averiguação.

Os componentes 4 e 5 representam as atitudes dos alunos em relação à preparação para o exame extraclasse. O componente 4, caracterizado como *Preparação individual focada no exame fora da sala de aula*, engloba itens relacionados à preparação com atividades no mesmo formato do exame, utilizando os recursos disponíveis no *site* do EPLIS. O componente 5, *Preparação em grupo de natureza variada fora da sala de aula*, correlacionou atividades preparatórias variadas, tais como, conversar em inglês com os colegas, procurar os professores para tirar dúvidas e participar de grupos de estudos. As respostas dos alunos a esses itens indicam que eles são um pouco mais propensos a se engajarem em atividades variadas do que em atividades individuais com foco no teste. Enquanto as respostas dos alunos em relação à frequência com que conversam em inglês com colegas e tiram dúvidas com professores concentraram-se no nível "às vezes", mais de 80% dos alunos responderam que nunca utilizaram os vídeos e os exemplos disponíveis no *site* do EPLIS. Esse dado nos leva a pensar

que pode ainda haver deficiência na divulgação do *site* aos alunos ou que a aprovação no exame não é uma prioridade para esses alunos.

4.4.4 Construto 3 – Visões e percepções sobre o EPLIS

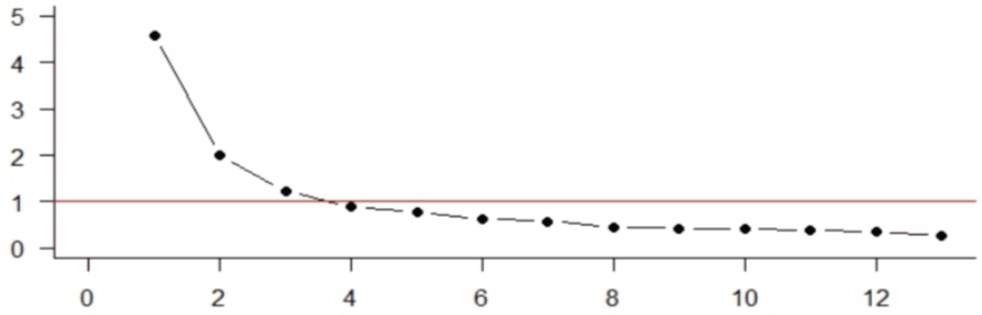
Compreender as percepções e visões dos alunos sobre o exame e suas exigências é crucial em um estudo sobre efeito retroativo, uma vez que elas podem modificar o grau de relevância imposto aos usuários pelos elaboradores do exame. No construto 3, 13 itens/variáveis avaliaram a influência do exame na motivação dos alunos, no nível de *stress* e na carga de estudos, bem como o nível de importância do exame e a preparação oferecida pela EEAR. Esses itens foram submetidos ao critério KMO e ao teste de Bartlett para certificar se a fatoração poderia ser empregada. Os resultados foram considerados positivos e podem ser verificados na Tabela 47.

Tabela 47 − KMO e Teste de Bartlett no construto 3

Medida de Adequação de Amostra Kaiser-Meyser-Olkin		0.852
Teste de Esfericidade de Bartlett	Aprox. Chi-quadrado	1806.571
	df	78
	Sig.	0.000

Foram extraídos três componentes que explicam 60% da variação total dos dados. O Gráfico 10 demonstra como foi feita a extração dos dados nesse construto com o emprego do critério Kaiser e do teste *Scree plot*. Os pontos acima da linha horizontal representam os três componentes extraídos.

Gráfico 10 − Extração dos componentes no construto 3



O resultado da análise paralela de Monte-Carlo confirmou a retenção desses três componentes, conforme indicado na Tabela 48. Vejam que os autovalores reais dos três

primeiros componentes superam os autovalores aleatórios. A partir do quarto componente, o autovalor real cai para menos de 1 e, conforme critério Kaiser, deve ser desconsiderado.

Tabela 48 − Análise paralela para retenção de componentes no construto 3

componentes	Autovalor aleatório (análise paralela)	Autovalor real (dados originais)
1	1.3156	4.596
2	1.2375	1.981
3	1.1744	1.224
4	1.1240	0.892
...	...	...
13	0.7194	.264

Na Tabela 49, estão dispostos os itens que se correlacionaram dentro de cada componente extraído e que alcançaram carga fatorial satisfatória, acima ou igual a 0.50. A confiabilidade dos itens nesse construto apresenta valor de Alfa Cronbach em 0.74.

Tabela 49 − Itens e cargas fatoriais no construto 3 com rotação varimax

Itens	Comp1	Comp2	Comp3	Comunalidades
Q52	0.62			0.63
Q53			0.59	0.53
Q54			0.78	0.68
Q55			0.80	0.65
Q56			0.56	0.56
Q57	0.64			0.62
Q58	0.56			0.49
Q59	0.68			0.57
Q60		0.83		0.76
Q61		0.84		0.75
Q62		0.81		0.70
Q63	-0.68			0.50

	Comp1	Comp2	Comp3
Autovalores	4.59	1.98	1.22
Variância Explicada (%)	35.35	15.24	09.41
Variância Acumulada (%)	35.35	50.59	60.01
Alfa Cronbach	0.79	0.84	0.65

O primeiro componente explica 35% da variância presente nos dados e é composto por cinco itens. Seu índice de confiabilidade interna alcançou 0.79, considerado bom. O componente 2 compreende a correlação existente entre três itens do questionário e detém 15% da variância total. A consistência interna nesse componente foi considerada elevada, sendo calculada em 0.84. O último componente é formado pela correlação entre quatro itens e responde a 9.5% da variância dos dados, com nível razoável de confiabilidade (α=0.65).

4.4.4.1 Interpretação dos componentes no construto 3

Os componentes no construto 3 foram interpretados como: *Significância do exame* (componente 1), *Preparação para o exame* (componente 2) e *Impacto negativo do exame* (componente 3). No Quadro 16, podemos observar o detalhamento dos itens relacionados a cada um dos componentes interpretados.

Quadro 16 − Estrutura fatorial do construto 3

Construto 3 – Visões e percepções sobre o exame (α=0.74) – 3 componentes	
1. Significância do exame (α=0.79)	Q52 O EPLIS me motiva a melhorar meu nível de inglês. 0.62
	Q57 O EPLIS representa bem as tarefas de um controlador de tráfego aéreo. 0.64
	Q58 O EPLIS avalia o nível de inglês dos alunos precisamente. 0.56
	Q59 O sucesso no EPLIS é importante para o bom desempenho de um controlador de tráfego aéreo. 0.68
	Q63 A realização do EPLIS durante o curso de formação deve ser opcional. -0.68
2. Preparação para o exame (α=0.84)	Q60 As aulas de inglês na EEAR preparam bem os alunos para o EPLIS. 0.83
	Q61 Os projetos desenvolvidos nas disciplinas de língua inglesa na EEAR contribuem para o sucesso no EPLIS. 0.84
	Q62 As avaliações de inglês na EEAR estão em consonância com o EPLIS. 0.81
3. Impacto negativo do exame (α=0.65)	Q53 O EPLIS modificou minha maneira de estudar inglês.0.59
	Q54 O EPLIS me força a estudar inglês fora da sala de aula.0.78
	Q55 O EPLIS aumenta consideravelmente minha carga de estudos. 0.80
	Q56 EPLIS aumenta o meu nível de ansiedade/*stress*. 0.56

O componente 1 engloba questões ligadas à motivação dos alunos, suas percepções sobre a autenticidade, confiabilidade e credibilidade do exame, bem como sobre a possibilidade de o exame se tornar opcional. O item com maior carga fatorial corresponde à importância da aprovação no teste, o que fez com que esse componente fosse interpretado como *Significância do exame*. Por carregar a maior parte da variância total dos dados, esse componente representa a principal faceta nas visões e percepções do aluno sobre o EPLIS.

Na análise descritiva das respostas dos alunos aos itens nesse componente, identificamos que a maioria dos alunos, 67%, disseram se sentir motivados pelo exame. Em relação à representatividade das tarefas do controlador no exame, 45% dos alunos não concordaram nem discordaram, enquanto 31% demonstraram concordância. Sobre o EPLIS avaliar precisamente o nível de proficiência dos alunos, as respostas mostraram maior variação: 37% concordaram, 23% discordaram e 33% não concordaram nem discordaram com essa alegação. Podemos presumir que isso pode estar ocorrendo devido à carência de conhecimentos sobre o EPLIS e sobre a função desempenhada por um controlador de tráfego aéreo, ambas

hipóteses necessitam ser investigadas mais profundamente. Em relação ao exame ser opcional, as respostas dividiram-se: 43% concordaram, enquanto 42% discordaram. Esse item se correlacionou negativamente com os demais. Isso significa que os alunos que discordaram de que o exame deva se tornar opcional atribuíram níveis de respostas mais altos aos itens ligados à motivação, autenticidade, confiabilidade e relevância do teste. Da mesma forma, os que concordaram com o caráter opcional do teste atribuíram níveis mais baixos para os demais itens que compõem a *significância do exame*.

Sabemos que a percepção do participante sobre a relevância de um exame contribui para a intensidade de seu efeito retroativo (GREEN, 2007a, 2013). Assim, acredito que seu impacto será, possivelmente, mais intenso para os alunos que veem o exame como um bom instrumento de avaliação, cujos resultados são consistentes e condizentes com o desempenho em atividades futuras e cuja aprovação é mais desejada. Esses se sentirão mais motivados a se prepararem para o teste.

O segundo componente, denominado *Preparação para o exame*, correlacionou itens sobre o alinhamento entre as aulas, os projetos, as avaliações desenvolvidas na EEAR e o EPLIS. As respostas dos alunos aos itens desse componente concentraram-se no nível de concordância. Porém, boa parte dos alunos demonstrou-se indecisa e optou pelo critério "nem concordo nem discordo". Esses dados também podem estar ligados à carência de conhecimentos sobre o teste, já exposta pelos alunos, o que pode acarretar a redução da intensidade do efeito retroativo para esses participantes.

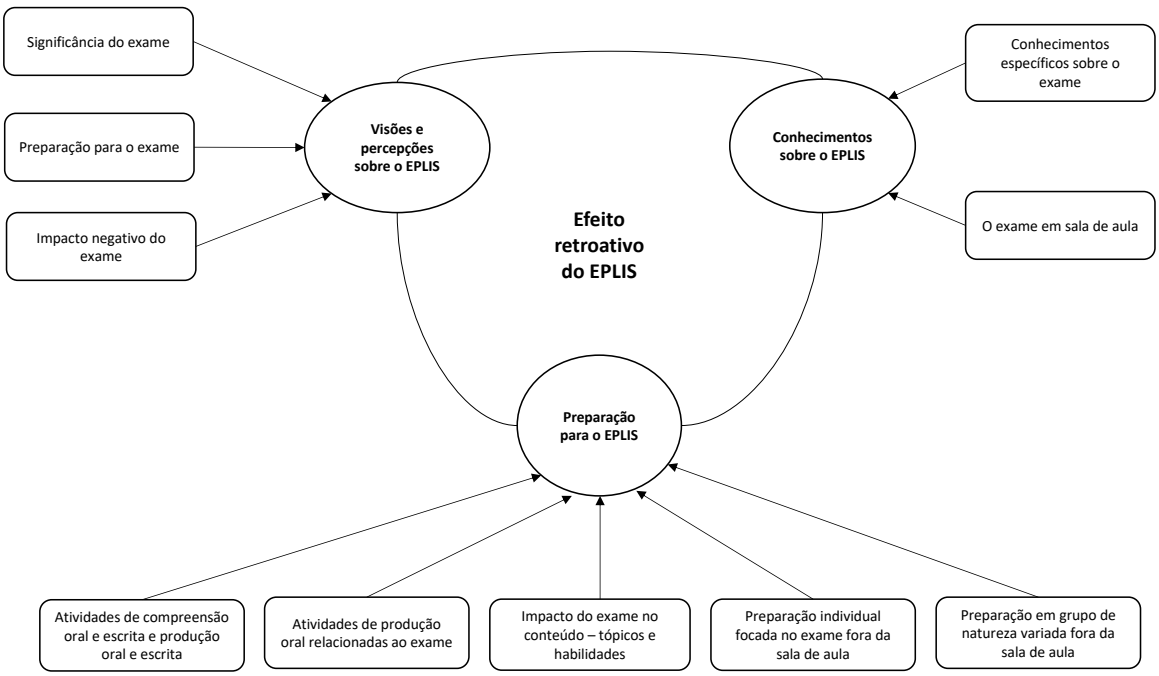
A última dimensão refere-se às percepções dos alunos sobre *o Impacto negativo do exame*. O item com maior carga fatorial aborda o aumento considerável na carga de estudos dos alunos provocado pelo teste. 43% dos alunos concordaram com essa situação, ao passo que 31% nem concordaram nem discordaram. Em relação ao EPLIS forçá-los a se dedicarem à aprendizagem de inglês fora do ambiente escolar, 52% dos alunos responderam afirmativamente, enquanto 24% discordaram desse fato. Somado a isso, 60% dos alunos indicaram que o exame aumenta o nível de ansiedade enquanto apenas 20% responderam negativamente à essa situação.

Como pontuei anteriormente, qualquer situação de avaliação irá causar algum impacto negativo, principalmente se a aprovação no exame for considerada importante para o participante. Todavia, o impacto negativo provocado por um teste deve ser minimizado para que práticas preparatórias indesejáveis sejam evitadas. 39% dos alunos indicaram que o exame os levou a alterar sua forma de estudar, sendo, portanto, necessário analisar mais profundamente a natureza dessa modificação e se ela condiz com os efeitos pretendidos.

4.4.5 Diagrama da estrutura fatorial do QA

A Figura 10 resume a estrutura fatorial presente no questionário do aluno. Para investigar os construtos do efeito retroativo do EPLIS na percepção dos alunos, foram analisadas as respostas dos participantes às 55 variáveis, divididas em três grandes construtos: *conhecimentos sobre o EPLIS*, *preparação para o EPLIS* e *visões e percepções sobre o EPLIS*. Das 55 variáveis analisadas, foi possível extrair 10 componentes principais que explicam os dados com perda mínima de informação.

Figura 10 − Diagrama da estrutura fatorial do questionário do aluno



4.4.6 Padrões de resposta entre os alunos

Sabemos que o efeito retroativo de um exame não ocorre da mesma forma para todos os envolvidos, podendo ser mais intenso para uns do que para outros (ALDERSON; WALL, 1993; ALDERSON; HAMP-LYONS, 1996). Por outro lado, Burrows (2004) argumenta que, mesmo reagindo diferentemente à influência de um exame, os participantes, no caso os professores, podem apresentar certos padrões de resposta. Acredito que o mesmo

também pode ocorrer com os alunos e, por essa razão, optei por investigar diferenças nas percepções e atitudes desses participantes em função do nível de proficiência declarado por eles, do nível de conhecimento geral sobre o EPLIS e da proximidade do teste. Os resultados da análise de componentes principais foram imprescindíveis para que essa averiguação pudesse ser realizada, pois, ao reduzirem 55 variáveis a 10 componentes, facilitaram a aplicação e a interpretação de testes de hipóteses que indicassem se a relação entre as variáveis era significativa em termos estatísticos. Foram utilizados dois tipos de testes de hipóteses: o teste *t* de *student* para amostras independentes e a análise de variância Anova com testes *post-hoc* de Scheffe e Hochberg. Essas técnicas compararam as médias das respostas dos seguintes subgrupos em relação aos componentes principais em busca de relações de significância entre eles:

1) alunos com baixa e alta proficiência linguística;
2) alunos com baixo e alto conhecimento sobre o EPLIS; e
3) alunos da 1ª, 2ª, 3ª e 4ª séries.

Em relação ao teste *t* de *student* para amostras independentes, foi necessário ajustar o valor de alfa (significância estatística), dividindo-o pelo número de componentes em cada construto a fim de reduzir o erro tipo 1.[61] Dessa forma, para o construto 1, consideramos p=0.025 (0.05/2); para o construto 2, p=0.01 (0.05/5); e para o construto 3, p=0.016 (0.05/3). Relembro que *p*-valor ou Sig indica se a hipótese nula será rejeitada ou não. Ou seja, se *p*-valor (Sig) for igual ou menor que o critério estabelecido para cada construto, a hipótese nula é rejeitada. Ou seja, existe diferença significativa entre as médias de respostas dos subgrupos analisados. Caso contrário, a hipótese nula não é rejeitada. Para a Anova, o *p*-valor utilizado foi de 0.05. Complementei os resultados do *p*-valor com o reporte do tamanho de efeito (TDE) para todas as comparações realizadas, garantindo, assim, mais rigor estatístico ao estudo.

4.4.6.1 Proficiência linguística autodeclarada

Para analisar a interferência do nível de proficiência linguística nas percepções e atitudes dos alunos em relação à influência do exame, dividi-os em dois subgrupos: com alta e baixa proficiência declarada. No subgrupo com alta proficiência, estão aqueles que

[61] O erro tipo I ocorre em testes de hipóteses quando a hipótese nula é rejeita mesmo sendo verdadeira, popularmente conhecido como "falso negativo".

responderam possuir nível intermediário ou avançado na aprendizagem da língua inglesa. No subgrupo com baixa proficiência, encontram-se os que se consideraram iniciantes ou no nível básico. Foram descartados aqueles que reportaram nível pré-intermediário por se posicionarem no centro. Na Tabela 50, é possível comparar as médias de respostas dos dois subgrupos às variáveis pertencentes a cada componente. Lembro que N se refere ao número total de respondentes em cada subgrupo. Na coluna "M", mostramos as médias de respostas de cada subgrupo de alunos por componente. DP é a abreviação de desvio padrão e refere-se ao quanto as respostas se distanciam da média reportada. EPM é a abreviação utilizada para erro padrão da média.

Tabela 50 − Média e desvio padrão das respostas dos subgrupos por proficiência

Construtos	componentes	Subgrupos	N	M	DP	EPM
Construto 1	Conhecimentos específicos sobre o exame	Baixa	113	2,413	0.838	0.078
		Alta	155	2.748	0.791	0.063
	O exame em sala de aula	Baixa	113	2.734	0.480	0.045
		Alta	155	3.051	0.473	0.037
Construto 2	Atividades de compreensão oral e escrita e produção oral e escrita	Baixa	113	2.991	1.004	0.094
		Alta	155	3.348	1.035	0.083
	Atividades de produção oral relacionadas ao exame	Baixa	113	3.182	0.855	0.094
		Alta	155	3.455	0.780	0.083
	Impacto do exame no conteúdo – tópicos e habilidades	Baixa	113	4.017	0.747	0.070
		Alta	155	4.187	0.676	0.054
	Preparação individual focada no exame fora da sala de aula	Baixa	113	1.512	0.785	0.073
		Alta	155	1.511	0.893	0.071
	Preparação em grupo de natureza variada fora da sala de aula	Baixa	113	2.316	0.816	0.076
		Alta	155	2.416	0.892	0.071
Construto 3	Significância do exame	Baixa	113	3.157	0.907	0.085
		Alta	155	3.432	0.821	0.066
	Preparação para o exame	Baixa	113	3.059	0.980	0.092
		Alta	155	3.352	0.856	0.068
	Impacto negativo do exame	Baixa	113	3.404	0.818	0.077
		Alta	155	3.190	0.880	0.070

Primeiramente, podemos observar que as médias de respostas do subgrupo mais proficiente (Alta) tendem a ser maiores do que as médias do subgrupo julgado com baixa proficiência (Baixa) na maior parte dos componentes extraídos. Essa diferença é mais expressiva nos componentes *Conhecimentos específicos sobre o exame*, *O exame em sala de aula*, *Atividades de compreensão oral e escrita e produção oral e escrita*, *Atividades de produção oral relacionadas ao exame*, *Significância do exame* e *Preparação para o exame*, o que nos leva a supor que o subgrupo com maior proficiência conhece melhor o exame, realiza mais atividades preparatórias em sala de aula, considera mais importante a aprovação, se sente mais motivado, tem visões mais positivas sobre o teste e avalia melhor o alinhamento entre ensino e avaliação externa. Por outro lado, o subgrupo menos proficiente parece possuir níveis mais baixos de conhecimentos sobre o exame, afirma realizar menos atividades de produção

oral em sala de aula, atribui ao exame menor importância, além de reportar níveis mais altos de impacto negativo, quando comparado ao subgrupo mais proficiente.

Esses resultados podem estar ligados à percepção do nível de dificuldade do exame por alunos e professores, o que interfere diretamente na intensidade do efeito retroativo (CHENG, 2005; GREEN, 2007a). Por considerarem o exame difícil, alunos com baixa proficiência tendem a desistir de qualquer tipo de preparação, o que faz com que o efeito retroativo diminua para esse subgrupo. Em relação aos professores, pelo fato de os alunos estarem distribuídos em turmas pelo nível de proficiência, podemos supor que, nas turmas mais proficientes, o professor ou instrutor desenvolve mais atividades de produção oral relacionadas ao exame, por considerar que esses alunos apresentam maior potencial para serem aprovados, além de mencionar e discutir aspectos do exame com mais frequência.

Em relação à preparação individual focada no exame fora da sala de aula, praticamente não houve diferença nas médias de respostas dos subgrupos com alta e baixa proficiência. As médias no componente *Atividades preparatórias de natureza variada fora da sala de aula* também são muito próximas, porém um pouco mais elevadas, indicando que os alunos são mais propensos a realizar atividades de natureza variada do que voltadas para o formato do teste. De qualquer forma, as médias em ambos são consideradas baixas e sugerem que, independente do nível de proficiência, os alunos raramente se engajam em atividades preparatórias para o exame fora do ambiente de sala de aula, o que pode estar relacionado à percepção de relevância.

Em termos estatísticos, a Tabela 51 resume os resultados obtidos para o teste *t* de *student* para amostras independentes. O *t*-valor e o Sig. 2 tailed (*p*-valor) indicam se a diferença entre as respostas dos subgrupos pode ser considerada significativa ou não. Como regra, quanto maior o *t*-valor, menor será o *p*-valor. Como mencionei anteriormente, para os componentes dos construtos 1, 2 e 3, utilizei como índice de significância $p <= 0.025$, $p <= 0.010$ e $p <= 0.016$, respectivamente. Os valores marcados com sinal de asterisco (*) indicam que a diferença entre as médias analisadas apresentou significância estatística.

Podemos observar que as diferenças entre médias são consideradas estatisticamente significantes em seis componentes: *Conhecimentos específicos sobre o exame*, *O exame em sala de aula*, *Atividades de compreensão oral e escrita e produção oral e escrita*, *Atividades de produção oral relacionadas ao exame*, *Significância do exame* e *Preparação para o exame*. Esses resultados validam as suposições de que os alunos mais proficientes possuem mais conhecimentos sobre o exame, se engajam com mais frequência em atividades voltadas para o exame em sala de aula, consideram mais importante o sucesso no exame, se sente mais

motivados, têm visões mais positivas sobre a autenticidade e a confiabilidade do teste e percebem melhor o alinhamento entre ensino e avaliação externa.

Tabela 51 − Resultados do teste t para o nível de proficiência

					t-test para igualdade de médias		
Construtos	**componentes**	**Sub grupos**	**N**	**SD**	**t (t-valor)**	**df**	**Sig.(2tailed) (p-valor)**
Construto 1 p<= 0.025	Conhecimentos específicos sobre o exame	Baixa	113	0.838	-3.338	266	0.001*
		Alta	155	0.791			
	O exame em sala de aula	Baixa	113	0.480	-5.381	266	0.000*
		Alta	155	0.473			
Construto 2 p<=0.010	Atividades de compreensão oral e escrita e produção oral e escrita	Baixa	113	1.004	-2.584	266	0.005*
		Alta	155	1.035			
	Atividades de produção oral relacionadas ao exame	Baixa	113	1.004	-2.710	266	0.007*
		Alta	155	1.035			
	Impacto do exame no conteúdo – tópicos e habilidades	Baixa	113	0.747	-1.936	266	0.54
		Alta	155	0.676			
	Preparação individual focada no exame fora da sala de aula	Baixa	113	0.785	0.13	266	.990
		Alta	155	0.893			
	Preparação em grupo de natureza variada fora da sala de aula	Baixa	113	0.816	-0.936	266	0.818
		Alta	155	0.892			
Construto 3 p<=0.016	Significância do exame	Baixa	113	0.907	-2.590	266	0.010*
		Alta	155	0.821			
	Preparação para o exame	Baixa	113	0.980	-2.605	266	0.010*
		Alta	155	0.856			
	Impacto negativo do exame	Baixa	113	0.818	2.028	266	0.044
		Alta	155	0.880			

Esses resultados alinham-se aos alcançados por Cheng et al (2011) e Stenlund et al (2017) em estudos essencialmente quantitativos. Cheng et al (2011) também constatou que alunos mais proficientes apresentavam maior motivação para o estudo, realizavam mais atividades dentro e fora da sala de aula, possuíam visões mais fortes sobre exames externos e se sentiam menos preocupados com os desafios impostos pelo teste, quando comparados aos alunos com proficiência julgada menor. No estudo de Stenlund et al (2017), examinandos com alto desempenho reportaram níveis mais altos de conhecimentos sobre o teste e seu conteúdo, disseram se sentir mais motivados e menos ansiosos, comparados aos examinandos com baixo desempenho. O único aspecto que parece diferenciar esta pesquisa e os estudos de Cheng et al (2011) e Stenlund et al (2017) refere-se à ansiedade provocada pelo teste. Apesar de termos notado diferença entre as médias dos subgrupos no componente *Impacto negativo*, que engloba questões ligadas à ansiedade e *stress* provocados pelo teste, essa diferença não foi considerada estatisticamente significativa.

Como explicado no capítulo anterior, apesar do *p*-valor informar a existência de efeito, rejeitando, assim, a hipótese nula de que as médias são iguais, ele não revela a extensão dessa diferença. Além disso, o *p*-valor pode ser afetado pelo tamanho da amostra (N), ou seja, um valor baixo de *p*-valor pode ser apenas reflexo de uma amostra grande. Por essa razão,

reporto aqui os cálculos referentes aos tamanhos do efeito (*d* de Cohen) para cada um dos casos analisados, inclusive para aqueles em que não foi constatada significância estatística entre os componentes principais e o nível de proficiência dos alunos (Tabela 52). Marquei com um asterisco (*) as diferenças que tiveram significância estatística confirmadas pelo *p*-valor.

Tabela 52 − Tamanhos de efeito para a relação entre proficiência e componentes principais

Construto	**componentes**	***d* de Cohen**
Construto 1	Conhecimentos específicos sobre o exame	0.41*
	O exame em sala de aula	0.66*
Construto 2	Atividades de compreensão oral e escrita e produção oral e escrita	0.34*
	Atividades de produção oral relacionadas ao exame	0.33*
	Impacto do exame no conteúdo – tópicos e habilidades	0.24
	Preparação individual focada no exame fora da sala de aula	0.02
	Preparação em grupo de natureza variada fora da sala de aula	0.12
Construto 3	Significância do exame	0.32*
	Preparação para o exame	0.32*
	Impacto negativo do exame	0.25

De acordo com os índices de referência propostos por Cohen (1988), podemos notar que, com exceção do componente *O exame em sala de aula*, o tamanho do efeito para os casos em que as diferenças foram plotadas como significativas é considerado pequeno ($0.20 < d < 0.49$). Isso significa que o efeito, embora não seja trivial, é pouco significativo. Já para a relação entre a proficiência e o exame em sala de aula, o efeito é considerado médio ($0.50 < d < 0.79$), aumentando, assim, também a sua significância prática. Ressalto que essas hipóteses, mesmo as que apresentaram tamanho de efeito pequeno, serão exploradas na fase qualitativa que, por sua vez, trará mais subsídios para rejeitá-las ou confirmá-las.

4.4.6.2 Nível de conhecimento sobre o EPLIS

O grau de conhecimento sobre o exame é outro fator que pode interferir na intensidade do efeito retroativo. Por essa razão, examinei se as atitudes e percepções dos alunos variam em função do quanto eles sabem a respeito do teste. Dividi, então, os alunos em dois subgrupos de acordo com o nível declarado de conhecimento geral sobre o EPLIS. O subgrupo com baixo conhecimento é formado por aqueles que assinalaram no questionário os níveis "nenhum" ou "pouco" para a questão 10.[62] O subgrupo com alto conhecimento está composto

[62] Q10 – De forma geral, como você avalia o seu nível de conhecimento sobre o EPLIS?

por aqueles que responderam “bom” ou “ótimo”. Foram desprezadas as respostas no nível “razoável” por estarem em uma posição intermediária.

Na Tabela 53, é possível comparar as médias de respostas dos dois subgrupos nos componentes extraídos. Foi preciso retirar dessa análise o construto *Conhecimentos sobre o EPLIS* para evitar resultados redundantes, uma vez que os subgrupos estão divididos pelo nível de conhecimento que possuem sobre o exame. Primeiramente, notamos uma diferença expressiva no número de participantes em cada grupo (N): enquanto 55 alunos assumem que conhecem bem o teste, aproximadamente três vezes esse número, ou seja, 169 alunos alegam possuir pouco conhecimento.

Tabela 53 − Média e desvio padrão das respostas dos subgrupos por nível de conhecimento sobre o exame

Construtos	componentes	Subgrupos	N	Média	DV	EPM
Construto 2	Atividades de compreensão oral e escrita e produção oral e escrita	Baixo	169	3.198	1.024	0.078
		Alto	59	2.890	1.161	0.155
	Atividades de produção oral relacionadas ao exame	Baixo	169	3.017	0.772	0.059
		Alto	59	3.655	0.806	0.105
	Impacto do exame no conteúdo – tópicos e habilidades	Baixo	169	3.960	0.785	0.060
		Alto	59	4.227	0.650	0.084
	Preparação individual focada no exame fora da sala de aula	Baixo	169	1.352	0.645	0.049
		Alto	59	2.067	1.141	0.148
	Preparação em grupo de natureza variada fora da sala de aula	Baixo	169	2.263	0.770	0.059
		Alto	59	2.745	1.045	0.136
Construto 3	Significância do exame	Baixo	169	3.187	0.890	0.068
		Alto	59	3.364	0.801	0.104
	Preparação para o exame	Baixo	169	3.179	0.879	0.067
		Alto	59	3.265	0.870	0.113
	Impacto negativo do exame	Baixo	169	3.226	0.834	0.064
		Alto	59	3.457	0.830	0.108

O subgrupo que alega ter bom ou ótimo conhecimento sobre o exame apresenta médias maiores quando comparado ao grupo com pouco conhecimento, sendo essa diferença mais expressiva em relação aos componentes pertencentes ao construto 2, com exceção do primeiro componente *Atividades de compreensão oral e escrita e produção oral e escrita*. Isso sugere que os alunos que conhecem melhor o teste, participam de mais atividades preparatórias dentro e fora da sala de aula, estão mais conscientes da importância da aprovação e mais motivados e apresentam atitudes mais positivas para com a avaliação externa.

O teste *t* de *student* apontou que a diferença é significativa nos seguintes componentes: *Atividades de produção oral relacionadas ao exame*, *Preparação individual focada no exame fora da sala de aula* e *Preparação em grupo de natureza variada fora da sala de aula*, cujos *p*-valores são menores que 0.010, como podemos observar na última coluna da Tabela 54. Esses resultados apontam que os alunos que julgam melhor conhecer o exame

tendem a realizar, em sala de aula, mais atividades de produção oral semelhantes àquelas encontradas no teste, e se engajam com mais frequência em atividades preparatórias fora da sala de aula. Xie e Andrews (2012) encontraram resultados parecidos que conectam a consciência sobre as exigências do teste ao envolvimento em atividades preparatórias, sendo essa relação mediada pelo valor atribuído à relevância do exame e à expectativa de sucesso.

Tabela 54 − Resultados do teste t para o nível de conhecimento sobre o exame

					t-test para igualdade de médias		
Construtos	**componentes**	**Subgrupos**	**N**	**DP**	**t (t-valor)**	**df**	**Sig.(2tailed) (p-valor)**
Construto 2 p<=0.010	Atividades de compreensão oral e escrita e produção oral e escrita	Baixo	169	1.024	-1.924	226	0.56
		Alto	59	1.161			
	Atividades de produção oral relacionadas ao exame	Baixo	169	0.772	-5.393	226	0.000*
		Alto	59	0.806			
	Impacto do exame no conteúdo–tópicos e habilidades	Baixo	169	0.785	-2.338	226	0.20
		Alto	59	0.650			
	Preparação individual focada no exame fora da sala de aula	Baixo	169	0.645	-5.893	226	0.000*
		Alto	59	1.141			
	Preparação em grupo de natureza variada fora da sala de aula	Baixo	169	0.770	-3.756	226	0.000*
		Alto	59	1.045			
Construto 3 p<=0.016	Significância do exame	Baixo	169	0.890	-1.344	226	0.18
		Alto	59	0.801			
	Preparação para o exame	Baixo	169	0.879	-0.645	226	0.519
		Alto	59	0.870			
	Impacto negativo do exame	Baixo	169	0.834	-1.836	226	0.068
		Alto	59	0.830			

Ressalvo, entretanto, que, apesar de haver diferença significativa nas atitudes dos alunos com maior ou menor nível de conhecimento geral sobre o exame em relação à preparação fora do ambiente escolar, a média de resposta para a prática de atividades no mesmo formato do teste é ainda baixa (2.067), o que pode indicar pouca relevância do exame para os alunos de modo geral. Em relação aos componentes do construto 3, Visões e percepções sobre o exame, as médias de resposta estão muito próximas entre os dois subgrupos e, por essa razão, não se constatou relação entre o conhecimento geral sobre o EPLIS e o modo como os alunos percebem a significância do exame, a preparação recebida e seu impacto negativo. Os resultados para tamanhos de efeito para as relações entre o conhecimento sobre o EPLIS e os componentes podem ser observados na Tabela 55.

Os tamanhos dos efeitos foram considerados altos (d > 0.80) nos componentes *Atividades de produção oral relacionadas ao exame* e *Preparação individual focada no exame fora da sala de aula*, demonstrando que essas relações também apresentam alta significância prática. A *Preparação em grupo de natureza variada fora da sala de aula*, por sua vez, alcançou nível médio de significância prática. Com exceção da relação envolvendo a *preparação para o exame* (construto 3), cujo tamanho de efeito foi considerado insignificante (d<0.19), as demais

associações demonstraram tamanhos de efeito pequenos (0.20 < d < 0.49). No entanto, pelo fato de seus valores de significância estatística (*p*-valor) terem excedido o limite máximo adotado, elas não foram consideradas hipóteses a serem investigadas mais a fundo na fase qualitativa do estudo.

Tabela 55 − Tamanhos de efeito para a relação entre conhecimento sobre o exame e componentes principais

Construto	**componentes**	***d* de Cohen**
Construto 2	Atividades de compreensão oral e escrita e produção oral e escrita	0.29
	Atividades de produção oral relacionadas ao exame	0.81*
	Impacto do exame no conteúdo– tópicos e habilidades	0.35
	Preparação individual focada no exame fora da sala de aula	0.89*
	Preparação em grupo de natureza variada fora da sala de aula	0.57*
Construto 3	Significância do exame	0.20
	Preparação para o exame	0.10
	Impacto negativo do exame	0.28

4.4.6.3 Proximidade do exame

A proximidade do exame é outro fator capaz de interferir na intensidade do efeito retroativo. Watanabe (1996), por exemplo, argumenta sobre a possibilidade de os métodos de ensino sofrerem mudanças em função do quão perto estão alunos e professores da data de realização do teste. Xie e Andrews (2012, p. 55, tradução nossa), com base em Prodomou (1995), discutem dois tipos de efeito retroativo relacionados à proximidade do exame – o implícito (*covert*) e o explícito (*overt*), e pontuam que próximo à realização do teste o efeito retroativo torna-se mais "intensivo, observável, e portanto, explícito".[63] Verificamos, portanto, se a série em que os alunos se encontravam apresentava alguma diferença na maneira como eles percebiam e agiam em relação ao EPLIS.

A Tabela 56 apresenta as médias e os desvios padrão das respostas dos alunos em cada componente de acordo com a série. Na maioria dos componentes, observamos que há um aumento progressivo das médias, com pequenas variações em alguns casos, corroborando a suposição de que o efeito retroativo do exame se intensifica quanto mais próximos estão os alunos de sua realização. No entanto, as médias relativas à *Significância do exame* e à *Preparação para o exame* diminuem com a proximidade do teste, indicando resultados atípicos.

[63] No original: "intensive, observable, and hence overt".

Tabela 56 − Média e desvio padrão das respostas dos subgrupos por série

Construtos	componentes	Subgrupos	N	Média	DP	EPM
Construto 1	Conhecimento específico sobre o exame	1ª série	116	2.262	0.784	0.072
		2ª série	116	2.615	0.831	0.077
		3ª série	109	2.689	0.774	0.074
		4ª série	46	3.332	0.705	0.103
	O exame em sala de aula	1ª série	116	3.063	0.601	0.055
		2ª série	116	3.287	0.618	0.057
		3ª série	109	3.501	0.567	0.054
		4ª série	46	3.841	0.528	0.077
Construto 2	Atividades de compreensão oral e escrita e produção oral e escrita	1ª série	116	2.135	0.820	0.076
		2ª série	116	3.577	0.877	0.081
		3ª série	109	3.563	0.695	0.066
		4ª série	46	3.688	0.808	0.119
	Atividades de produção oral relacionadas ao exame	1ª série	116	2.793	0.766	0.071
		2ª série	116	3.354	0.802	0.074
		3ª série	109	3.452	0.679	0.065
		4ª série	46	4.021	0.644	0.095
	Impacto do exame no conteúdo– tópicos e habilidades	1ª série	116	3.772	0.738	0.068
		2ª série	116	4.177	0.720	0.066
		3ª série	109	4.249	0.621	0.059
		4ª série	46	4.282	0.684	0.100
	Preparação individual focada no exame fora da sala de aula	1ª série	116	1.387	0.644	0.059
		2ª série	116	1.413	0.688	0.063
		3ª série	109	1.433	0.713	0.068
		4ª série	46	2.361	1.157	0.170
	Preparação em grupo de natureza variada fora da sala de aula	1ª série	116	2.437	0.872	0.081
		2ª série	116	2.327	0.861	0.080
		3ª série	109	2.502	0.867	0.083
		4ª série	46	2.608	0.842	0.124
Construto 3	Significância do exame	1ª série	116	3.489	0.846	0.078
		2ª série	116	3.107	0.953	0.088
		3ª série	109	3.422	0.787	0.075
		4ª série	46	3.016	0.858	0.126
	Preparação para o exame	1ª série	116	3.453	0.776	0.072
		2ª série	116	3.106	0.991	0.092
		3ª série	109	3.452	0.833	0.079
		4ª série	46	2.818	0.882	0.130
	Impacto negativo do exame	1ª série	116	3.290	0.875	0.081
		2ª série	116	3.198	0.897	0.083
		3ª série	109	3.415	0.766	0.073
		4ª série	46	3.434	0.721	0.106

A fim de atestar se as diferenças plotadas na Tabela 56 são estatisticamente significativas, utilizei a análise de variância conhecida como Anova que permite testar a hipótese de duas ou mais médias serem iguais ou não. A hipótese nula (H_0) assume que as médias são iguais, enquanto a hipótese alternativa (H_1) indica que, pelo menos, uma média é diferente. No entanto, a Anova não é capaz de mostrar quais subgrupos se diferem. Para isso, empreguei os testes *post-hoc* de Scheffe e de Hotchberg, dos quais somente os resultados do primeiro são reportados neste estudo. Dessa forma, pude identificar quais subgrupos compartilhavam os mesmos traços ou apresentavam características distintas. O nível de significância utilizado foi α=0.05. Além da significância estatística, também reporto os tamanhos de efeito para cada par ou tratamento considerado analisado. As Tabelas de 57 a 66

apresentam os resultados obtidos nas comparações entre as séries, combinando todas as possibilidades. Tratamentos com *p*-valor igual ou abaixo de 0.05 indicam, portanto, hipótese alternativa, ou seja, o comportamento daqueles dois subgrupos se difere dentro do componente analisado. O tamanho do efeito é apresentado na última coluna (*d* de Cohen).

Podemos observar, na Tabela 57, que o teste de hipótese Anova indica diferenças significativas entre as médias dos subgrupos analisados, com *p*-valor (Sig.) inferior a 0.05. O teste de Scheffe, por sua vez, retrata quais subgrupos se diferem estatisticamente. O sinal de asterisco (*) no valor das médias marca que houve significância, e o sinal negativo (-), na coluna tratamento, demonstra que a média do primeiro subgrupo é menor que a do segundo subgrupo. Por exemplo, na comparação entre as médias das 1ª e 2ª séries, a diferença foi calculada em -.352*. O asterisco indica significância, confirmada pelo *p*-valor de 0.009. O sinal negativo registra que a média da 1ª série é menor que a da 2ª série. Se voltarmos à Tabela 55, veremos que a média para a 1ª série, no componente *Conhecimentos específicos sobre o exame*, foi de 2.262, enquanto a 2ª série alcançou média 2.615.

Tabela 57 − Diferença entre séries no componente 1 do construto 1

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Conhecimentos específicos sobre o exame	20.788	.000	1ª – 2ª	-0.352*	0.103	0.009	0.43
			1ª – 3ª	-0.426*	0.105	0.001	0.54
			1ª – 4ª	-1.069*	0.137	0.000	1.40
			2ª – 3ª	-0.073	0.105	0.920	0.09
			2ª – 4ª	-0.716*	0.137	0.000	0.90
			3ª – 4ª	-0.642*	0.138	0.000	0.85

Esses resultados revelam que a 1ª e a 4ª séries se distinguem significativamente das demais em relação ao nível de conhecimentos específicos que possuem sobre o teste. Assim, a 1ª série é a que menos conhece o exame, o que coincide com o programa curricular da escola que prevê o ensino de inglês geral para esse subgrupo. A 4ª série é a que alcançou média mais alta e se diferiu das demais, sugerindo que esse subgrupo, pela proximidade do teste, adquire níveis mais elevados de conhecimentos sobre suas características. Ademais, os tamanhos dos efeitos para todas as associações em que a 4ª série aparece são considerados bastante expressivos (*d* de Cohen > 0.80), indicando que essas diferenças apresentam alta significância prática. Observem também que, na relação entre a 1ª e a 4ª série, o TDE alcança o valor mais alto (*d =1.40*), sugerindo que essa diferença é a mais relevante dentre todas.

Em relação à abordagem do *O exame em sala de aula*, a Anova indicou diferença significativa entre os subgrupos com *p*-valor menor que 0.05, conforme mostrado na Tabela 58.

Observamos, novamente, que a média da 1ª série se difere significativamente das 2ª, 3ª e 4ª séries, sugerindo que o exame é menos mencionado e menos discutido pelos professores em sala de aula naquela série. A 4ª série também se distinguiu das demais, obtendo a maior média de respostas. Isso nos leva a inferir que os professores que atuam nessa série procuram mencionar e discutir aspectos do exame com os alunos com maior frequência dada a sua proximidade. Os tamanhos de efeito são maiores nas comparações entre séries mais distantes. Por exemplo, entre a 1ª e a 4ª série a extensão do efeito supera 1.30, evidenciando um grau alto de significância prática. Entre séries muito próximas, como é o caso da 1ª e 2ª séries, a extensão do efeito é considerada pequena ($0.20 < d < 0.50$), pontuando que a diferença entre as médias dessas duas séries é menor e, portanto, não tão relevante em termos práticos.

Tabela 58 − Diferença entre séries no componente 2 do construto 1

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
O exame em sala de aula	22.650	.000	1ª – 2ª	-0.224*	0.077	0.040	0.36
			1ª – 3ª	-0.438*	0.078	0.000	0.75
			1ª – 4ª	-0.777*	0.102	0.000	1.33
			2ª – 3ª	-0.214	0.078	0.061	0.36
			2ª – 4ª	-0.553*	0.102	0.000	0.93
			3ª – 4ª	-0.339*	0.103	0.014	0.61

No componente *Atividades de compreensão oral e escrita e produção oral e escrita*, (Tabela 59), a 1ª série apresenta diferença significativa comparada às demais séries, atingindo a menor média. Esse resultado coincide com o enfoque dado ao ensino de inglês geral nessa série. Lembro que os itens nesse componente se referem a atividades cujos temas giram em torno da aviação e do controle de tráfego aéreo, tais como, leitura de textos da área, compreensão de áudios de radiotelefonia e produção de diálogos entre piloto e controlador. Da 2ª série em diante, é introduzido o ensino para fins específicos e, por essa razão, esses tipos de atividades são trabalhados em uma frequência bastante semelhante nessas séries.

Tabela 59 − Diferença entre séries no componente 1 do construto 2

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Atividades de compreensão oral e escrita e produção oral e escrita	89.000	.000	1ª – 2ª	-1.442*	0.105	0.000	1.70
			1ª – 3ª	-1.428*	0.107	0.000	1.88
			1ª – 4ª	-1.552*	0.140	0.000	1.90
			2ª – 3ª	0.013	0.107	0.999	0.01
			2ª – 4ª	-0.110	0.140	0.891	0.12
			3ª – 4ª	-0.124	0.141	0.856	0.17

Os tamanhos dos efeitos nas comparações envolvendo a 1ª série alcançaram valores muito altos ($d > 0.80$), indicando que a diferença entre essa série e as demais é extremamente relevante em termos de significância prática. Para as demais associações, os valores para *d* de Cohen foram considerados insignificantes ($d < 0.19$).

Na Tabela 60, estão reportadas as diferenças entre as médias das quatro séries em relação a *Atividades de produção oral relacionadas ao exame*. Percebemos, primeiramente, que a 1ª série se difere significativamente das demais, uma vez que nesse grupo não são trabalhados assuntos pertinentes ao controle de tráfego aéreo. A 4ª série também se distingue das 1ª, 2ª e 3ª séries, o que pode significar que os professores na 4ª série tendem a desenvolver mais atividades de produção oral voltadas para o EPLIS, tais como descrição de figuras, relato de histórias, oferta de sugestão para situações de tráfego aéreo, entre outras. Em termos de tamanho de efeito, com exceção da relação entre a 1ª e 2ª séries, cuja extensão foi considerada mediana ($0{,}50 < d < 0.79$), todas as demais comparações envolvendo a 1ª e a 4ª série demonstraram valores para *d* de Cohen superiores a 0.80, assegurando que as diferenças entre as médias contém alta significância prática, além de significância estatística atestada pelo *p*-valor inferior a 0.05.

Tabela 60 − Diferença entre séries no componente 2 do construto 2

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Atividades de produção oral relacionadas ao exame	34.352	.000	**1ª – 2ª**	-0.560*	0.097	0.000	0.71
			1ª – 3ª	-0.658*	0.098	0.000	0.91
			1ª – 4ª	-1.227*	0.129	0.000	1.68
			2ª – 3ª	-0.097	0.098	0.805	0.13
			2ª – 4ª	-0.666*	0.129	0.000	0.87
			3ª – 4ª	-0.569*	0.130	0.000	0.85

No componente *Impacto do exame no conteúdo – tópicos e habilidades*, foi encontrada diferença significativa entre a 1ª e as demais séries, conforme podemos verificar na Tabela 61. Esse resultado parece estar relacionado ao fato de o conteúdo abordado na 1ª série privilegiar temas gerais, tais como descrição física e psicológica de pessoas, turismo, vida de famosos, bastante diferentes dos temas e tópicos específicos de tráfego aéreo tratados nas outras séries. Os tamanhos dos efeitos para as comparações estatisticamente significantes (*p*-valor < 0.05) foram considerados medianos, com *d* de Cohen entre 0.50 e 0.79. As demais relações apresentaram valores para TDE insignificantes.

Não obstante, a média de respostas da 1ª série (3.772) nesse componente, apesar de ser menor que as médias dos demais subgrupos, é ainda considerada alta, dado que o currículo

nessa série é voltado para o inglês geral. Acredito que isso possa ter ocorrido devido à presença de temas relacionados à aviação como, por exemplo, refeições a bordo de aeronaves, nessa série. Embora esses temas não façam parte da comunicação aeronáutica, os alunos, por possuírem ainda poucos conhecimentos sobre o EPLIS e sobre o ambiente profissional, podem terem sido induzidos a avaliar como razoável a frequência com que os professores trabalham tópicos relacionados ao controle de tráfego aéreo.

Tabela 61 − Diferença entre séries no componente 3 do construto 2

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Impacto do exame no conteúdo – tópicos e habilidades	11.755	.000	**1ª – 2ª**	-0.405*	0.091	0.000	0.55
			1ª – 3ª	-0.477*	0.092	0.000	0.70
			1ª – 4ª	-0.510*	0.121	0.001	0.70
			2ª – 3ª	-0.071	0.092	0.896	0.10
			2ª – 4ª	-0.105	0.121	0.861	0.14
			3ª – 4ª	-0.033	0.122	0.995	0.05

Em relação à *Preparação individual focada no exame fora da sala de aula*, é possível observar, na Tabela 62, diferença estatisticamente significativa (p-valor < 0.05) entre a 4ª série e as demais, indicando que os alunos que estão prestes a realizarem o exame tendem a se engajar mais em atividades preparatórias no mesmo formato do teste. A significância prática das associações consideradas estatisticamente significantes também é confirmada com valores de *d* de Cohen superiores a 0.80, registrando uma extensão de efeito grande. Ainda assim, notamos que as médias de respostas dos alunos, em geral, são baixas nesse componente, variando de 1.387 (1ª série) a 2.361 (4ª série). Em outras palavras, mesmo havendo uma diferença significativa entre a 4ª série e as demais, fica evidente que os alunos costumam se engajar pouco em práticas preparatórias extraclasse, o que pode estar relacionado a uma provável baixa relevância do exame entre esses participantes.

Tabela 62 − Diferença entre séries no componente 4 do construto 2

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Preparação individual focada no exame fora da sala de aula	21.567	.000	**1ª – 2ª**	-0.025	0.098	0.995	0.03
			1ª – 3ª	-0.046	0.100	0.975	0.06
			1ª – 4ª	-0.974*	0.131	0.000	1.23
			2ª – 3ª	-0.020	0.100	0.998	0.02
			2ª – 4ª	-0.948*	0.131	0.000	1.15
			3ª – 4ª	-0.927*	0.132	0.000	1.09

No componente *Preparação em grupo de natureza variada fora da sala de aula* (Tabela 63), não foi encontrada nenhuma diferença de médias estatisticamente significativa nas

respostas dos alunos das quatro séries. As médias foram muito semelhantes, com pequena variação de 2.437 (1ª série) a 2.608 (4ª série), e os tamanhos dos efeitos foram considerados insignificantes ou pequenos. Como dito anteriormente, são médias relativamente baixas que indicam que os alunos não costumam participar de atividades fora do ambiente da sala de aula com o objetivo de se prepararem para o EPLIS. Esse dado também nos aponta que, provavelmente, o exame não está alcançando um nível de relevância expressiva no cenário escolar.

Tabela 63 − Diferença entre séries no componente 5 do construto 2

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Preparação em grupo de natureza variada fora da sala de aula	1.425	.235	**1ª – 2ª**	0.109	0.113	0.816	0.12
			1ª – 3ª	-0.064	0.115	0.957	0.07
			1ª – 4ª	-0.171	0.150	0.731	0.19
			2ª – 3ª	-0.174	0.115	0.514	0.20
			2ª – 4ª	-0.281	0.150	0.324	0.32
			3ª – 4ª	0.106	0.151	0.921	0.12

Em relação à *significância do exame*, as 1ª e 4ª séries divergiram significativamente, assim como as 1ª e 2ª séries (Tabela 64). Os tamanhos dos efeitos dessas duas associações foram considerados medianos. No entanto, a variação decrescente entre as médias, percebida pelos valores positivos na coluna "Diferença entre médias", causa preocupação. Enquanto a 1ª série apresentou média de 3.489, os alunos da 4ª série alcançaram média de 3.016. Esse dado parece sugerir que os alunos da 4ª série estão incertos quanto à importância da aprovação no exame, a autenticidade de suas tarefas e a confiabilidade de seus resultados, indo ao contrário da expectativa de que, com a proximidade do teste, os alunos se tornariam mais conscientes de sua relevância. A meu ver, essa situação pode indicar que, com o passar das séries, o exame perde sua significância, provavelmente, pela presença de um outro fator ainda desconhecido e que necessita ser esclarecido.

Tabela 64 − Diferença entre séries no componente 1 do construto 3

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Significância do exame	6.119	.000	**1ª – 2ª**	0.381*	0.1137077	0.011	0.42
			1ª – 3ª	0.067	0.1155188	0.953	0.08
			1ª – 4ª	0.472*	0.1508875	0.021	0.55
			2ª – 3ª	-0.314	0.1155188	0.062	0.36
			2ª – 4ª	0.091	0.1508875	0.947	0.09
			3ª – 4ª	0.405	0.1522571	0.070	0.50

Ademais, outro dado que requer atenção é o fato de a comparação entre 3ª e 4ª séries ter alcançado um tamanho médio de efeito (*d*=0.50). Embora esse valor seja considerado o mínimo dentro da referência (0,50 < d < 0.79), seria ingênuo desconsiderá-lo, uma vez que o *p*-valor também atingiu valor próximo ao nível de significância adotado (α=0.05). Isso aponta uma queda considerável entre a 3ª e 4ª séries em relação à significância do teste, respaldando a hipótese levantada anteriormente de que a relevância do exame é diminuída na última série do curso.

No componente *Preparação para o exame*, que aborda questões relacionadas ao alinhamento entre o ensino e o exame, notamos diferenças significativas nos seguintes subgrupos: 1ª e 2ª séries, 2ª e 3ª séries, 1ª e 4ª séries e 3ª e 4ª séries, conforme Tabela 65. Os tamanhos dos efeitos das associações envolvendo a 4ª série alcançam valores quase duas vezes maiores do que os valores das demais associações, sugerindo que aquelas diferenças são mais expressivas.

Tabela 65 − Diferença entre séries no componente 2 do construto 3

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Preparação para o exame	8.738	.000	**1ª – 2ª**	0.347*	0.114	0.028	0.39
			1ª – 3ª	0.001	0.116	1.000	0.01
			1ª – 4ª	0.635*	0.152	0.001	0.78
			2ª – 3ª	-0.346*	0.116	0.033	0.37
			2ª – 4ª	0.287	0.152	.313	0.29
			3ª – 4ª	0.633*	0.153	0.001	0.74

Ao verificarmos as médias das respostas novamente (Tabela 56), percebemos que a 4ª série possui a menor média (abaixo de 3), enquanto as demais alcançaram média acima de 3. Esse resultado pode estar relacionado ao fato de que os alunos da 4ª série afirmam possuir maior conhecimento sobre o exame e, por estarem mais próximos da realização do teste, conseguem perceber discrepâncias entre o ensino oferecido e a avaliação externa com mais nitidez do que os demais. É necessário, no entanto, confirmar essa hipótese e analisar a natureza dessas discrepâncias na perspectiva dos alunos.

Por fim, embora a 4ª série apresente média maior no componente *Impacto negativo do exame* não foi encontrada diferença estatisticamente significativa entre as séries, conforme constatado na Tabela 66, na qual se vê valores de *p*-valor (Sig.) superiores ao nível de significância estatística adotado (α=0.05). Os tamanhos dos efeitos foram considerados pequenos ou insignificantes. Portanto, não existem evidências estatísticas de que o impacto

negativo em termos de ansiedade, *stress* e sobrecarga aumente com a proximidade da realização do exame, o que dá mais respaldo à suposição de que o teste carece em relevância.

Tabela 66 − Diferença entre séries no componente 3 do construto 3

componente	Anova		Scheffe				*d* de Cohen
	F	Sig.	Tratamentos	Dif. médias	SEM	Sig.	
Impacto negativo do exame	1.621	.184	1ª – 2ª	0.092	0.109	0.870	0.10
			1ª – 3ª	-0.124	0.111	0.744	0.15
			1ª – 4ª	-0.143	0.145	0.807	0.17
			2ª – 3ª	-0.216	0.111	0.288	0.25
			2ª – 4ª	-0.236	0.145	0.452	0.27
			3ª – 4ª	-0.019	0.147	0.999	0.02

Em suma, a análise da relação entre a proximidade e os componentes principais sugeriu que os alunos da 4ª série possuem níveis mais elevados de conhecimentos específicos sobre o EPLIS. Em suas aulas, o exame parece ser mais mencionado e discutido, e tarefas com foco na produção oral, semelhantes às do teste, são trabalhadas com maior frequência. Eles também se engajam mais em práticas preparatórias individuais no formato do exame fora da sala de aula. Esses resultados nos fazem supor que, nessa série, o efeito retroativo se torna mais explícito (XIE; ANDREWS, 2012) e intenso e corroboram as conclusões de Cheng et al (2011) de que a proximidade do teste apresenta maior impacto sobre as situações de ensino e aprendizagem. Todavia, na contramão do que era esperado, constatamos que o nível de significância do teste parece diminuir nessa série e os alunos atribuem níveis mais baixos ao alinhamento entre a preparação recebida e a avaliação externa. Essas suposições serão analisadas mais profundamente na fase qualitativa da pesquisa com o intuito de compreendermos melhor as razões que levam os alunos a se expressarem dessa forma.

4.5 Percepções e atitudes de professores e alunos comparadas

Os dados coletados por meio dos questionários aplicados em 2016 também possibilitaram realizar um comparativo sobre como os participantes da pesquisa – professores (professor e instrutor) e alunos, percebem a influência do exame na sala de aula e em suas atitudes. A Tabela 67 apresenta as porcentagens de respostas desses dois grupos, na qual 1 significa “nenhum”, 2 “pouco”, 3 “razoável”, 4 “bom” e 5 “ótimo”. É notável a discrepância existente na forma como os participantes avaliam o nível de conhecimento geral que possuem sobre o exame. Enquanto a grande maioria dos alunos se dividiu nos níveis “pouco” e “razoável”, a maioria dos professores selecionou “bom” (51%) e “ótimo” (31%).

Tabela 67 − Nível de conhecimento de alunos e professores sobre o EPLIS

	1	2	3	4	5
Alunos	0.02	0.41	0.41	0.15	0.00
Professor/instrutor	0.00	0.12	0.06	0.51	0.31

Em relação à frequência com que o professor menciona o exame em sala de aula, na Tabela 68, é possível observar as respostas dos dois grupos, na qual 1 representa "nunca", 2 "raramente", 3 "às vezes", 4 "frequentemente" e 5 "sempre". As respostas tanto dos professores quanto dos alunos concentraram-se no nível "frequentemente", não havendo diferenças expressivas na forma como ambos avaliaram a frequência com que o exame é mencionado em sala de aula.

Tabela 68 − Frequência com que o professor menciona o EPLIS em sala de aula

	1	2	3	4	5
Alunos		0.03	0.22	0.53	0.22
Professor/instrutor		0.06	0.19	0.50	0.25

Quando perguntados, porém, sobre a frequência com que aspectos do exame são discutidos pelos professores em sala de aula, percebemos uma pequena variação em suas respostas (Tabela 69). Enquanto a maioria dos alunos respondeu "às vezes", a maioria dos professores e instrutores concentrou-se no nível "frequentemente".

Tabela 69 − Frequência com que o professor discute aspectos do EPLIS em sala de aula

	1	2	3	4	5
Alunos	0.02	0.12	0.43	0.30	0.12
Professor/instrutor		0.12	0.31	0.38	0.19

Sobre a frequência com que o professor trabalha tarefas semelhantes às do EPLIS em sala de aula, as respostas dos dois grupos pareceram concentradas nos níveis "frequentemente" e "às vezes", como pode ser observado na Tabela 70.

Tabela 70 − Frequência com que o professor trabalha questões semelhantes às do EPLIS em sala de aula

	1	2	3	4	5
Alunos	0.02	0.15	0.36	0.38	0.09
Professor/instrutor	0.12	0.06	0.31	0.38	0.12

Esses dados revelam que, apesar de, em termos gerais, alunos e professores concordarem que o exame é mencionado, discutido e trabalhado frequentemente em sala de aula, os alunos avaliaram o nível de conhecimento geral sobre o EPLIS como "pouco" ou

"razoável". Essa contradição nos chama a atenção para o quê de fato pode estar interferindo nesse processo, o que será explorado na fase qualitativa do estudo.

Também foi possível traçar um paralelo entre as respostas de alunos e professores/instrutores em relação à frequência com que tarefas consideradas semelhantes às do exame são tratadas em sala de aula. A Tabela 71 mostra as proporções de respostas desses dois grupos a seis tipos de atividades (Q35, Q26, Q32, Q33, Q29 e Q30), relacionadas às fases I e II do EPLIS.

Tabela 71 − Frequência com que questões semelhantes às do EPLIS são trabalhadas em sala de aula

		1	2	3	4	5
Q35 responder questões sobre um áudio ou vídeo envolvendo uma conversa entre piloto e controlador	Alunos	0.12	0.10	0.17	0.28	0.33
	Professor/instrutor	0.00	0.06	0.12	0.38	0.44
Q26 responder, oralmente, perguntas colocadas pelo professor sobre aviação e controle de tráfego aéreo	Alunos	0.01	0.08	0.24	0.42	0.25
	Professor/instrutor	0.00	0.00	0.12	0.25	0.62
Q32 explicar, oralmente, uma situação-problema de tráfego aéreo	Alunos	0.07	0.17	0.26	0.36	0.14
	Professor/instrutor	0.00	0.06	0.25	0.44	0.25
Q33 fazer uma sugestão, oralmente, para um problema de tráfego aéreo	Alunos	0.11	0.19	0.34	0.24	0.13
	Professor/instrutor	0.00	0.06	0.19	0.44	0.31
Q29 descrever uma figura relacionada à aviação e ao controle de tráfego aéreo	Alunos	0.04	0.16	0.32	0.27	0.21
	Professor/instrutor	0.00	0.06	0.12	0.38	0.44
Q30 contar uma história com base em uma figura relacionada à aviação e ao controle de tráfego aéreo	Alunos	0.10	0.25	0.31	0.25	0.10
	Professor/instrutor	0.00	0.12	0.12	0.44	0.31

Em relação à compreensão oral de conversas entre piloto e controlador, as respostas dos professores e alunos concentraram-se nos níveis "sempre" e "frequentemente", com pouca variação. Quanto a responder, oralmente, perguntas sobre aviação e controle de tráfego aéreo, enquanto a grande parte dos alunos (42%) alegou ser bastante frequente essa atividade em sala de aula, 62% dos professores responderam que sempre a trabalham. No que concerne à atividade de explicar, oralmente, um problema de tráfego aéreo, as respostas de ambos os grupos se concentraram em "frequentemente", com uma proporção de respostas dos professores um pouco maior no nível "sempre". Para a atividade de propor uma solução para o problema de tráfego aéreo, a maior parte dos professores escolheu o nível "frequentemente" (44%), seguido de "sempre" (31%), enquanto 34% dos alunos responderam "às vezes" e 24% disseram ser frequente. Uma variação um pouco mais expressiva pode ser observada em relação à atividade de descrição de figura. Os professores concentraram suas respostas no nível "sempre" com 44%, e 38% no nível "frequentemente". As respostas dos alunos, por sua vez, se repartiram nos níveis "às vezes" e "frequentemente". Algo parecido ocorreu com a atividade de criação de histórias: as respostas dos alunos dividiram-se nos níveis "raramente", "às vezes" e

"frequentemente", enquanto os professores afirmaram realizar esse tipo de atividade frequentemente (44%) ou sempre (31%).

Com base nesses dados, é possível perceber que os professores avaliam mais positivamente a frequência com que trabalham atividades semelhantes às do exame do que os alunos. Essa diferença pode estar relacionada à série em que os alunos se encontravam na época da geração de dados, ou seja, algumas atividades podem ser mais trabalhadas em uma série do que em outra. Por essa razão, verifiquei a frequência das respostas dos alunos às atividades por série e os resultados encontram-se na Tabela 72, na qual 1 representa "nunca", 2 "raramente", 3 "às vezes", 4 "frequentemente" e 5 "sempre". Destaco, em cinza escuro, as duas maiores proporções de respostas por série.

Tabela 72 − Proporção de atividades realizadas por série relacionadas ao exame

		1 série	2 série	3 série	4 série
Q35	**1**	0.371	0.026	0.000	0.000
	2	0.224	0.086	0.028	0.022
	3	0.190	0.147	0.220	0.022
	4	0.164	0.284	0.376	0.370
	5	0.052	0.457	0.376	0.587

		1 série	2 série	3 série	4 série
Q26	**1**	0.017	0.009	0.000	0.000
	2	0.138	0.086	0.046	0.022
	3	0.336	0.293	0.147	0.065
	4	0.388	0.302	0.578	0.413
	5	0.121	0.310	0.229	0.500

		1 série	2 série	3 série	4 série
Q32	**1**	0.207	0.026	0.000	0.022
	2	0.293	0.112	0.138	0.065
	3	0.241	0.267	0.294	0.174
	4	0.190	0.431	0.431	0.478
	5	0.069	0.164	0.138	0.261

		1 série	2 série	3 série	4 série
Q33	**1**	0.267	0.069	0.037	0.000
	2	0.233	0.216	0.156	0.065
	3	0.319	0.302	0.413	0.326
	4	0.112	0.233	0.321	0.348
	5	0.069	0.181	0.073	0.261

		1 série	2 série	3 série	4 série
Q29	**1**	0.043	0.026	0.055	0.000
	2	0.164	0.207	0.174	0.022
	3	0.414	0.267	0.404	0.022
	4	0.241	0.302	0.294	0.239
	5	0.138	0.198	0.073	0.717

		1 série	2 série	3 série	4 série
Q30	**1**	0.198	0.078	0.064	0.000
	2	0.353	0.241	0.220	0.043
	3	0.302	0.336	0.339	0.196
	4	0.095	0.241	0.312	0.478
	5	0.052	0.103	0.064	0.283

Notamos um movimento nas respostas dos alunos dos níveis mais baixos de frequência para os níveis mais altos de uma série para a outra. O exemplo mais característico dessa flutuação ocorre na Q30 (contar uma história com base em uma figura relacionada à aviação e ao controle de tráfego aéreo). As respostas dos alunos da 1ª série concentraram-se no nível "raramente", enquanto os alunos da 2ª série optaram por responder "às vezes". Da 2ª para a 3ª séries, observamos um aumento na proporção de respostas no nível "frequentemente", de 0.241 (24%) para 0.312 (31%). A mesma situação se repete da 3ª série para a 4ª série, na qual há um aumento de respostas no nível "frequentemente" (de 31% na 3ª série para 47% na 4ª série) e no nível "sempre" (de 6% na 3ª série para 28% na 4ª série).

Outro caso interessante acontece na Q29 (Descrever uma figura relacionada à aviação e ao controle de tráfego aéreo). Apenas 7% dos alunos da 3ª série responderam "sempre" para essa atividade. Na 4ª série, essa proporção saltou para 71%, sugerindo que essa é uma atividade bastante recorrente na última série do curso. A descrição de figura foi a atividade que mais recebeu respostas no nível "sempre" dos alunos da 4ª série, seguida da compreensão oral de conversas entre piloto e controlador (Q35), com 58%.

Dessa forma, nas percepções dos alunos, a frequência de atividades relacionadas ao exame e trabalhadas pelos professores em sala de aula tende a aumentar de uma série para a outra, com uma grande concentração nos níveis mais altos de frequência na última série do curso, quando os alunos estão próximos à realização do EPLIS, o que respalda a hipótese levantada pela análise de variância Anova de que os professores intensificam a preparação para o exame de acordo com sua proximidade.

Comparei, também, as respostas dos alunos e professores quanto à adequação da preparação oferecida pela escola para a realização do exame. Na Tabela 73, podemos visualizar as proporções das respostas dos dois grupos de participantes. 94% dos professores afirmaram que as aulas de inglês na EEAR preparam bem os alunos para o EPLIS e 44% desses concordaram completamente com essa alegação. Em relação aos alunos, 42% concordaram que as aulas contribuem para o exame, desses 9% concordaram completamente. Porém, uma grande parte dos alunos, 35%, demonstrou nem concordar nem discordar. Outro ponto que requer nossa atenção para posterior averiguação é que 23% dos alunos se mostraram contrários a essa afirmação, enquanto nenhum professor discordou.

Tabela 73 − Adequação da preparação para o exame oferecida aos alunos

	1	**2**	**3**	**4**	**5**
Alunos	0.07	0.16	0.35	0.33	0.09
Professor/instrutor			0.06	0.50	0.44

Se por um lado essa situação pode ser explicada pelo pouco conhecimento sobre o EPLIS que os alunos, em geral, alegam possuir, o que os impediria de avaliar a adequação da preparação recebida, por outro lado, os alunos da 4ª série, que se atribuíram níveis mais altos de conhecimentos sobre o exame, apresentaram a menor média de resposta nesse componente. Isso nos faz supor que outro fator também está em jogo, fazendo com que os alunos, mesmo aqueles com mais conhecimentos sobre o teste, avaliem menos positivamente o alinhamento entre ensino e avaliação externa, diferentemente dos professores e instrutores.

Quanto às visões e percepções sobre o exame, questões semelhantes nos questionários do professor e do aluno também possibilitaram traçar diferenças entre esses dois grupos em relação aos seguintes assuntos: motivação dos alunos, carga de estudos, *stress* e ansiedade nos alunos, autenticidade do exame, confiabilidade dos resultados, exame como opcional, exame como requisito para aprovação, e importância da aprovação. A Tabela 74 mostra as proporções de respostas para cada um desses aspectos.

Tabela 74 − Comparativo das visões de professores/instrutores e alunos sobre o EPLIS

		1	**2**	**3**	**4**	**5**
Motivação para o estudo da língua	Alunos	0.06	0.08	0.18	0.42	0.25
	Professor/instrutor	0.00	0.06	0.12	0.50	0.31
Aumento da carga de estudos	Alunos	0.09	0.17	0.31	0.30	0.13
	Professor/instrutor	0.00	0.38	0.38	0.12	0.12
Aumento nos níveis de *stress* e ansiedade dos alunos	Alunos	0.10	0.10	0.20	0.29	0.31
	Professor/instrutor	0.00	0.12	0.31	0.38	0.19
Autenticidade do exame	Alunos	0.07	0.16	0.45	0.24	0.07
	Professor/instrutor	0.00	0.12	0.38	0.31	0.19
Confiabilidade dos resultados	Alunos	0.07	0.23	0.33	0.26	0.11
	Professor/instrutor	0.00	0.25	0.25	0.31	0.19
Exame como opcional	Alunos	0.16	0.27	0.15	0.21	0.21
	Professor/instrutor	0.25	0.56	0.00	0.19	0.00
Exame como requisito para a aprovação no curso	Alunos	0.48	0.19	0.12	0.12	0.10
	Professor/instrutor	0.19	0.50	0.12	0.12	0.06
Importância da aprovação no exame	Alunos	0.09	0.17	0.31	0.30	0.13
	Professor/instrutor	0.00	0.00	0.25	0.50	0.25

Em termos de motivação, a maioria dos professores/instrutores e dos alunos concordaram que o EPLIS é um instrumento de motivação para o estudo do inglês. No entanto, uma fatia maior de alunos (14%) discordou dessa afirmação. Quanto ao aumento na carga de estudos, ambos apresentaram um número expressivo de respostas no nível "nem concordo nem discordo". Porém, enquanto 43% dos alunos concordaram que o exame aumenta sua carga de estudos, 38% dos professores afirmaram o contrário. No caso do *stress* e ansiedade causados pelo exame, as respostas dos alunos e dos professores foram bastante aproximadas, concentrando-se nos níveis de concordância, com uma parcela maior de alunos concordando completamente com essa alegação.

A maior parte dos alunos (45%) buscou não opinar sobre a autenticidade do EPLIS, o que pode estar relacionado novamente à falta de conhecimentos sobre o exame ou até mesmo sobre a profissão. Já os professores concentraram-se suas respostas nos níveis de concordância, embora uma parcela ainda expressiva (31%) tenha optado por não se posicionar, provavelmente, pelas mesmas razões atribuídas aos alunos. A respeito da confiabilidade, os alunos distribuíram suas respostas de forma equilibrada em todos os níveis: uma margem de 30% discordou que o exame avalia precisamente a proficiência, 33% disseram não concordar

nem discordar, e 37% concordaram. Apesar de os professores/instrutores terem também atribuído respostas aos níveis de discordância e neutralidade, 50% deles afirmaram que os resultados do exame são consistentes.

Sobre o interesse dos participantes em o exame se tornar opcional, enquanto a imensa maioria dos professores (81%) discordou, 41% dos alunos afirmaram que gostariam que houvesse a possibilidade de optarem pela realização ou não do exame na escola. Quanto a incluir o exame como pré-requisito para a aprovação no curso, professores e alunos se comportaram da mesma forma, com a grande maioria se mostrando contrária. Esses resultados parecem sugerir que os professores e instrutores externam visões um pouco mais positivas sobre o exame, o que provavelmente está relacionado ao fato desses profissionais serem menos afetados pelos resultados do teste do que os alunos, como iremos discutir no próximo capítulo.

4.6 Resumo do capítulo

Dediquei este capítulo ao estudo quantitativo do efeito retroativo do EPLIS nas percepções e atitudes de professores, instrutores e alunos. Analisei a influência do exame em diversos aspectos: no conteúdo, nos materiais didáticos, nas atividades, nos métodos de ensino, nas avaliações de rendimento, no nível de conhecimentos sobre o exame e nas percepções dos participantes em relação à significância do teste, seu impacto negativo e a adequação da preparação. Reportei, primeiramente, os resultados da análise dos dados de base referentes aos professores que revelaram que o efeito retroativo era ainda bastante limitado, antes da implementação do teste no contexto escolar. Dois fatores foram considerados intermediadores da intensidade do efeito: o nível de conhecimento geral sobre o EPLIS e a experiência no ensino de inglês aeronáutico dos professores. Os resultados da análise estatística descritiva, referentes às respostas dos professores e instrutores ao questionário de 2016, foram mostrados em seguida e descreveram um cenário bastante distinto do anterior, com aumento expressivo na intensidade do efeito retroativo, causado, provavelmente, pela implementação do exame no contexto escolar e pelas melhorias em sua divulgação. No entanto, diferenças relacionadas à intensidade do efeito ainda foram encontradas nas comparações entre professores e instrutores e entre profissionais mais e menos experientes no ensino de inglês aeronáutico.

Em relação aos alunos, o número alto de respostas ao questionário possibilitou o emprego da análise fatorial dos componentes principais visando à redução dos dados e de testes de hipóteses. Por meio da análise dos componentes principais, reduzi 55 variáveis originais a 10 componentes principais que resumiram a maior parte de variação presente nos dados do

questionário. Relacionei todos esses componentes a três diferentes fatores: o nível de proficiência declarado pelos alunos, o nível de conhecimento geral sobre o EPLIS e a proximidade do exame. O teste de hipótese *t* de *student* para amostras independentes e a Anova revelaram diversas relações de significância estatística entre os componentes e os fatores analisados. Também foram calculados os tamanhos de efeito para essas associações, com intuito de complementar os resultados e atestar a eles também significância prática. Padrões e tendências foram constatados, hipóteses essas que necessitam ser explicadas com mais profundidade e corroboradas no estudo qualitativo. Por fim, tracei um paralelo entre professores e alunos. Os resultados revelaram que o efeito retroativo do EPLIS, na perspectiva dos professores e instrutores, é um pouco mais intenso do que na perspectiva dos alunos. Além disso, os professores e instrutores parecem externar visões mais positivas sobre o exame.

CAPÍTULO V

AS PERCEPÇÕES, ATITUDES E AÇÕES DE PROFESSORES, INSTRUTORES E ALUNOS: UM ESTUDO QUALITATIVO

5.1 Considerações iniciais

No Capítulo IV, apresentei e discuti os resultados da análise quantitativa dos dados coletados nos questionários sobre o impacto do EPLIS nas percepções e atitudes dos participantes. Neste capítulo, examino os dados qualitativos com o objetivo de explicar, expandir, refinar e corroborar as alegações e suposições feitas anteriormente, atendendo, assim, aos propósitos de complementaridade e de triangulação da pesquisa de métodos mistos para este estudo. Para isso, analisei os comentários dos professores, instrutores e alunos ao preencherem o questionário, dos alunos no grupo focal, e dos professores, instrutor, coordenação e chefia nas entrevistas individuais. Esses relatos foram organizados e interpretados por meio da análise de conteúdo, proposta por Bardin (1977), e confrontados com os dados coletados nas observações de aulas, a fim de se obter uma visão mais completa sobre como o EPLIS atua no cenário escolar. Discuto, primeiramente, as percepções e atitudes dos alunos, professores, instrutor, coordenação e chefia, traçando um contraponto com os resultados estatísticos. Em seguida, detalho os impactos do exame nas ações em sala de aula e os comparo com as percepções e atitudes dos participantes. Discorro, também, sobre a extensão com que os efeitos pretendidos são atendidos, a presença ou não de efeitos indesejáveis e não pretendidos, assim como os tipos de preparação oferecidos.

5.2 As percepções e as atitudes dos participantes

Como visto no Capítulo IV, as técnicas estatísticas empregadas possibilitaram a identificação de tendências e padrões nas percepções e atitudes dos alunos acerca da influência do EPLIS, em função do nível de proficiência por eles declarado, do nível de conhecimento geral sobre o exame e de sua proximidade. Em relação às percepções dos professores e instrutores, a análise estatística descritiva demonstrou um possível aumento na intensidade do efeito retroativo após a implementação do exame no contexto escolar. No entanto, essa intensidade ainda pareceu variar de professor para instrutor e de profissionais mais experientes

para profissionais menos experientes. Além disso, óbices em relação ao nível de conhecimentos sobre o EPLIS e ao acesso, oferta e disponibilidade de recursos ainda foram constatados nas respostas dos participantes, o que pode afetar negativamente os efeitos pretendidos. Esses resultados iniciais apontaram caminhos a serem investigados em maior profundidade sobre como o EPLIS influencia os processos dos quais alunos, professores e instrutores participam. A análise qualitativa das percepções e atitudes desses participantes está organizada em torno dos seguintes temas: *Conhecimentos sobre o EPLIS*, *Significância do exame*, *As situações de ensino e aprendizagem*, *As avaliações de rendimento na EEAR*, e a *Preparação para o exame e sua natureza*.

5.2.1 Conhecimentos sobre o EPLIS

O estudo quantitativo demonstrou problemas envolvendo o nível de conhecimentos dos alunos sobre o EPLIS, principalmente, em relação às suas características específicas, o que pode ocasionar prejuízos para a intensidade do efeito retroativo. Nas respostas ao questionário, quando indagados sobre o quanto conheciam o exame, os alunos se dividiram entre os níveis "razoável" e "pouco". Em relação aos conhecimentos que possuíam sobre o seu formato, os tipos de tarefas e o conteúdo, o cenário se deteriorou um pouco mais, uma vez que a maioria das respostas dos alunos se concentraram no nível "pouco", embora tenham considerado bastante frequente a menção ao exame em sala de aula feita pelos professores e instrutores.

Nos comentários ao questionário, os alunos justificaram que o exame é frequentemente mencionado em sala de aula, porém de maneira superficial, sem oferecer pormenores. No trecho abaixo, uma aluna da 3ª série ressalta que não houve uma apresentação oficial do exame aos alunos, na qual fossem abordadas suas especificidades, o significado de seus resultados, bem como os impactos deles decorrentes.

> Apesar de no momento atual do curso, 3ª série, o exame ser frequentemente comentado em sala de aula, o mesmo não é passado para os alunos de maneira oficial e clara. Sentimos falta de um momento direcionado especialmente para tratar como o exame acontece, especificar as fases, como elas ocorrem e de que maneira os resultados são obtidos. Além disso, que os alunos sejam informados sobre o impacto real que o EPLIS terá na vida como controlador de tráfego aéreo dentro da unidade (Questionário, comentário de uma aluna da 3ª série BCT, turma A).

Essa fala parece remeter a duas características vistas como essenciais na construção da significância do exame pelos participantes, *status* e relevância (STONEMAN, 2006). Realizar uma apresentação oficial sobre o teste, com o objetivo de fornecer esclarecimentos

sobre a aceitação, o uso e os impactos de seus resultados, além de garantir transparência ao processo, é uma forma de atestar o *status* e a relevância atribuído ao exame pelas autoridades competentes. A ausência de algo nesse sentido, a meu ver, pode afetar negativamente a credibilidade do teste entre os participantes, principalmente, para aqueles que estão próximos de sua realização, como é o caso dessa aluna, podendo resultar em uma diminuição de sua influência.

Em outro comentário, uma aluna da 2ª série reforça que, embora o teste seja bastante mencionado em sala de aula, as informações são precárias e limitadas a alguns tipos de tarefas, aos níveis de proficiência avaliados e à periodicidade de sua realização. O ponto mais importante de sua fala está na sua percepção sobre a falta de consciência dos alunos acerca da importância do inglês no controle de tráfego aéreo, o que contradiz o esperado em termos de efeito.

> Em geral, eles (professores e instrutores) vivem falando a palavra Eplis, mas não explicam com eficácia como é ou o que é. Os instrutores dão pista, como: vai ter figuras para descrever e a gente vai conversar com a avaliador sem vê-lo. Apenas sabemos que isso vai influenciar em uma nota que vai determinar de quanto em quanto tempo faremos outros testes em inglês... acho que outro Eplis. Não sei. Eles dizem que quem atinge o nível 5, faz o teste novamente em 4 anos (eu acho... por aí); nível 6 não faz mais esse teste. […]. Sinceramente acho que ainda não temos a consciência do quão importante essa língua é para nossa profissão (Questionário, comentário de uma aluna da 2ª série BCT, turma C).

A análise quantitativa também mostrou que, estatisticamente, os conhecimentos que os alunos possuem sobre o exame variam de acordo com a sua proximidade. A 1ª série alegou conhecer pouco o EPLIS, na comparação com as demais, enquanto os alunos da 4ª série apresentaram níveis mais elevados, inclusive em relação a informações específicas. Além disso, foi também pontuado que na 4ª série o exame é mais mencionado e discutido e tarefas semelhantes são trabalhadas com maior frequência. No trecho abaixo, um aluno da 2ª série, da turma A, descreve que seu primeiro contato com o EPLIS ocorreu logo após sua entrada no curso. Ele explicita que os professores da 1ª série do curso chegaram a abordar o exame em sala de aula, porém não deixa claro o conteúdo que foi discutido.

> eu descobri sobre o Eplis assim que eu cheguei aqui (+) o único teste de proficiência (incompreensível) que eu conhecia era o Toefl (+) e aqui falaram que tem dois outros, né (+) inclusive o Eplis (+) não fazia ideia para quê que servia (+) nem conhecia (+) nunca tinha ouvido falar (+) assim que eu cheguei aqui o pessoal falou pra gente bastante sobre o Eplis principalmente as professoras de inglês da 1ª série (Grupo focal, comentário de um aluno da 2ª série BCT, turma A).

Outro aluno da 1ª série argumenta que ficou sabendo sobre o EPLIS por acaso, em uma conversa informal com uma instrutora de inglês durante a aplicação de um teste oral. Ele reforça ainda que nada a respeito havia sido dito na sua turma de inglês e afirma que a maioria dos seus colegas desconhece a existência do EPLIS. Assim, relata:

> mas isso não foi pra todo mundo (+) porque pra mim não passaram (não mencionaram o EPLIS) (+) eu descobri que tinha essa prova porque (+) a gente fez um teste oral (+) acho que em todas as séries tem, né (+) a gente fez um teste oral com a sargento (+) aí eu tava conversando com ela e aí ela falou pra mim que existia uma prova (+) no final da 4ª série (+) fiquei sabendo tudo o que era EPLIS (+) mas SÓ eu (+) porque eu conversei com a sargento e ela me contou (+) mas ela não contou pra sala (+) contou só pra mim naquela prova oral por isso que eu sabia (+) mas ninguém passou na minha sala (+) ninguém falou nada (+) acho que se perguntar na 1ª série (+) pelo menos pra turma fox (+) se alguém sabe (+) ninguém levanta a mão (+) ninguém sabe (Grupo focal, comentário de um aluno da 1ª série, turma F).

Um aluno da 4ª série diz, durante o grupo focal, se recordar de o EPLIS ter sido citado na 1ª série do curso como sendo um exame de proficiência a ser realizado no último semestre, no entanto, ele afirma que não foram abordados detalhes sobre o teste: "na 1ª (série) nem sabia o que era (EPLIS) (+) simplesmente falaram que tinha uma prova no final do curso que (+) que seria avaliado o nosso nível de inglês e nada mais".

Um quarto comentário respalda a hipótese estatística de que os alunos da 4ª série possuem maior nível de conhecimentos sobre o exame e suas características, em comparação às demais séries. A fala dessa aluna mostra indícios de que os professores e instrutores que lecionam nessa série explicam e discutem aspectos do exame com os alunos. Outro ponto interessante é a afirmação de que quase tudo que é feito na sala de aula da 4ª série apresenta relação com o exame, indicando, provavelmente, uma preparação em mesmo formato.

> mas agora orientação de realmente como' seria o teste eh (+) como o teste seria abordado eh (+) coisas desse aspecto foram BEM mais na 4ª série (+) 4ª série praticamente int / 4ª série inteira' do inglês foi voltada pro Eplis como sendo o Eplis (Grupo focal, comentário de uma aluna da 4ª série BCT, turma B).

Com base nesses relatos, podemos assumir que, primeiramente, não existe padronização do tratamento dado ao EPLIS nas séries iniciais. O exame parece ser mencionado ainda no início do curso, mas suas características tendem a ser exploradas em sala de aula somente na 4ª série, sugerindo que o efeito retroativo é mais explícito na última série do curso. Na entrevista com a coordenadora e professora Pilar, ela explica que, de fato, o exame é apenas citado no momento de apresentação do curso como um evento futuro, com o objetivo aparente de não causar pressão nos alunos.

> a gente fala pra eles a respeito do exame quando vai apresentar o curso' (+) olha gente, na 4ª série vocês vão ter o exame que vai rodar na escola (+) mas assim, nada que Ó vocês VÃO ter que estudar pro exame (+) não, aquilo que vai acontecer (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

Todavia, a meu ver, a decisão de adiar para a 4ª série informações mais detalhadas sobre o exame pode diminuir sua relevância, ao invés de elevá-la, dificultando, assim, a ocorrência dos efeitos pretendidos. Os demais professores e instrutor também confirmam a percepção dos alunos quanto à carência de informações sobre o exame no início do curso. O instrutor João também afirma que os alunos das primeiras séries do curso são informados apenas sobre a necessidade de realização do EPLIS e que, a partir da 3ª série, os alunos começam a procurá-lo para adquirirem mais informações sobre o teste. Assim, relata:

> no final da 3ª série ele sabe'(+) ele JÁ sabia desde a 1ª série que ele teria que fazer o EPLIS na 4ª série (+) mas quando ele chega na 3ª série ele começa a ficar com aquela dúvida'(+) e se eu não conseguir o nível quatro" /.../ então, da 3ª para a 4ª série quando ele começa a sentir que ele vai se formar (+) na escola (+) e chega o EPLIS na 4ª série (+) ele começa a perguntar (+) então, eu noto isso MUITO nos alunos da 2ª série (+) não, da 3ª série (+) desculpa (+) a 3ª série pergunta MUITO sobre o EPLIS (+) MUITO (+) principalmente para nós controladores que já fizemos (Entrevista com João, instrutor na EEAR desde 2016)

Essa fala de João, além de sustentar a hipótese levantada no estudo quantitativo de que os alunos preferem sanar dúvidas sobre o exame com os instrutores de inglês aeronáutico por esses já terem realizado a prova como candidatos, demonstra que a iniciativa pela busca por conhecimentos específicos tende a partir dos alunos e, não, dos professores. Dessa forma, ainda que a intenção dos professores de não causar *stress* seja válida, os próprios alunos querem mais esclarecimentos sobre o teste e o quanto antes isso ocorrer mais intenso pode ser o seu efeito.

O fato de os alunos da 4ª série possuírem mais conhecimentos específicos sobre o EPLIS, tais como informações sobre sua estrutura, tipos de tarefas e conteúdo abordado, em comparação às demais séries, pode estar relacionado à percepção de que a 4ª série, diferentemente das anteriores, tem por objetivo a preparação dos alunos para o exame. Em um dos relatos anteriores, um aluno já havia comentado a respeito, a coordenadora e professora Pilar, então, confirma:

> a 4ª série é a única série que (+) digamos assim, (+) está VOLTADA para a preparação pro exame (+) então, a gente faz um apanhado (+) um *review* de todas as temáticas (+) de todos os temas abordados na 2ª e 3ª série (+) e a gente dá essa pincelada do EPLIS (+) a gente treina para o Eplis veementemente (+) é a única série que está voltada exclusivamente para o exame (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

A preparação para um exame não constitui uma prática ruim se ela se propuser a aumentar os resultados obtidos por meio do desenvolvimento das capacidades e habilidades que o teste se dispõe a avaliar (MESSICK, 1982). Contudo, existem intervenções consideradas indesejáveis, pois inflam as notas sem uma contrapartida em termos de proficiência ou estreitam demasiadamente os métodos de instrução e o currículo em sala de aula. Nesse sentido, a afirmação feita por Pilar de que na 4ª série os professores e instrutores "treinam veementemente para o EPLIS" deve ser tomada com cautela. Nas observações em sala de aula, investiguei melhor a natureza desse treinamento e as diferenças para com as aulas ministradas nas demais séries.

Outra hipótese suscitada no estudo quantitativo de que quanto mais alto o nível de proficiência dos alunos, mais conhecimentos eles possuem sobre o exame encontra respaldo no comentário de uma aluna da 2ª série do curso, da turma Eco, ao questionário. A aluna relata que a maior parte dos professores tende a desenvolver melhor o conteúdo relacionado ao EPLIS nas turmas consideradas mais proficientes, o que provavelmente pode ter contribuído para níveis mais elevados de conhecimentos sobre o exame e de sua importância nesses grupos. Ela também pontua que os alunos com baixa proficiência não são motivados pelos professores no que se refere ao teste. Cabe lembrar que, a partir da 2ª série do curso, os alunos são distribuídos em turmas de acordo com o nível de proficiência na língua. É utilizado o alfabeto fonético da aviação para identificá-las. A turma Alfa (A) é considerada a mais proficiente, a turma Bravo (B) vem em seguida e assim por diante. Essa aluna pertence à turma Eco (E). Assim, comenta:

> Um outro fator, também preocupante que ocorre no galpão (local onde as aulas de inglês acontecem), é que a maioria dos professores trabalham e desenvolvem melhor os conteúdos relacionados ao EPLIS, com as turmas mais avançadas. As classes que possuem o nível de inglês inferior não são estimuladas a melhorar o interesse pelo inglês (Questionário, comentário de uma aluna da 2ª série BCT, turma E).

Como discutido anteriormente, ao considerar que a percepção do nível de dificuldade do teste pode interferir na intensidade do efeito retroativo em sala de aula (GREEN, 2007a), podemos presumir que os professores e instrutores são mais propensos a discutir o exame e engajar os alunos em práticas preparatórias com maior frequência nas turmas mais avançadas linguisticamente, por acreditarem que existam chances reais de aprovação. Da mesma maneira, os alunos mais proficientes, ao perceberem a possibilidade de serem bem-sucedidos no exame, podem demandar de seus professores e instrutores mais informações a seu respeito. Por outro lado, nas turmas menos proficientes, ao assumir que o nível do exame está muito além do nível de proficiência dos alunos, os professores e instrutores podem optar por

não o enfatizar, a fim talvez de evitar sentimentos de frustração. Esses alunos, por sua vez, podem se conformar com o fato de que o exame está acima de sua capacidade atual e acabam perdendo qualquer interesse.

Nas entrevistas com os professores e instrutor, perguntei se eles percebiam relações entre a abordagem de ensinar e o nível de proficiência dos alunos. A professora Ana e o instrutor João confirmaram que, nas turmas mais proficientes, é possível atribuir mais atividades de produção oral livre, inclusive atividades extras, dado que os alunos são capazes de se expressar com mais facilidade. Nas turmas menos proficientes, segundo Ana, é necessário ter mais controle, o que faz com que o professor acabe por dominar o tempo de fala. No trecho abaixo, Ana discorre sobre as diversas possibilidades de desenvolvimento da produção oral em turmas com maior proficiência. Além da atividade de descrever, ela menciona outras tarefas para a promoção da interação, como emitir opiniões e defender um ponto de vista. Essas propostas são bastante semelhantes àquelas encontradas na parte 3 da fase II do exame, denotando provável efeito retroativo.

> nas turmas avançadas eu acho que a gente / DÁ tempo de fazer outras atividades (+) mas também dá tempo de praticar mais o *speaking* (+) assim' (+) até mesmo com o próprio material (+) você explora mais a figura (+) a descrição de figura (+) você explora mais a perguntinha que ali tá (+) sabe, você vai fazer uma pergunta em cima de pergunta (+) tem mais discussão (+) eles expressam melhor as opiniões (+) o ponto de vista deles (+) a gente pode até dividir a sala (+) fica um grupo defendendo um ponto de vista (+) o outro grupo defende o outro ponto de vista (+) e promove debate em sala de aula (+) então, na turma mais fluente isso é possível (Entrevista com Ana, professora na EEAR desde 2013).

Fica evidente que a obtenção de informações claras e precisas sobre o exame intensifica o seu potencial para exercer os efeitos pretendidos. Mas, a quem caberia a tarefa de prover esclarecimentos sobre o teste, conscientizando, assim, os alunos sobre sua importância? Os comentários dos alunos parecem conferir aos professores e instrutores essa incumbência. Sem dúvida, eles são os principais agentes nos processos de efeito retroativo, por fazer a ponte entre o ensino e a avaliação (BAILEY, 1996; BURROWS, 2001, 2004). Burrows (2001, p. 171) define o professor como o "meio pelo qual qualquer mudança causada pela implementação de um exame deve passar",[64] e estudos como os de Bartholomeu (2002) e Correia (2003) confirmam a importância do seu papel nesse processo. Contudo, questiono se as condições necessárias têm sido dadas a esses participantes. Fulcher (2012) nos lembra do poder investido nos testes, principalmente, os externos, em manipular o comportamento de professores e

[64] No original: "Teacher as a medium through which any change brought about by the implementation of assessment must pass".

colocá-los como os responsáveis pelo alcance das metas estabelecidas nas políticas linguísticas educacionais. Para ele,

> os elaboradores de política frequentemente acreditam que mudanças podem ser implementadas de forma relativamente rápida e barata sem ter que empreender desenvolvimento curricular ou mudança nas práticas em sala de aula por meio de programas educacionais para professor[65] (FULCHER, 2012, p. 114).

Diversas são, portanto, as situações associadas aos professores que podem dificultar a ocorrência dos efeitos esperados como, por exemplo, a falta de conhecimentos sobre o exame, a falta de domínio sobre o conteúdo cobrado e a falta das habilidades pedagógicas necessárias, além de questões ligadas à resistência e à disposição para a inovação (ALDERSON; HAMP-LYONS, 1996; BURROWS, 2001; GREEN, 2007a).

Sobre isso, nas discussões levantadas no grupo focal, os alunos ressaltaram que o modo como o exame é abordado em sala de aula varia de professor para professor, respaldando a hipótese de Alderson e Wall (1993) de que os professores não se comportam da mesma forma em face da influência de um exame, e atribuem essa diferença aos níveis de conhecimentos desses participantes sobre o EPLIS e sobre a comunicação aeronáutica. Na visão desses alunos, os instrutores de inglês aeronáutico, por serem controladores de tráfego aéreo, conhecem melhor o exame e, por isso, preparam os alunos mais adequadamente, quando comparados aos professores com graduação em Letras. No trecho abaixo, transcrito do grupo focal, um aluno da 4ª série pontua que os instrutores de inglês aeronáutico trouxeram mais informações sobre as características do EPLIS, associando a vinda deles a um maior esclarecimento sobre o teste.

> a gente foi entender o que realmente era o Eplis quando chegaram os sargentos de fora (controladores de tráfego aéreo) (+) no caso da verde na 2ª série (+) na 2ª série vieram três sargentos e (+) REALMENTE explicaram ali como que acontecia (+) ih: o que que era o Eplis (+) foi realmente a introdução do Eplis pra mim (+) na minha turma ali foi (+) a vinda dos sargentos que (+) são Bct' (controladores de tráfego aéreo) e falaram' o Eplis é isso (+) eu fiz o Eplis e é assim que acontece' eh: 'essas são as etapas'(+) depois na 3ª série foi também bastante lembrado (+) pelos sargentos como que funciona (+) quais são as etapas (+) como funciona a primeira fase (+) como que funciona a 2ª fase (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Em relação aos professores formados em Letras, os alunos os subdividem em dois grupos. De um lado, estão aqueles que receberam algum tipo de capacitação em inglês aeronáutico ou que, por meios próprios, buscaram conhecimentos sobre a área. Eles se igualam

[65] No original: "Policymakers frequently believe that changes can be implemented relatively quickly and cheaply without having to undertake curriculum development or change classroom practices through teacher education programs".

ao grupo de controladores de tráfego aéreo em termos de conhecimentos sobre o exame e sobre inglês aeronáutico. O restante dos professores, na visão dos alunos, apresenta conhecimentos precários e, por essa razão, costumam tratar o EPLIS em sala de aula como um exame bastante difícil, cuja aprovação é uma meta quase inatingível, o que faz com que os alunos desenvolvam sentimentos de medo e intimidação em relação ao teste, como podemos perceber no relato abaixo de um aluno da 4ª série.

> a forma que eles (os professores da EEAR) falavam do Eplis (+) eu tinha uma visão que era um negócio impossível de chegar no nível 4 (+) eu não vou chegar no nível 4 nunca (+) e a forma que eles falam que tem que ser MUITO (+) bom (+) […] eu vi ali a forma que era abordado o Eplis (+) o medo que eles colocavam em nós (+) eu achava que era impossível chegar no nível 4 (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Outro aluno da mesma série compara o professor com baixos níveis de conhecimentos sobre o teste e o instrutor que, por ser controlador de tráfego aéreo e ter realizado o EPLIS, o conhece bem. No seu comentário, que respalda o anterior, o instrutor é visto como alguém capaz de desmitificar o exame, tranquilizar os alunos e mostrar caminhos para uma preparação adequada. Por outro lado, os professores com poucos conhecimentos criam um ambiente hostil. Essa situação, a meu ver, pode não apenas comprometer os impactos pretendidos, uma vez que faz com que os alunos percam o interesse pelo EPLIS, mas, também, desenvolver nos alunos certa aversão ao teste.

> os instrutores […] que não tem essa formação específica da aviação (+) eles às vezes acabam criando um bicho de sete cabeças (+) e aí (+) mhm: não deixam o aluno a par do que é o Eplis (+) e acaba deixando os alunos com medo (+) preocupados (+) e sem saber o que que é (+) e aí às vezes vem um instrutor que é tipo' (+) sargento controlador que já sabe (+) já passou por isso (+) e ele diz assim' não, (+) é tranquilo (+) faz assim' (+) dá o caminho certo pra estudar (Grupo focal, comentário de um aluno da 4ª série BCT, turma C).

As percepções dos alunos quanto a níveis distintos de conhecimentos sobre o EPLIS entre professores e instrutores coincidem com os resultados alcançados no estudo quantitativo. As respostas dos professores e instrutores ao questionário revelaram diferenças entre esses dois grupos, tendo os instrutores atribuído níveis mais altos a todos os aspectos concernentes ao exame (objetivo, conteúdo, tarefas, tabela de níveis, desempenho necessário, e Doc 9835 AN/453). Dentro do grupo de professores, embora tenha havido um aumento expressivo entre 2014 e 2016, com a maioria atribuindo níveis "bom" ou "ótimo", foram observadas respostas nos níveis "pouco" e "razoável", indicando que ainda existem arestas a serem preenchidas. Esse

aumento na percepção dos professores em relação ao nível de conhecimentos que possuem sobre o EPLIS pode ser decorrente de uma oficina oferecida pelo ICEA em parceria com a EEAR, em 2015, aberta a todos os professores de inglês, bem como das melhorias implementadas na divulgação do teste. No entanto, a oficina teve uma duração bastante curta, dois dias apenas, sendo que nem todos os professores participaram integralmente das atividades, indicando que podem haver, sim, deficiências nesse sentido.

Pilar, coordenadora pedagógica e professora, ao ser questionada sobre o papel da coordenação na orientação dos professores e instrutores a respeito do EPLIS, explica que, apesar do nome, a coordenação não exerce uma função pedagógica e, sim, administrativa. Em suas palavras:

> esse papel do coordenador na escola de especialista é mais aquele que gerencia a distribuição, a atribuição de aulas eh:: (+) gerencia substituições (+) algumas mudanças de material didático (+) enfim, gerenciamento da parte eh: de como as coisas vão acontecer [a logística] sim, a logística (+) e não veementemente a sala de aula (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

Sendo assim, Pilar diz que apenas sugere aos professores e instrutores, principalmente àqueles que irão atuar na 4ª série do curso, que leiam o manual do candidato e visitem o *site* do exame para obter informações mais precisas e acrescenta que costuma compartilhar todo o material e atividades que elabora, porém como sugestão.

> agora quanto ao Eplis (+) o que eu procuro também falar PRINCIPALMENTE para a equipe que tá trabalhando com a 4ª série (+) que eu tenho uma cópia aqui no galpão (manual do EPLIS) e eu sempre falo no meu briefing que todos os professores que trabalham com a 4ª série PRECISAM conhecer o exame /.../ então assim' eu peço para que os professores conheçam a apostila (manual do candidato) (+) entrem no *site*.
>
> eu procuro oferecer tudo aquilo que eu produzo (material didático) pros meus professores (+) então, eu compartilho todos os meus arquivos pra todo mundo (+) eh: compartilho todas as atividades que faço com meus alunos durante as aulas (+) mas como uma sugestão (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

Podemos dizer que o que é feito em termos de orientação pedagógica parece ser muito pouco frente às pretensões da política linguística da OACI de inovação do ensino e aprendizagem de inglês no contexto aeronáutico. A ausência de um acompanhamento e direcionamento mais efetivo, além de poder diminuir, de modo geral, a intensidade do impacto, possibilita uma maior variação entre os professores, uma vez que cada um terá sua própria interpretação do teste, seu propósito e seu impacto em sala de aula.

Na entrevista, a professora Ana confirma que grande parte do seu conhecimento sobre o EPLIS proveio da leitura do manual do candidato e do suporte dado pela coordenação e por colegas mais experientes. No final de sua fala, Ana destaca que conhece bem o exame, porém deixa claro que esse conhecimento não é fruto de treinamento específico, mas, sim, de situações mais informais.

> ai, […], é assim'+) eu conheço o Eplis lendo aqui o manual que a gente tem no no galpão (+) né, (+) e conversando com as meninas que trabalhavam já com a 4ª série /.../ assim que a gente chegou na 4ª série ela (coordenação) sentou e teve o cuidado de explicar (+) como que é (+) como funciona o Eplis (+) então assim, isso foi legal (+) eu acho que eu conheci BEM (+) assim'(+) mediante explicações de pessoas que JÁ trabalhavam há tempo com isso (Entrevista com Ana, professora na EEAR desde 2013).

Assim como Ana, o instrutor João, controlador de tráfego aéreo e instrutor de inglês aeronáutico, afirma que vem construindo conhecimentos sobre o EPLIS por meio do manual do candidato, das informações vinculadas no *site* do exame e da troca de experiências com colegas. João também menciona, positivamente, as mudanças ocorridas no processo de divulgação.

> o meu conhecimento do Eplis ele tem sido aprimorado (+) eu comecei a me atentar mais a:ao manual do candidato (+) que tá no *site* (+) e também o ICEA por ter disponibilizado mais recursos recentemente (+) colocando mais *quiz* da parte da primeira fase (+) o manual foi desmembrado (+) tem mais itens (+) tá mais explicado (+) então, eu comecei a aumentar (+) e também começou a aumentar meu conhecimento pela troca de experiência (Entrevista com João, instrutor na EEAR desde 2016).

A chefe da Subseção de Idiomas e professora de inglês Isabel, com 6 anos de experiência na EEAR, ao ser perguntada sobre orientações em relação ao EPLIS para os professores, relata que se reuniu com a última turma de professores, recebida em 2016, para explicar o funcionamento do exame. Além disso, afirma tê-los incluído como assistentes na aplicação da fase I para que pudessem, dessa maneira, conhecê-lo melhor. Na sua fala, ela também cita que um dos professores recém-chegados foi avaliador do EPLIS e sua experiência contribuiu para que a equipe tivesse contato com casos reais de candidatos.

> especificamente com essa última turma que a gente recebeu com novos professores (+) a gente fez uma reunião (+) eu expliquei como é o Eplis (+) como que funciona (+) inclusive no último agora (+) eu coloquei MAIS oficiais para participarem da aplicação do Eplis (1ª fase do Eplis) (+) porque é bem né (+) é uma logística bem grande por causa da quantidade de alunos (+) então, eles conhecem (+) a gente discute (+) tem alguns que já eram avaliadores como o […](+) ele pôde falar melhor ali de situações REAIS que aconteceram (+) que pra gente fica muito mais fácil de mensurar

> ali (+) esse aqui é nível 4 (+) isso aqui é o 5 (+) por causa da prática já do dia a dia (+) alguns ainda estão pegando (+) por exemplo, teve uma […] que deu tudo cinco cinco cinco cinco (+) e eu falei' qual turma você pegou" (+) ah turma tal (+) essa turma não é cinco (+) eles são bem criteriosos (+) eles avaliam muito a estrutura (+) a fluência (+) pra tirar cinco é bem difícil na verdade (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

Outro ponto ressaltado por Isabel em sua fala está relacionado à atribuição de níveis de proficiência aos alunos pelos professores, com base na tabela da OACI. Ela afirma que a prática torna a distinção entre os níveis mais fácil. Sem o propósito de desmerecer os conhecimentos construídos por Isabel e Pilar, o fato é que elas não tiveram a oportunidade de participar de cursos mais específicos sobre o exame que as habilitassem a capacitar outros profissionais. Essa situação pode acarretar inconsistências no entendimento das demandas do teste como, por exemplo, a percepção de que a correção gramatical é enfatizada, presente na fala de Isabel. Por outro lado, a presença desse ex-avaliador do EPLIS na equipe de professores e instrutores pode mudar esse cenário.

Vemos, portanto, nos relatos acima, que os conhecimentos sobre EPLIS se devem, majoritariamente, à busca do próprio professor ou instrutor por informações, seja lendo o manual ou visitando o *site* do exame, seja por meio de informações passadas por colegas mais experientes e pela prática do dia a dia. Ainda que as melhorias implementadas no *site* de divulgação possam ter potencializado efeitos positivos, o que foi feito é, a meu ver, minúsculo diante da alta relevância do EPLIS e, como muitos estudos têm mostrado (ALDERSON; HAMP-LYONS, 1996; CHENG, 1999; SCARAMUCCI, 1999; WATANABE, 2004b; QI, 2004; RETORTA, 2007), o exame por si só não é capaz de introduzir as mudanças esperadas em sua plenitude. Assim, a pouca ou ausência de oferta de capacitação especializada e contínua aos professores sobre o propósito do teste, seu conteúdo, os tipos de tarefa, e a grade de avaliação pode fazer com que visões distorcidas ou mitos sobre o exame sejam criados, o que levaria a efeitos indesejáveis, como reconhece a literatura.

Quanto ao nível de conhecimentos sobre inglês aeronáutico, as respostas dos professores ao questionário também apresentaram dispersão, com 34% dos professores marcando "pouco" ou "razoável", confirmando a visão dos alunos de que alguns não dominam por completo o conteúdo que ensinam. Isso gera preocupação, uma vez que, por ser um exame de proficiência com propósitos específicos, entender o EPLIS pressupõe entender a natureza da linguagem na comunicação aeronáutica. A proficiência do professor sobre o uso da língua nesse contexto é uma das condições que irá favorecer ou não a ocorrência do efeito retroativo e determinar seu tipo e intensidade.

De fato, o ensino de inglês para fins específicos (ESP), mais especificamente de inglês aeronáutico, impõe tarefas às quais o professor de inglês geral não necessariamente está habituado. Basturkmen (2017) lista uma série de atividades que fazem parte do repertório do professor de inglês para fins específicos, além, claro, de ministrar aulas: entender as necessidades dos alunos, conhecer a natureza da linguagem utilizada por eles e desenvolver cursos e materiais específicos.

Quando perguntada sobre o processo de capacitação dos professores nessa área, a coordenadora e professora Pilar explica que costumam ser oferecidas aulas sobre controle de tráfego aéreo, em português, para aqueles que irão atuar nas 2ª, 3ª e 4ª séries. Além disso, Pilar reporta que, ao longo desses anos, alguns professores tiveram a oportunidade de participar de um curso de formação de instrutores de inglês aeronáutico no ICEA. Esse curso a que Pilar se refere, ainda que colabore positivamente para que os professores tenham maior clareza sobre as características da comunicação aeronáutica e para que se atualizem em relação às metodologias mais recentes no ensino de línguas, não foi elaborado, especificamente, para esse propósito. Pelo contrário, ele assume que os participantes (na grande maioria, controladores de tráfego aéreo) têm domínio sobre o inglês aeronáutico, porém necessitam de capacitação pedagógica para atuarem como instrutores.

A pouca ou ausência de ofertas de desenvolvimento profissional, somada à falta de um trabalho mais adequado em termos de coordenação pedagógica e às deficiências reportadas por alguns professores no entendimento do exame e de seu construto, parece apontar para uma política linguística ainda fraca, o que, consequentemente, pode levar à redução dos impactos pretendidos ou até mesmo à ocorrência de efeitos negativos. Portanto, se desejarmos que a intensidade do efeito retroativo seja, de fato, extensa, é necessário, primeiramente, preencher essas lacunas, começando por oferecer a todos aqueles envolvidos nas atividades de ensino, considerados os principais agentes na condução desse processo, a formação adequada, com foco na natureza da linguagem avaliada, nas necessidades dos alunos, no que deve ser ensinado e aprendido e, principalmente, no como fazer. Somente assim, eles se sentirão mais preparados e dispostos a refletirem sobre sua prática pedagógica e a promoverem as inovações pretendidas.

5.2.2 A significância do exame

Proposto por Stoneman (2006), o conceito amplo de significância de um teste (*test significance*) reúne os conceitos de *status*, lastro, consequência e relevância, conforme salientei no Capítulo II, na seção 2.4. O *status* de um exame é definido por três características:

reconhecimento e reputação atribuídos por seus usuários, sanção oficial dada pelo governo ou por outras autoridades no assunto e escopo de implementação. O nível de relevância está intimamente ligado à percepção das consequências advindas de seus resultados. A relevância é considerada um conceito subjetivo, pois varia de um participante para outro. O *lastro*, por sua vez, integra o grau de aceitação dos resultados e uso que se faz deles.

Na análise fatorial das repostas dos alunos ao questionário, o componente interpretado como *Significância do exame* recebeu maior expressividade no construto 3 (*Visões e percepções sobre o EPLIS*) e relacionou itens ligados à motivação, autenticidade do exame, confiança nos seus resultados e importância da aprovação para o desempenho da função profissional. Os alunos da 4ª série foram os que atribuíram menor nível de concordância a esses itens, o que causou certa surpresa, uma vez que, por terem adquirido mais conhecimentos sobre o exame e estarem próximos à sua realização, esperávamos que a avaliação desses itens fosse mais positiva em relação às demais séries.

Durante o grupo focal com os alunos, surgiram diversos comentários relacionados à percepção de *status* e relevância do EPLIS. Na fala de uma aluna da 4ª série, por exemplo, a escolha pelo termo *determinante* indica que ela parece estar consciente do que o exame representa para a sua carreira. Ademais, ela confere credibilidade ao teste ao ver a aprovação como uma comprovação de que o profissional está qualificado para desempenhar sua função no que se refere à proficiência em inglês, o que lhe proporciona segurança e confiança. Assim, relata:

> algo determinante (+) é algo que vai (+) não (apenas) determinar sua vida de controlador (+) mas sim de (+) para mostrar para si mesmo (+) o quão (+) não é focado mas o quanto você se preparou (+) saber que você (+) tem uma segurança maior também (+) em relação a (+) eu vou pegar um voo' internacional (+) não quero porque eu não tenho segurança de que eu sei (+) esse teste acaba sendo para nós mesmo uma: comprovação de que eu consigo passar no EPLIS (+) então tem alguma coisa ai (+) acho que nesse sentido (Grupo focal, comentário de uma aluna da 4ª série BCT, turma D).

A coordenação, a chefia, os professores e o instrutor entrevistados também consideram que a relevância atribuída pelos alunos ao exame tem aumentado, de modo geral, no contexto escolar e foram unânimes ao dizer que a entrada do EPLIS na escola fez com que os alunos se tornassem mais conscientes da importância do inglês para o bom desempenho na futura profissão. Nas palavras da coordenadora e professora Pilar, é possível observar a menção feita ao poder exercido pelo teste de aumentar a percepção de importância da proficiência

quando diz que a implementação do exame "foi uma maneira de (+) forçá-los a enxergar que isso é realmente importante", o que evidencia, possivelmente, efeito positivo.

> os aspectos positivos eu acho que (+) a questão da consciência (+) porque eu acho que a juventude hoje (+) ela / a gente precisa (+) a gente precisa desse INPUT (+) da importância eh: das nossas / de executar bem as nossas tarefas (+) então, é algo que já vai incutindo no aluno que eh: (+) essa qualidade (+) essa proficiência que ele necessita para executar tarefas simples e também tarefas eh: que podem não ser tão simples assim, né, (+) no caso de emergência ou situações não rotineiras (+) da importância do inglês (+) então eu acho que foi uma maneira de (+) forçá-los a enxergar que isso é realmente importante (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

A chefe da Subseção de Idiomas e professora Isabel confirma que a aplicação do EPLIS na escola contribuiu para que os alunos tratassem com mais seriedade o curso oferecido: "então, eu acho que (+) pra escola (+) no caso nosso aqui (+) acho que é só deu mais aspecto positivo (+) na questão do aluno levar mais a sério o Eplis (+) levar mais a sério o curso". O professor Eduardo compartilha da mesma visão: "MAIS (conscientes) do que antes (+) principalmente depois que o EPLIS começou a acontecer aqui na escola". A professora Ana, por sua vez, atribui esse aumento também à participação dos instrutores de inglês aeronáutico nas 2ª e 3ª séries do curso que, segundo ela, por serem controladores de tráfego aéreo, costumam partilhar com os alunos suas experiências sobre o EPLIS, conscientizando-os da importância do exame cada vez mais cedo, o que é considerado fundamental para a percepção de relevância do teste. Assim, relata:

> eles ESTÃO (+) preocupados (+) eles estão sim (+) eh mesmo porque no:o na 3ª e 2ª série agora a gente tá trazendo instrutores que SÃO controladores (+) e foram alunos aqui (+) passaram por toda a experiência que eles irão passar (+) e os instrutores TÊM dado testemunhos assim pra eles' (+) como é que foi (+) olha, a importância do inglês (+) como é que foi pra mim (+) então assim, eles estão dando BASTANTE importância sim (+) ficam preocupados (Entrevista com Ana, professora na EEAR desde 2013).

O instrutor João menciona que o EPLIS faz com que os alunos se dediquem mais ao estudo do idioma: "o fato do Eplis ser na 4ª série (+) pelo menos eles se dando conta que eles têm que estudar (+) eles estão estudando alguma coisa (+) alguma coisa de inglês pra aviação eles estão estudando". Assim como João, a professora Ana destaca o papel do EPLIS como um elemento motivador nesse cenário:

> eu acho que só ajuda né, (+) eu acho que é um elemento motivador (+) pra gente conseguir CHAMAR a atenção dos alunos pra importância da aula (+) olha'(+) você tem que prestar atenção porque você tem esse teste pra fazer (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

Os professores também relatam que quanto mais próximos os alunos estão do EPLIS, mais importância eles atribuem a ele. Sobre isso, a professora Ana reporta: "eu vejo mais na 4ª série (+) eles acordam pra isso aí na 4ª série (+) o Eplis vai chegando e eles começam a se preocupar". Essa visão é endossada pela coordenadora e professora Pilar:

> na 2ª série eles não se preocupam ainda (+) quando chegam na 3ª série eles se preocupam mas também é bem pontual (+) não são todos (+) certo, (+) eh: na 4ª série eles já ficam BEM preocupados (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Tanto a fala da aluna quanto os comentários da coordenadora, da chefia, dos professores e do instrutor sustentam que o nível de relevância do exame tem aumentado entre os alunos, alcançando seu auge na 4ª série, o que contraria os resultados quantitativos apresentados no início desta seção. Os relatos desses participantes evidenciam que os alunos da 4ª série parecem estar mais cientes em relação ao teste, o veem como um instrumento confiável na avaliação da proficiência e um fator de motivação no estudo do idioma e na conscientização sobre a importância da língua no serviço de tráfego aéreo.

Entretanto, no decorrer das discussões no grupo focal, pude notar que existe certa diferença de significado entre importância e relevância atribuídas ao exame. Em outras palavras, o fato de os alunos estarem mais conscientes da importância do inglês e de serem bem-sucedidos no exame, embora demonstre efeito positivo, não significa, necessariamente, que a relevância atribuída ao exame, em termos de consequências, é alta. Pelo contrário, os alunos deixam claro que a relevância dada ao EPLIS é menor do que a relevância alcançada pelas avaliações de rendimento no contexto escolar, já que os resultados no exame não têm implicações imediatas para a vida profissional.

O nível alto de relevância atribuído às avaliações de rendimento reside no fato de os resultados alcançados nessas provas comporem a média geral do aluno no curso BCT, juntamente com as notas obtidas nas demais disciplinas. Essa média final define a classificação dos alunos, estabelecendo uma ordem de prioridade na escolha do local (cidade) onde eles irão atuar como controladores de tráfego aéreo. Ou seja, o aluno que atinge a maior média geral no curso é o primeiro a escolher a localidade e, assim, sucessivamente. As vagas são determinadas pelo DECEA e estão distribuídas por todo o território nacional. Além disso, o aluno que não obtém o rendimento mínimo em qualquer disciplina corre o risco de ser desligado do curso, inclusive nas disciplinas de inglês. No caso do EPLIS, os alunos argumentam que se eles não obtiverem a nota mínima, eles terão a oportunidade de refazê-lo no próximo ano, sem sofrerem qualquer penalidade.

No trecho abaixo, um aluno da 4ª série explicita que as provas internas são mais relevantes do que o EPLIS pelas sérias consequências que seus resultados acarretam, podendo levar até mesmo à exclusão do curso. O aluno também se mostra informado de que a reprovação no exame externo não traz implicações imediatas, sendo possível refazê-lo após certo período de tempo, sem sofrer qualquer penalidade.

> a relevância aqui das provas na escola é realmente passar (+) porque o pensamento é exatamente esse (+) porque a prova (+) se não passar aqui na prova como aluno (+) na prova da escola (+) vai ser desligado (+) porém o Eplis (+) o cara não passou nem no primeiro nível do Eplis (fase I) (+) o cara vai chegar na unidade depois e refazer o Eplis (+) pode ser que futuramente seja até que seja um nível 6 (+) mas na escola não tem segunda chance (+) não deu não deu ele vai embora (Grupo focal, comentário de um aluno da 4ª série, turma B).

Essa atitude dos alunos para com as provas de rendimento e o EPLIS é endossada pelos professores que admitem que, realmente, no cenário escolar, as avaliações de rendimento ganham mais relevância para os alunos pela possibilidade de desligamento do curso e pela preferência na escolha de localidades. Essa constatação fica ainda mais evidente na fala de Pilar quando perguntada sobre a relevância do EPLIS: "(as provas de rendimento) são BEM mais importantes (+) definem nota (+) definem nota (+) definem o lugar que ele vai morar (+) definem a VIDA". A professora Ana acrescenta que os alunos na EEAR são orientados pela nota nas avaliações em sala de aula:

> aqui eles estão muito focados na NOTA (+) eles querem nota (+) então, eles ficam preocupados com a PDP (provas de rendimento) (+) o Eplis (+) eles sabem que depois vão ter outras chances de fazer (+) então, o que eles fazem" (+) eles pegam o que é importante AGORA' (+) gente, o importante agora é eu estudar (+) tirar nota na prova pra ter uma boa classificação (+) e escolher uma localidade legal pra mim /../ porque eles não conseguem colocar o Eplis na frente disso (+) o foco é outro (+) o Eplis fica como segundo plano (+) eles têm preocupação SIM (+) mas ainda como segundo plano (Entrevista com Ana, professora na EEAR desde 2013).

De fato, a reprovação no EPLIS não impede que o aluno conclua o curso de formação e tampouco que prossiga para o estágio operacional nas unidades de controle de tráfego aéreo para as quais serão designados. Como bem sabem os alunos, até o momento, é possível refazer o exame sem que isso acarrete prejuízos profissionalmente. A ausência de consequências atreladas aos resultados do EPLIS faz com a percepção de sua relevância seja, portanto, diminuída entre os alunos, professores, instrutores e instituição, em comparação às avaliações escolares, o que pode reduzir a intensidade de seu efeito. Em resumo, com base no conceito de significância de Stoneman (2006), podemos assumir que o EPLIS perde relevância,

principalmente entre os alunos da 4ª série que estão próximos à conclusão do curso, em razão de que, embora os resultados obtidos pelos alunos sejam aceitos pelo DECEA, eles não são exigidos por essa instituição e, tampouco, impedem que os alunos recebam o certificado de conclusão do curso de formação.

Estão em andamento tratativas acerca da possibilidade de os resultados alcançados pelos alunos no EPLIS serem computados na média final da disciplina de inglês da 4ª série, na tentativa de fortalecer a política linguística no contexto escolar, elevando, assim, as notas obtidas pelos alunos. Em relação ao posicionamento da escola sobre esse assunto, a chefe e professora Isabel revela que há argumentos contrários e a favor:

> o […]falou assim'(+) não, isso aí tem que entrar na nota do aluno (+) pra escolher vaga (+) mas é uma discussão porque tem uns que concordam e outros que não (+) acham que por exemplo, (+) que tem controles que não são internacionais (+) e daí vai prejudicar um aluno que não passou" (+) vai desligá-lo por causa que ele não passou" (+) ou então vai privilegiá-lo pra ele escolher melhor a vaga" (+) então, são vários questionamentos aí (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

Quanto aos professores, instrutores e alunos, somente 22% dos alunos e 18% dos professores concordaram que o desempenho no EPLIS deveria ser considerado para a aprovação na disciplina de inglês da 4ª série. Essa decisão, a meu ver, não deve ser tomada sem que antes haja estudos sistemáticos sobre a validade de uso desses resultados nesse sentido, a fim de evitar que injustiças sejam cometidas. Como exemplo, remeto à fala de Isabel e proponho a seguinte reflexão: não seria um contrassenso ou até mesmo uma prática injusta desclassificar um aluno por não ter sido aprovado no EPLIS, dado que existem órgãos de controle de tráfego aéreo no Brasil que lidam exclusivamente com voos domésticos?

Outro ponto que requer atenção é o fato de a política linguística do DECEA ser, atualmente, compensatória e não-punitiva, no sentido de que vantagens são oferecidas aos aprovados ao invés de sanções aos reprovados, o que pode fazer com que aqueles com maior nível de proficiência na língua se sintam mais motivados e, consequentemente, atribuam maior relevância ao exame. Sobre isso, no estudo quantitativo, constatei uma diferença significativa no modo como os alunos mais proficientes tratam a significância do teste em relação aos menos proficientes. Além de afirmarem possuir mais conhecimentos sobre o exame e se engajarem com maior frequência em práticas preparatórias, os alunos mais proficientes se julgam mais motivados, avaliam mais positivamente a avaliação externa e a importância de serem bem-sucedidos.

Durante o grupo focal, os alunos, de modo geral, mostraram-se cientes dos benefícios que a aprovação traz e reconheceram neles o fator motivador. No trecho abaixo, um aluno menciona que a flexibilidade nas trocas de serviço possibilitaria a ele, por exemplo, frequentar uma universidade. Vale saber que o regime de escala a que os controladores são submetidos dificulta o envolvimento em atividades que requerem certa rotina. Então, a permissão de permutar o turno de trabalho com outros colegas é vista como vantajosa.

> se eu não me engano (+) tem (+) tem que ter um nível 4 em cada equipe de controle, né (+) isso acaba deixando até as trocas (+) dependendo (+) se tá fazendo uma faculdade e tal (+) deixa uma flexibilidade maior na escala (+) são fatores que incentivam bastante a gente (Grupo focal, comentário de um aluno da 2ª série BCT, turma F).

Para o instrutor João, não existe uma relação direta entre a proficiência dos alunos e o nível de relevância atribuída por eles ao exame, porém reconhece que os alunos apresentam atitudes distintas para com o teste em função do nível de proficiência apresentado. Segundo João, os alunos mais proficientes demonstram maior confiança, enquanto os menos proficientes, desânimo e preocupação.

> sim (+) os menos proficientes, aqueles que / de todos os alunos que sabem do Eplis (+) que já tão com a consciência de que eles vão fazer uma prova (+) da importância da prova (+) os menos proficientes estão desanimados (+) e preocupados (+) e os mais proficientes estão TÃO confiantes que querem o nível dez (risos) (+) né, (Entrevista com João, instrutor na EEAR desde 2016).

Na opinião de João, os alunos menos proficientes, mesmo estando conscientes da importância do EPLIS, são obrigados a focar suas energias, primeiramente, nas avaliações de rendimento para que possam ser aprovados no curso. Já os alunos mais proficientes não têm a necessidade de engendrar grandes esforços para obterem notas altas nas provas de inglês. Assim, reporta:

> João: os mais proficientes querem saber mais sobre a prova (+) é o que eu observo
> Pesquisadora: será que eles pensam que' (+) eu tenho chance de passar (+) então, eu vou me dedicar
> João: exato (+) e o que não tem ele espera o tempo passar (+) ele diz' eu tô mais preocupado em tirar a nota na escola (+) porque eu preciso tirar a nota seis ou sete pra ser aprovado no curso de formação (+) do que propriamente no exame lá fora (+) porque o mais proficiente já pensa' (+) eu já estou garantido (+) o meu dez ou nove e meio (+) agora eu vou correr atrás da minha parte profissional (Entrevista com João, instrutor na EEAR desde 2016).

A percepção de João encontra respaldo em vários estudos que analisaram a relação entre o efeito retroativo e as características individuais dos alunos, tais como motivação, autoconfiança e conhecimento do teste. No estudo de Bartholomeu (2002), os alunos com maiores níveis de autoconhecimento, autoeficácia e autonomia apresentaram atitudes mais positivas para com o exame. Retorta (2007), por sua vez, concluiu que, em contextos de ensino privilegiados nos quais as chances de aprovação no vestibular eram mais elevadas, o impacto do exame se intensificava, com alunos e professores mais motivados e mais interessados em obter informações sobre o exame e a se prepararem ou prepararem seus alunos. O estudo de Green (2007b) demonstrou que o alto desempenho obtido por alunos no exame tinha sido influenciado mais por suas características individuais do que pelo tipo de preparação recebido. Os resultados alcançados por Cheng et al (2011), por sua vez, revelaram que os alunos com níveis mais altos de proficiência apresentavam maior motivação para o estudo da língua, realizavam mais atividades preparatórias dentro e fora da sala de aula, sentiam-se menos preocupados e possuíam visões mais positivas acerca de exames externos.

Tanto no estudo de Cheng et al (2011) quanto nesta pesquisa, os níveis de proficiência foram definidos pelos próprios alunos, ou seja, são níveis de proficiência percebida e, por essa razão, podem estar associados à autoconfiança ou autoeficácia. O conceito de autoeficácia é definido como a confiança que uma pessoa deposita em sua própria capacidade de completar com sucesso determinada tarefa ou resolver determinado problema. A relação entre a autoeficácia e preparação para o teste foi investigada por Xie e Andrews (2012). Esses autores apontaram que as percepções dos indivíduos sobre as demandas do exame impactaram mais a preparação do que a percepção de seus usos (relevância). Ou seja, a percepção de que as exigências do teste eram possíveis de serem atingidas, associada a um alto grau de autoeficácia, fez com que houvesse um maior engajamento dos indivíduos na preparação para o teste.

A meu ver, o que parece estar ocorrendo na EEAR se aproxima das conclusões de Xie e Andrews (2012) no sentido de que, frente à baixa relevância imediata do EPLIS para os alunos, o efeito retroativo é mais intenso para aqueles mais proficientes que, ao conhecerem melhor as demandas do exame, veem chances reais de aprovação (autoeficácia) e, por essa razão, sentem-se mais motivados, demandam mais informações, demonstram atitudes mais positivas para com o teste e decidem se engajar em atividades preparatórias. Os menos proficientes, por considerarem as chances de aprovação pequenas, acabam mostrando pouco interesse pelo EPLIS, o que faz com que o impacto do teste seja menor entre esses alunos.

5.2.3 As situações de ensino e aprendizagem

Como já discutido no Capítulo II, as características do teste podem influenciar as situações de ensino e aprendizagem dentro e fora do ambiente de sala de aula em termos de conteúdos, habilidades e tarefas. Na análise das respostas dos alunos ao questionário, na fase quantitativa, constatei, primeiramente, diferença significativa em relação à 1ª série e às demais em termos de conteúdo (tópicos e habilidades). Esse resultado, no entanto, não causou espanto dado que o currículo escolar programático da 1ª série contemplava, na época da pesquisa, o ensino de inglês geral, com a justificativa de oferecer base linguística aos alunos provenientes de diferentes regiões do país e classes sociais. Além disso, embora o currículo contemplasse alguns tópicos ligados à aviação, esses não estavam relacionados com o repertório de uso do controlador de tráfego aéreo.

Abordei essa questão no grupo focal e os alunos se colocaram bastante conscientes da inadequação desse currículo para as necessidades e interesses profissionais. No trecho abaixo, um aluno da 1ª série argumenta que as funções comunicativas desenvolvidas em sala de aula não têm aplicabilidade no ambiente profissional e cita como exemplo "saber servir em um restaurante". Em relação aos tópicos de aviação abordados, ele também se mostra capaz de, mesmo estando no primeiro semestre do curso, perceber que o conteúdo ensinado sob o rótulo de inglês aeronáutico em sala de aula não condiz com suas futuras funções. Assim, enfaticamente, afirma que não será um comissário de voo, mas, sim, um controlador de tráfego aéreo (BCT).

> na 1ª série o que a gente tá tendo (+) acho que em todas as salas (+) é áudio sei lá (+) de restaurante (+) de comida de [exatamente] não faz sentido (+) [e cardápio] eu não vou trabalhar em restaurante (+) eu não vou trabalhar servindo comida japonesa (risos) (incompreensível) (+) começou um pouquinho de vocabulário agora (+) mas é quase nada (+) uma coisinha aqui uma coisinha ali (+) mais uma situaçãozinha de *flight attendant*, né (+) que são [comissárias de voo] comissária de bordo lá (+) eu também não vou ser comissário de bordo (+) eu vou ser BCT (controlador de tráfego aéreo) (+) eu quero saber como que eu vou trabalhar em uma situação de emergência (+) como que eu vou poder falar com o piloto em uma situação que ele vai' tá' falando (+) mais (+) não descontraído (+) desesperado no caso (+) e não como eu vou servir uma comida (+) o que eu vou pedir no restaurante (+) isso não faz sentido (Grupo focal, comentário de um aluno da 1ª série BCT, turma E).

Nessa falta, também é possível notar a preocupação do aluno em relação à capacidade de se comunicar com o piloto em uma situação de risco, o que demonstra consciência da importância da língua para a prestação do serviço de tráfego aéreo e de sua

responsabilidade em saber lidar com as demandas comunicativas impostas pela natureza emergencial da situação.

Outro ponto levantado pelos alunos no grupo focal diz respeito à falta de motivação para a aprendizagem de inglês causada pela incompatibilidade entre o currículo da 1ª série e as características da comunicação aeronáutica, conforme podemos verificar no excerto abaixo.

> isso de fato desmotiva o aluno (+) a gente tava conversando (+) a gente da 2ª série (+) ih: a gente chegou a essa conclusão (+) que a gente vem muito desmotivado da 1ª série por não ter (+) em relação ao inglês, né (+) por não ter o contato pleno com o inglês de aviação (+) com o Eplis (Grupo focal, comentário de um aluno da 2ª série BCT, turma F).

Nesse trecho, notamos que o aluno deixa explícito que essa opinião não é individual, mas, sim, de um grupo que compartilha a visão de que o ensino de inglês geral não satisfaz suas necessidades e interesses enquanto futuros profissionais: "a gente tava conversando (+) a gente da 2ª série (+) ih: a gente chegou a essa conclusão".

A apatia reportada pelos alunos está claramente associada ao fato de o currículo de inglês geral não atender as demandas profissionais, o que é discutido no Doc 9835 AN/453 da OACI que, por sua vez, argumenta a favor de programas de ensino de línguas para fins específicos como fator motivador. Segundo o documento, até em níveis mais baixos linguisticamente, os materiais relacionados ao contexto profissional são mais atraentes por suscitar nos alunos a sensação de propósito e relevância do que está sendo ensinado e aprendido. Tornar os programas de ensino de inglês no contexto aeronáutico mais específicos para atender às características da comunicação entre piloto e controlador de tráfego aéreo constitui, portanto, um dos efeitos pretendidos da política da OACI de implementação de exames de proficiência para esses profissionais. Ademais, o planejamento de ensino em torno das necessidades e interesses dos alunos é um dos princípios norteadores da abordagem comunicativa (ALMEIDA-FILHO, 2002; VIEIRA-ABRAHÃO, 2015).

Apesar dessa recomendação oficial, o ensino de inglês geral na 1ª série foi justificado no currículo da disciplina pela necessidade de oferecer uma base linguística aos alunos. Essa decisão pelo inglês geral me parece ser pautada em uma visão estruturalista de ensino que tem como foco principal o desenvolvimento da competência linguística do aluno de forma hierarquizada e descontextualizada. Na abordagem comunicativa, a competência linguística é parte integrante da competência comunicativa e, portanto, seu ensino deve, sim, ser promovido em sala de aula, porém em nível discursivo e pragmático. Dessa forma, não vejo

como inconciliáveis o desenvolvimento da base linguística dos alunos e um currículo específico.

Ademais, ainda que não haja consenso em relação à inserção do ensino de inglês para fins específicos nos níveis iniciais de aprendizado do inglês, particularmente, acredito que, nos cursos de formação de controladores de tráfego aéreo, em que restrições de tempo e recursos estão presentes, devem ser privilegiadas em sala de aula as necessidades do futuro profissional, a fim de que ele possa desempenhar suas funções comunicativas na língua inglesa dentro dos parâmetros minimamente aceitáveis. Não obstante, sei que algumas limitações irão existir, dado que os alunos ainda estão em processo de construção de seus conhecimentos técnicos, mas é possível alinhar o currículo de língua inglesa ao que está sendo estudado em outras disciplinas, maximizando, assim, os tempos de aulas dedicados ao estudo do inglês para aquilo que será, de fato, utilizado pelos alunos nos órgãos de controle de tráfego aéreo.

Ademais, os alunos também demonstram clareza sobre os benefícios na implementação do ensino de inglês aeronáutico desde o início do curso. No trecho abaixo, retirado dos comentários ao questionário, um aluno da 2ª série argumenta que os resultados no EPLIS seriam melhores se a 1ª série também focasse no ensino de inglês para fins específicos, assim como as demais.

> Porém, na 1ª série, os assuntos abordados são totalmente distantes da aviação. Penso que seria de extrema importância e de grande valia que os alunos Bcts da 1ª série pudessem ter um inglês mais voltado para a aviação, assim como ocorre nas outras séries. […] se os métodos aplicados na 2ª, 3ª e 4ª séries fossem aplicados já na primeira série, tenho certeza que os resultados no Eplis seriam muito mais satisfatórios e teríamos profissionais de tráfego aéreo mais bem preparados (Questionário, comentário de um aluno da 2ª série BCT, turma D).

A percepção de distanciamento existente entre os conteúdos abordados na 1ª série, e o contexto do controle de tráfego aéreo se estende também aos professores. Eduardo enfatiza a miscelânea de assuntos abordados que não condizem com a realidade profissional dos alunos e deixa subentendido que também é a favor do ensino de inglês específico desde o início do curso: "enfim, a 1ª série é todo meio (+) atirando pra tudo quanto é lado né (+) eles aprendem pizza (+) técnicas de relaxamento (+) animais da fazenda (+) cores (+) sabe assim,". A professora Ana, que também atua na 1ª série do curso, quando perguntada sobre a influência do EPLIS nas diferentes séries, responde: "a 1ª série eu vejo um pouquinho mais distante (+) que lida com o inglês geral (+) até entrar um pouquinho só de aeronáutico". Esse "pouquinho de aeronáutico" a que Ana se refere são aqueles tópicos de aviação geral que não pertencem ao domínio de uso do controlador de tráfego aéreo, como explicado anteriormente.

Sobre as habilidades desenvolvidas na 1ª série, o desenvolvimento da compreensão oral não consta explicitamente nos objetivos do currículo da disciplina, porém a fala de uma aluna, a respeito das aulas, especificamente, nessa série, sugere que essa prática é recorrente. Os áudios e vídeos a que ela se refere abordam assuntos gerais ou ligados à aviação, porém sem relação com o controle de tráfego aéreo.

> eh, você tem compreensão de áudio e vídeo" (+) ah, tem realmente (+) tem muito áudio (+) tem muito vídeo (+) áudio pra caramba até (+) com certeza (+) cinco (referindo-se à escala *likert* utilizada no questionário, no qual 5 é o nível mais alto) (+) inclusive muitos (Grupo focal, comentário de uma aluna da 4ª série BCT, turma D).

Dessa forma, se por um lado a ênfase colocada na compreensão oral pode estar associada à conscientização dos professores de que essa habilidade é de grande importância no controle de tráfego aéreo, por outro o fato de os assuntos tratados nos áudios e vídeos não terem ligação com o ambiente profissional atesta que o impacto da política linguística e do EPLIS é ainda considerado fraco na 1ª série do curso, utilizando a definição de intensidade de Watanabe (2004a), pois não mantém uma relação estrita com as características específicas do teste e do seu conteúdo.

Em termos de atividades de compreensão oral e escrita e produção oral e escrita, relacionadas ao controle de tráfego aéreo como, por exemplo, simular oralmente ou compreender diálogos entre piloto e controlador, o estudo quantitativo também constatou diferença significativa entre a 1ª série e as demais. Esse resultado faz sentido perante a dissonância dos conteúdos trabalhados na 1ª série e o controle de tráfego aéreo, conforme constatado nas falas dos alunos e professores.

No que tange à 1ª série, verificamos, portanto, que a influência do EPLIS no currículo é praticamente nula em termos de tópicos e tarefas e considerada fraca em termos de habilidades. Entretanto, essa situação parece mudar para as demais séries do curso. No trecho abaixo, um aluno da 4ª série ressalta a diferença nos conteúdos trabalhados. Enquanto a 1ª série se dedica ao ensino de inglês geral, as 2ª, 3ª e 4ª séries contemplam o ensino de inglês aeronáutico. Nessa fala, também percebemos a relação entre a proximidade do exame e a intensidade de seu efeito. As aulas da 4ª série são apontadas como as mais voltadas ao que é abordado no teste, confirmando os resultados alcançados no estudo quantitativo.

> quando a gente chega aqui na 1ª série é algo mais geral, né (+) abrange mais aspecto geral do inglês (+) inclusive a parte na minha opinião mais difícil, né (+) do inglês é a primeira série realmente é mais os aspectos gerais (+) na segunda série já é um inglês MUITO mais voltado para a aviação (+) TOTALMENTE diferente (+) aí que a gente

> começa realmente a ter contato com o que seria o Eplis (+) porém mais no aspecto de ganhar vocabulário, né (+) terceira série (+) acho que daria para ter (+) voltado um pouco mais para o EPLIS (+) não é tão voltado para o Eplis ainda' (+) mas sim MUITO voltado pro inglês de aviação assim como na 2ª série (+) mas agora orientação de realmente como' seria o teste eh (+) como o teste seria abordado eh (+) coisas desse aspecto foram BEM mais na 4ª série (+) (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Além de sugerir que o efeito retroativo é mais explícito na 4ª série, esse comentário sinaliza que essa série é vista como um curso preparatório para o EPLIS, o que já foi apontado pela coordenadora e professora Pilar. A natureza dessa preparação e como ela se difere do que ocorre nas 2ª e 3ª séries, conforme já mencionado, foram investigados nas observações de aulas e os resultados serão reportados mais adiante no capítulo.

Os tópicos abordados pelos professores e instrutores nas aulas das 2ª, 3ª e 4ª séries estão previstos no currículo programático dessas disciplinas e refletem diretamente aqueles encontrados nos livros didáticos adotados pela escola para essas séries, conforme explicitado, no grupo focal, por uma aluna da 4ª série, turma D: "existe um Pud (programa de unidades didáticas) (+) um conteúdo programado que a gente tem em sala (+) esse conteúdo que tá programado é baseado no LIVRO (+) o que tem no LIVRO".

Embora alunos e professores tenham afirmado no questionário que o EPLIS exerce influência sobre os tópicos desenvolvidos em sala de aula, a escolha por esses materiais não tem relação direta com o exame. Deveu-se ao disposto no Programa de Elevação de Nível na Língua Inglesa, publicado pelo DECEA em 2008, do qual o EPLIS é o mecanismo garantidor da implementação, que solicitava a reformulação do currículo dos cursos de formação de controladores de tráfego aéreo para atender às necessidades linguísticas desses futuros profissionais. Esses livros eram os únicos disponíveis no mercado com o objetivo de suprir as demandas comunicativas no diálogo entre pilotos e controladores, determinadas pela OACI no Doc 9835 AN/453, e não há informações de que foram elaborados com base em um exame específico. Porém, dado que os tópicos e temas avaliados no EPLIS são determinados por esse mesmo documento, há, certamente, sobreposição entre os livros didáticos e o exame.

Por outro lado, o fato de esses livros didáticos terem como público-alvo pilotos e controladores de tráfego aéreo faz com que adaptações sejam necessárias. Ana, por exemplo, comenta que o material atende em grande parte às necessidades dos alunos, mas ressalta que certos tópicos são mais específicos para pilotos, exigindo do professor ajustes.

> eu gosto (+) gosto de trabalhar com ele sim (material didático) (+) gosto (+) eu acho que é adequado apesar de ter umas ênfases ALI que a gente percebe que é pro piloto (+) mas depende da unidade (+) não é no livro TODO (+) não, (+) aí eu acho que o

> vocabulário é bem útil pra eles como futuros controladores (+) e quando dá ênfase MUITO pra piloto (+) a gente pode / a gente tem essa liberdade de adaptar'né, (+) não precisa durante a:a atividade dar foco (+) se o exercício é foco pra piloto (+) a gente pode dar / fazer uma adaptação daquele exercício (Entrevista com Ana, professora na EEAR desde 2013).

A fala de Ana também indica que os professores e instrutores não são escravos do material e tem a liberdade de modificá-lo sempre que julgarem conveniente. Contudo, ao ser indagada sobre o papel do livro didático nas 2ª, 3ª e 4ª séries, a coordenadora e professora Pilar esclarece que devido às avaliações serem iguais para todas as turmas e possuírem alta-relevância no contexto escolar, existem conteúdos comuns nos livros didáticos que devem ser, obrigatoriamente, cobertos por todos os professores, em todas as turmas. Uma vez coberto o conteúdo comum, Pilar ressalta que os professores teriam autonomia para trabalhar algo diferenciado de acordo com a necessidade dos alunos, como pontuou Ana.

> eu costumo já eh dizer pra equipe toda vez que eles vão pegar uma lição (+) quais são os exercícios (+) quais são as atividades que não podem ser (+) eh: não podem ser / não devem ser / desculpa (+) que devem ser seguidas à RISCA (+) porque assim" o que acontece" (+) a avaliação é igual pra todos (+) então, tem aquele conteúdo que não pode DESCARTAR (+) o RESTO eu deixo assim' bem livre (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

Ainda que Pilar ateste que há abertura em relação aos conteúdos abordados, acredito que a necessidade de cobrir um conteúdo comum restringe os professores e instrutores a levar outros assuntos para dentro da sala de aula, principalmente, em turmas menos proficientes que demandam mais tempo para a realização das atividades. O papel do livro didático na sala de aula será mais discutido na análise das observações de aulas.

Quanto à composição do chamado conteúdo comum, a professora Ana explica que, em relação à 4ª série, as tarefas com foco na compreensão oral e as propostas de prática oral (descrição de figuras) são sempre trabalhadas por ela, assim como o vocabulário trazido na unidade.

> sim (+) eh:: (+) essas partes de atividades de *listening* a gente nunca / eu nunca pulo (+) procuro / a gente tem que trabalhar TODAS (+) essa parte de *listening* e a parte de *describe the picture* (+) isso a gente:ee não a gente não elimina da aula (+) de forma alguma (+) vocabulário eh:: (+) TEM que passar o vocabulário pros alunos senão eles ficam usando o vocabulário que eles já aprenderam (+) não tem nada novo (+) né, pra tá inserindo (+) então, essa ordem'(+) vocabulário (+) eh: depois as atividades de *listening* e *describe the picture* (+) são as exigidas (Entrevista com Ana, professora na EEAR desde 2013).

Essa fala de Ana corrobora as orientações passadas pela coordenadora e professora Pilar à equipe de professores que atuam nas 2ª, 3ª e 4ª séries. Pilar diz que costuma se reunir com os professores e instrutores para enfatizar a importância de que as habilidades de compreensão e produção oral sejam sempre privilegiadas em sala de aula. Assim, descreve:

> eu sempre eh: no início do curso eu costumo fazer uma reunião a respeito do que é esperado que o professor (+) desempenhe em sala de aula (+) que tipos de atividades eles devem fazer, né (+) que a gente tem que tá pautado para a produção oral e atividades de *listening*, tá (+) que toda aula tem que ter produção oral, isso que eu passo para os professores /.../ eu foco SEMPRE na importância das duas habilidades (+) do *listening* e *speaking* /.../ eu sempre FALO que nós precisamos fazer os alunos falarem (+) tentar cada vez mais diminuir o *teacher talking time* (+) deixar os alunos fazerem / produzirem / pra trazer atividades que façam (o aluno falar) (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª, e 4ª séries e professora na EEAR desde 2005).

Outro ponto interessante na fala de Pilar refere-se à diminuição do tempo de fala do professor em sala de aula e, consequentemente, ao aumento do tempo de fala do aluno. Promover o envolvimento do aluno em atividades comunicativas em pares ou em grupos é um dos princípios norteadores da abordagem comunicativa (VIEIRA-ABRAHÃO, 2015) e um dos efeitos pretendidos da política linguística no contexto aeronáutico.

Ainda sobre habilidades, no estudo quantitativo, a frequência com que atividades de compreensão oral de conversas radiotelefônicas entre piloto e controlador foi o item em que os professores e instrutores mais concentraram suas respostas nos níveis "frequentemente" (38%) e "sempre" (44%). Os alunos também avaliaram essa atividade como bastante recorrente em sala de aula a partir da 2ª série do curso. Ao considerarmos que a fase I do EPLIS tem como foco, exatamente, a avaliação da compreensão de comunicações aeronáuticas, por meio de 30 questões de múltipla escolha, acredito que o EPLIS exerce influência nessas séries não apenas em termos de habilidade (compreensão oral), mas, também, em termos de gênero (comunicação radiotelefônica entre piloto e controlador) a ser desenvolvido.

As atividades relacionadas à prática de escrita foram as que atingiram menor média em relação à frequência com que são desenvolvidas, tanto na perspectiva dos professores e instrutores quanto na perspectiva dos alunos. Uma ênfase menor na produção escrita coincide com o fato de o exame não avaliar essa habilidade. Nas discussões do grupo focal, nenhum aluno se posicionou contrariamente a essa questão. Todavia, uma aluna da 2ª série criticou no questionário a falta de atividades dessa natureza. Na sua opinião, esse tipo de prática auxilia a organização do texto falado e deveria ser trabalhado com mais frequência em sala de aula. Assim, relata:

> Sinto falta de exercícios escritos pois são quase inexistentes e acho eles muito importante até porque eles ajudam a melhorar a organização de ideias e o melhor escolhas de palavras quando se cria uma história ou tem que contar uma. Quero enfatizar que os exercícios escritos são quase inexistentes e que precisam melhorar nesse aspecto. (Questionário, comentário de uma aluna da 2ª série BCT, turma A).

Na análise das observações de aulas, esse aspecto é mais bem discutido. Adianto que, com base nas visitas técnicas realizadas, em práticas de produção oral, alguns alunos, principalmente, aqueles com baixa proficiência na língua, recorrem à escrita para conseguirem realizar o que está sendo proposto em uma espécie de preparação para a fala.

Divergindo um pouco das percepções dos professores, o estudo quantitativo mostrou que, para os alunos, a atividade de leitura de textos relacionados à aviação e ao controle de tráfego aéreo é a mais recorrente em sala de aula, na frente, até mesmo, da compreensão oral de conversas entre piloto e controlador. Se por um lado, a prática da leitura nos mostra que a sala de aula não está restrita às habilidades avaliadas no exame, somado ao fato de que o texto escrito pode ser usado para fomentar a prática da produção oral, por outro lado, em termos de efeitos, esperava-se que fossem priorizadas as habilidades de compreensão e produção orais dos alunos e a integração entre elas devido à natureza do exame. Lembro que, apesar de o EPLIS avaliar a compreensão oral isoladamente na fase I, na fase II integra-se essa habilidade à produção oral.

Os alunos também externaram que o desenvolvimento das habilidades orais, especialmente, produção oral, não é tão bem enfatizado em sala de aula, coincidindo com resultados obtidos por Chini (2014), em um estudo que analisou as percepções dos alunos da 4ª série BCT sobre a consonância entre o curso de inglês da EEAR e as necessidades de aprendizagem. Ao serem indagados por Chini (2014) sobre qual habilidade deveria ter sido mais trabalhada nas disciplinas de inglês, 60% dos alunos afirmaram ser a produção oral e justificaram que não se sentiam ainda confiantes no domínio dessa habilidade. Em segundo lugar ficou a compreensão oral com 25% das respostas. Em seus comentários, os alunos mostraram-se bastante conscientes de suas necessidades, ao alegarem que essas duas habilidades são as mais importantes para o serviço de controle de tráfego aéreo e, por essa razão, deveriam ter sido mais bem exploradas no curso da EEAR.

Na tentativa de entender a razão pela qual a compreensão de textos escritos foi avaliada pelos alunos como a atividade mais recorrente em sala de aula, verifiquei o papel dado à leitura nos livros didáticos. Como argumentam Wall e Horak (2007) e Spratt (2005), o material didático pode direcionar os tópicos, habilidades e atividades dentro de um curso.

Ademais, vimos que, em termos de tópicos, o currículo das disciplinas reflete por completo as unidades presentes nos materiais didáticos.

Os livros *English for Aviation* e *Aviation English*, utilizados respectivamente nas 2ª e 3ª séries do curso, abordam a compreensão de textos escritos, pelo menos, uma vez por unidade didática, porém como forma de fomentar a produção oral. Ou seja, a leitura serve como ponto de partida para promover discussões em pares ou em grupos, além de fornecer insumo para o ensino contextualizado de vocabulário e funções. Em nenhum desses livros, a leitura é enfatizada em detrimento das habilidades orais. Essa também parece ser a conclusão de Bonifácio (2015) quanto ao livro didático *Aviation English*. Bonifácio fez uma análise avaliativa desse material e pontuou que há, em média, uma atividade de leitura por unidade com o objetivo de introduzir o tópico a ser estudado, enquanto a compreensão oral atinge, em média, quatro atividades por unidade, e a produção oral nomeia, explicitamente, duas atividades. Já o livro utilizado na 4ª série, *Check your Aviation English*, não aborda a leitura. Pelo contrário, privilegiam-se a compreensão e a produção orais, juntamente com o ensino de vocabulário. Assim sendo, assumi que a ênfase dada na leitura em sala de aula, na percepção dos alunos, pode estar relacionada aos conhecimentos, crenças e experiências de professores e instrutores ou ser influência das avaliações de rendimento, como veremos na seção 5.2.4.

Em relação às atividades de produção oral, o estudo mostrou que a 4ª série se diferencia significativamente das demais quanto à frequência com que essas atividades semelhantes ao exame são desenvolvidas em sala de aula. Durante o grupo focal, os alunos confirmaram que a produção oral voltada para o exame é bem mais explorada na 4ª série e que nas séries anteriores essa prática varia de acordo com o nível de conhecimentos dos professores e instrutores sobre o EPLIS, conforme podemos verificar no trecho abaixo:

> Pesquisadora: vocês acham que vocês têm muita prática de produção oral na sala de aula de inglês"
> Aluno: na 4ª série sim (+) porém antes da 4ª série não (+) antes da 4ª série só os sargentos de fora que vêm (+) que tem contato (+) que sabem o que é o EPLIS e abordam bastante isso com a gente (+) a parte prática de como funciona (+) seria a 2ª fase do EPLIS, né (Grupo focal, aluno da 4ª série do curso BCT, turma B).

Os alunos mencionaram diversos tipos de tarefas, desenvolvidas pelos professores em sala de aula, semelhantes àquelas encontradas no exame, dando ênfase, principalmente, na descrição de figuras e criação de histórias baseadas nessas figuras, que correspondem às tarefas presentes na parte 4 da fase II do EPLIS. No comentário de um aluno da 2ª série, podemos constatar que alguns professores e instrutores costumam inserir na sua prática pedagógica esses dois tipos de atividades na intenção de familiarizar os alunos com o teste. Essa fala também

endossa a alegação de que antes da 4ª série a influência do EPLIS varia em função dos conhecimentos que professores e instrutores possuem sobre o exame.

> pelo menos recentemente eu passei por alguns (professores/instrutores) que (+) mandavam a gente a descrever (+) falavam que caía muita descrição no Eplis e por isso era importante a gente tá acostumado a fazer descrição de imagem (Grupo focal, comentário de um aluno da 2ª série BCT, turma A).

A prática dessas duas atividades parece ser intensificada na 4ª série, quando se torna bastante recorrente nas aulas de todos os professores, conforme aponta um aluno da 4ª série no trecho abaixo.

> agora na 4ª série foi praticamente' SÓ as técnicas para fazer o EPLIS (+) em todas as aulas (+) tinha uma imagem e o instrutor pedia pro aluno descrever e tal (+) […] depois teria que contar uma história (+) como que (+) de acordo com aquela imagem conta uma história a respeito disso (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Sua fala também ressalta que técnicas para a realização do exame são ensinadas nessa série. Lembro que essas técnicas ou estratégias podem ser consideradas adequadas ou inadequadas. Quando inadequadas elas ajudam o aluno a responder a um item corretamente, utilizando para esse fim artifícios que não correspondem à sua capacidade linguística. São consideradas adequadas quando, por exemplo, auxiliam o aluno a lidar com determinados tipos de itens, reduzindo, assim, a variância irrelevante ao construto (MESSICK, 1982; POPHAM, 1991; HALADYNA et al, 1991) e a presença delas em sala de aula evidenciam efeito retroativo.

No trecho abaixo, outro aluno da 4ª série exemplifica os tipos de estratégias ensinadas para lidar com as tarefas de descrição de imagens e criação de histórias. Na opinião desse aluno, essas técnicas os ajudam a explorar os detalhes da figura tanto no primeiro quanto no segundo plano e a organizarem uma história de forma mais coerente e coesa, o que vejo como positivo na perspectiva da preparação para o teste.

> ih: por exemplo o professor […](+) ele passou umas técnicas excelentes ali de descrever a imagem (+) como usar (+) abordar (+) por exemplo *foreground* (+) *background* (+) as partes, né, da imagem (+) eh organizar as ideias (+) […] e também foram passadas as técnicas (+) principalmente pelo professor […] dee: como eu ia descrever isso (+) qual seria a ordem de descrição disso (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Esses tipos de estratégias são classificados por Cohen (2006) como estratégias de uso, ou seja, processos cognitivos empregados conscientemente para a realização de

determinada tarefa, e quando aplicadas à uma situação avaliativa se tornam estratégias para a realização de um exame e podem ser consideradas positivas.

Outras duas tarefas mencionadas pelos alunos, porém em menor frequência, são a explicação de uma situação problemática de tráfego aéreo e a oferta de solução, encontradas na parte 2 da fase II do exame. Os comentários abaixo respaldam a alegação de que esses tipos de atividades vêm sendo introduzidos em sala de aula por alguns professores e instrutores, no caso aqueles considerados mais conhecedores do exame, se tornando mais constantes nas práticas dos professores e instrutores envolvidos na 4ª série, assim como as atividades discutidas anteriormente, evidenciando que o efeito é mais explícito nessa série.

> eu fui entender a questão de (+) *what's the situation*" (+) *what's the solution*" (+) na 3ª série (+) porque os instrutores de FORA trouxeram (+) (Grupo focal, comentário de uma aluna da 4ª série BCT, turma D).

> na 4ª série era aquele negócio (+) por exemplo (+) *what's the situation"* (+) *give a suggestion* (+) porque é uma fase do EPLIS que pô (+) foda, né (+) chegava ali oo: tenente (professor) (+) é essa situação (+) e realmente fazia como se fosse o EPLIS (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

Para lidar com essas duas tarefas em específico, uma aluna da 4ª série reporta que uma instrutora os alertou sobre a presença de elementos inesperados nas situações abordadas no EPLIS:

> eh: o piloto falava com o controlador que tinha um chapéu na pista (+) *a hat on the runway* (+) ai na resposta tava ali (+) chapéu […] então foi uma dica que deixaram na 3ª série (+) a sargento […] que iam utilizar objetos que não tinha nada a ver com aviação com situações de aviação (Grupo focal, comentário de uma aluna da 4ª série, turma D).

Essa explicação, a meu ver, parece reducionista ou até mesmo um pouco distorcida em relação ao conteúdo do teste, como se as situações nele contempladas fossem totalmente inconcebíveis do ponto de vista profissional: "utilizar objetos que não tinha nada a ver com aviação". Isso pode levar os alunos, por exemplo, a optarem por uma resposta que contenha elementos estranhos ao controle de tráfego aéreo, sem, necessariamente ser a correta, o que seria visto como prejudicial.

Os alunos também relataram que alguns instrutores de inglês aeronáutico (controladores de tráfego aéreo) costumam trabalhar em sala de aula as tarefas de responder perguntas sobre a vida profissional e emitir opinião sobre assuntos ligados à aviação e ao controle de tráfego aéreo, ambas também presentes na fase II do exame. Os dois trechos abaixo

remetem a esse fato. No entanto, não foram feitas menções à prática dessas tarefas na 4ª série do curso.

> uma coisa também que (+) o sargento […] fazia bastante com a gente (+) era chegar e perguntar coisas da vida em inglês (+) e acontece no Eplis LOGO de cara (+) o avaliador já começa perguntando pra nós (+) eh: qual a importância do inglês na aviação (Grupo focal, comentário de um aluno da 4ª série, turma B).

> uma coisa positiva (+) da 3ª série (+) que eu achei da 3ª série (+) do inglês foi (+) inclusive tá no Eplis (+) o que a gente estudou em sala (+) literalmente caiu (+) a importância da tecnologia na aviação (+) tanto o sargento […] quanto o sargento […] bateram nessa tecla com a gente e caiu na 2ª fase do EPLIS (risos) (Grupo focal, comentário de uma aluna da 4ª série, turma D).

Além disso, outra aluna relata que um instrutor da 3ª série ensinou algumas expressões de ênfase, tais como, *let me stress* e *I want to emphasize that* para serem utilizadas na parte 3 da fase I do exame, quando necessitam emitir sua opinião sobre determinado assunto em um turno de fala mais longo, com o objetivo de manterem uma boa fluência, o que parece positivo, pois permitem a familiarização dos alunos com o teste e os ajudam a lidar com esse tipo de questão. Por outro lado, não posso deixar de mencionar aqui que as tarefas de emitir opinião e responder perguntas sobre a carreira são pouco representativas das situações de uso da língua na comunicação aeronáutica, como salientaram Scaramucci et al (2010), e um foco demasiado nesses tipos de atividades em sala de aula pode vir a constituir um efeito negativo.

Com base nesses reportes, fica evidente que, a partir da 2ª série do curso, alguns professores e instrutores começam a familiarizar seus alunos com os tipos de tarefas encontradas na fase II do EPLIS, ensinando estratégias e oferecendo dicas, algumas nem sempre adequadas. Suponho que isso ocorra porque, mesmo atribuindo níveis altos aos conhecimentos que possuem sobre o exame, muitos professores e instrutores não participaram de cursos de capacitação específicos sobre o EPLIS, o que aumenta as chances de que mitos e visões distorcidas sejam criados. Além disso, encontramos indícios de que ainda existem lacunas nos entendimentos acerca do exame.

Na 4ª série, os comentários indicaram que a prática de tarefas semelhantes àquelas encontradas na fase II do EPLIS é intensificada, passando a envolver todos os professores e instrutores, o que confirma a hipótese levantada no estudo quantitativo de que a ocorrência dessas atividades é significativamente mais frequente nessa série do que nas demais. Como disse anteriormente, a natureza dessa preparação será discutida na análise das aulas, a fim de investigar se existe a possibilidade de estreitamento das práticas em sala de aula ou de treinamento antecipado de prováveis respostas sem ganho em termos de proficiência.

No entanto, antes de associar esse aumento na ocorrência de atividades semelhantes ao exame a um provável efeito retroativo, fiz uma breve análise dos tipos de exercícios com foco na produção oral, propostos pelo livro didático adotado na 4ª série. Encontrei, pelo menos, três momentos para a prática da oralidade em cada uma das unidades do material. No início de cada lição, é pedido aos alunos que descrevam uma figura e respondam perguntas sobre ela. Em um segundo momento, é proposta a criação de um diálogo (*role-play*) entre um piloto e controlador. No final da unidade, são apresentadas algumas perguntas abertas sobre o tema da lição para discussão entre os alunos.

É possível notar algumas sobreposições de tarefas entre o EPLIS e o material. A descrição de imagens e a emissão de opinião sobre questões aeronáuticas estão presentes tanto no exame quanto no livro didático do aluno da 4ª série, apontando que o enfoque nessas atividades pelos professores e instrutores pode ser também influência do último. No entanto, contar uma história sobre a imagem apresentada, explicar uma situação problemática e oferecer uma sugestão para solucioná-la são próprias do EPLIS, e a presença dessas tarefas nos eventos em sala de aula atestam impacto do teste. Sobre isso, na entrevista conduzida com a professora Ana, quando perguntada sobre o papel do livro nas atividades que propõe em sala de aula, ela ressalta que a criação de história não é parte da proposta do material da 4ª série, porém ela e os demais professores costumam desenvolver essa atividade por demanda do EPLIS. Assim, comenta: "é (+) *telling a story* também não tá no livro e a gente acaba adaptando".

Isabel, chefe da Subseção de Idiomas e professora, também comenta em sua entrevista que, nos últimos três anos, é possível verificar o papel de destaque dado ao EPLIS em sala de aula por meio de menções ao teste e atividades preparatórias. Essa fala de Isabel corrobora os resultados obtidos na comparação realizada entre as percepções dos professores antes da implementação do EPLIS, coletadas há, aproximadamente, 3 anos, por meio de um questionário preliminar, e as percepções dos professores obtidas em 2016, de que houve uma mudança de posicionamento dos professores em relação ao exame e à sua influência no ambiente escolar, indicando um provável aumento da intensidade do efeito retroativo. Assim, Isabel comenta:

> então, a três anos atrás a gente não verificava (+) né (+) o Eplis assim' inserido em sala de aula (+) mas hoje a gente consegue verificar (+) porque foi implantado (+) não a metodologia (+) mas o modo que é dado / os moldes que é dado o Eplis foi implantado em sala de aula (+) no dia a dia (+) JÁ na 1ª série a gente começa a comentar do Eplis né (+) a importância do Eplis (+) que os alunos vão fazer uma vez por ano se eles não tiveram o nível 4 lá quando se formarem (+) então, hoje durante as aulas a gente já conversa (+) às vezes na 1ª série a gente trabalha o *describe the picture* (+) bem básico (+) aquele que é o inglês geral (+) pra eles saberem as

> preposições (+) e na 2ª e 3ª série já tem um projetinho também pra descrever e na 4ª já é mais a prova (+) a gente trabalha a figura (+) a transcrição do inglês por português (+) e também mostrar a questão (+) a importância do EPLIS NÃO fazendo terrorismo (+) mas falar que é da carreira dele (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

Também fica evidente na fala de Isabel que a intensidade do efeito retroativo em sala de aula aumenta com a proximidade do EPLIS, conforme já apontado pelos alunos e pela coordenadora e professora Pilar ao salientar que a 4ª série tem como foco a preparação dos alunos para o exame. Outro ponto mencionado por Isabel é o receio de que a avaliação externa possa exercer uma pressão muito forte nos alunos, o que poderia acarretar em práticas preparatórias indesejáveis. Acredito que essa seja uma das causas de os professores e instrutores optarem por não discutir aspectos do teste ou oferecer informações muito detalhadas nas séries iniciais do curso.

Na entrevista com Pilar, ela também reconhece que sua prática pedagógica sofreu mudanças em função da entrada do exame na escola, ocorrida no final de 2014. Algumas atividades que ela não costumava trabalhar com frequência estão hoje inseridas em suas aulas como, por exemplo, recontar uma história. Pilar explica que, mesmo não sendo exatamente como a proposta do EPLIS, esse tipo de atividade contribui para o desempenho dos alunos no exame. Desenvolver tarefas com foco no conteúdo do exame, porém em formatos variados, é considerado um tipo de preparação adequada, pois é capaz de desenvolver a proficiência do aluno (POPHAM, 1991).

> ah o EPLIS influenciou SIM (+) porque embora os alunos não PERCEBAM, né, que eu tô trazendo algo que pode ajudá-los no exame /.../ eh: durante as aulas também' (+) eu costumo perguntar (+) pedir pra eles eh:: recontar aquilo que eles ouviram (+) é' o que eu fazia menos (+) depois que o EPLIS começou a rodar na escola eu comecei a fazer mais tá (+) então houve esse impacto sim (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

O professor Eduardo, quando questionado sobre o impacto do EPLIS em sua prática pedagógica, também responde que ele é maior hoje do que anos atrás, pois passou a conhecê-lo melhor, e acrescenta que o exame guia suas aulas como, por exemplo, na atribuição de atividades de descrição de figuras, atividade essa que ele não consideraria desenvolver em sala de aula se não fosse pela presença no teste, o que se caracteriza como efeito retroativo.

> sim (+) BEM mais que antes porque agora eu sei MAIS do que antes né, (+) apesar de não saber (+) tanto quanto eu gostaria (+) porque eu não sou avaliador do EPLIS (+) nunca fiz nenhum curso (+) mas eu:u que guio / o EPLIS norteia minhas aulas sim (+) tá constantemente na minha cabeça né, /.../ uma coisa que eu normalmente não

> faria (+) antigamente' quando eu entrei aqui eu não faria *describe the pictur*e (+) hoje eu faço porque eu SEI que cai no Eplis né, (Entrevista com Eduardo, professor na EEAR desde 2010)

No entanto, o fato de Eduardo dizer que o exame "norteia" suas aulas pode ser preocupante no sentido de que uma preparação no formato do exame, ao contrário da preparação de natureza variada, é considerada prejudicial, pois capacita os alunos somente naquilo que o teste avalia. Esse problema pode ser ainda agravado caso o construto não esteja bem representado no exame, ou seja, quando nem todos os aspectos do uso da língua estão contemplados.

Assim como o professor Eduardo e a coordenadora e professora Pilar, a professora Ana também é capaz de perceber o impacto do EPLIS nas atividades que conduz em sua sala de aula. Ela expõe que, se não fosse pelo conhecimento que ela construiu sobre o exame, ela não teria essa preocupação em explicar aos alunos como lidar com a descrição de figuras e a criação de histórias.

> influencia (+) influencia sim'(+) a gente fica preocupado com / em passar pros alunos aquela parte de:e (+) como descrever figuras (+) como depois contar história (+) cuidado pra não misturar as coisas né, /.../ eu não iria passar para a aluna hoje esse cuidado que ela deve tomar (+) quando vai descrever a:a figura (+) só descrever (+) então assim, é porque eu conheci um pouquinho do exame (+) do Eplis (+) que eu consegui orientá-la dessa forma (+) acho que influencia dessa maneira aí (Entrevista com Ana, professora na EEAR desde 2013).

Nota-se nesses relatos uma ênfase nas atividades de descrição e relato de histórias, já verificada nos comentários dos alunos, o que pode sugerir que esses aspectos da avaliação externa são mais salientes do que outros em sala de aula, como foi também observado no estudo de Gimenez (1999). Esse efeito pode ser considerado negativo, caso seja constatado o estreitamento das oportunidades de aprendizagem em sala de aula. Em termos de autenticidade, esses dois propósitos comunicativos, descrição e relato, são encontrados no Doc 9835 AN/453, porém Scaramucci et al (2010) alertam que, em situações de uso real da língua, essas funções não partem de uma imagem isolada, mas, possivelmente, de uma solicitação do piloto ou de uma necessidade de reportar um fato que pode ocasionar uma alteração nos eventos previstos.

Outra mudança ocorrida na 4ª série, provavelmente em decorrência da fase I do EPLIS, deve-se à inserção de tarefas em português para a prática da compreensão oral. Conforme mencionado no Capítulo III, nessa fase, o candidato é exposto a 30 áudios de conversas entre piloto e controlador e para cada áudio existe uma questão de múltipla escolha com três alternativas elaboradas em português. Essa característica do formato do teste parece

ter afetado as práticas dos professores em sala de aula. Na fala da chefe e professora Isabel, ela relata que, na 4ª série, tem sido abordada com os alunos "a transcrição do inglês por português" na intenção de familiarizá-los com o EPLIS. Isabel esclarece que, a pedido da coordenação, os professores e instrutores que atuam na 4ª série começaram a elaborar questões de múltipla escolha em português para os áudios trabalhados em sala de aula.

> então, a primeira parte (+) do que é feita no áudio (+) que é a primeira fase do exame (+) a gente trabalha na sala de aula exercícios de inglês pro português (+) daí a gente já tá fazendo essa transcrição /.../ então, a gente já teve aquele briefing que a coordenadora falou'(+) não, TODO mundo tem que fazer as questões (em português) (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

A professora Ana também assinala a mudança do inglês para o português ao ser perguntada sobre diferenças no seu modo de ensinar entre as séries. Na sua fala, a 4ª série se distingue das anteriores em relação à prática de exercícios de compreensão oral redigidos em português, o que não é comum nas demais. Assim como no estudo de Qi (2004), percebemos o possível impacto do formato da fase I do exame nas atividades da última série do curso.

> acho que na 4ª série muda (+) a 2ª e a 3ª eu acho bem próxima (+) a forma de lidar com os alunos (+) com os exercícios em sala de aula (+) agora a 4ª série ela já (+) diferencia (+) diferencia (+) diferencia sim (+) um pouquinho (+) porque o foco passa a ser o Eplis (+) e tem aquela questão de:e eles agora fazerem a prova aqui (+) e eles lidarem com a prova em português (+) as questões estão em português (Entrevista com Ana, professora na EEAR desde 2013).

Fica evidente que a transição do inglês para o português é uma maneira de familiarizar os alunos com o exame, diminuindo, assim, a variância irrelevante ao construto, quando o método utilizado na avaliação, por exemplo, pode comprometer o desempenho do aluno. Além disso, a alternância de códigos (português-inglês ou inglês-português) é uma característica da comunicação aeronáutica no contexto brasileiro. Contudo, a intenção dos elaboradores em utilizar o português na elaboração dos itens foi fazer com que a capacidade de leitura do candidato não interferisse no seu desempenho. Acredito que esse efeito é considerado inteiramente inesperado e seu impacto será discutido na análise das aulas assistidas.

Também foram reportadas pelos alunos diferenças no modo como o EPLIS é abordado em sala de aula. Na visão deles, os instrutores – controladores de tráfego aéreo, e alguns professores, por conhecerem melhor o exame, preparam os alunos mais adequadamente. A coordenadora pedagógica e professora Pilar, ao ser perguntada sobre diferenças no modo como professores e instrutores conduzem suas aulas, confirma:

> eu não acompanho MUITO mas eu vejo que assim' (+) vejo diferença sim (+) eu vejo que (+) alguns professores eles (+) tentam (+) em especial agora com os instrutores quando vêm dar aula (+) eles tentam REALMENTE eh:: mudar a didática em função do EPLIS (+) então, eles fazem mais atividades DIRECIONADAS a algo que eles vão experienciar no EPLIS (+) ATÉ mesmo EU (+) eu também faço com eles (+) coloco algumas frases na lousa e pergunto pra eles (+) ok *what's going on*" (+) *what happened*" (+) *what's the situation*" (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Pilar, em sua fala, assinala que alguns professores modificaram sua maneira de ministrar aulas em função do EPLIS e, se colocando como parte desse grupo, cita a inclusão de tarefas semelhantes àquelas encontradas no exame como exemplo de mudanças em sua abordagem de ensino. A inclusão de certos tipos de tarefas na prática do professor não, necessariamente, implica que o seu modo de ensinar se modificou. Os estudos conduzidos por Cheng (1999) e Scaramucci (1999), dentre outros, demonstraram que as mudanças advindas do exame contemplaram o conteúdo e as atividades, porém não foram capazes de afetar a abordagem do professor em sala de aula.

Em relação à dedicação ao exame fora da sala de aula, no estudo quantitativo, diferenças estatisticamente significativas foram verificadas entre o nível de conhecimento sobre o EPLIS e a preparação extraclasse, sugerindo que os alunos com mais conhecimentos sobre o exame se engajavam com maior frequência em práticas preparatórias fora do ambiente de sala de aula, fossem elas focadas no formato do teste ou de natureza mais variada. Em relação à proximidade do exame, os testes estatísticos revelaram que os alunos da 4ª série se envolviam com maior frequência em práticas preparatórias em mesmo formato, provavelmente por conhecerem melhor as características do EPLIS. No entanto, apesar das diferenças no grau com que os alunos se dedicam ao exame extraclasse, as respostas ao questionário mostraram que, de forma geral, o engajamento nessas práticas é bastante limitado.

No grupo focal, quando perguntados sobre a dedicação ao EPLIS fora da sala de aula, os alunos listaram diversas atividades. No trecho abaixo, uma aluna da 4ª série expõe que, a partir da 3ª série, ela começou a estudar inglês no seu tempo livre, indicando que a preparação extraclasse tende a se intensificar com a proximidade do teste. Essa aluna cita atividades consideradas de natureza mais geral, tais como estudar regras gramaticais, solicitar a ajuda dos professores, assistir vídeos de conversas entre piloto e controlador, assistir filmes com legenda em inglês e escutar áudios do livro novamente. Como atividades mais específicas, ela menciona vídeos sobre aviação disponíveis na internet e escutas de conversas radiotelefônicas reais.

> quando eu cheguei na 3ª série (+) eu comecei a acordar pra vida'(+) eu comecei a estudar por fora (+) por fora (+) no desespero (+) eu comprei um livro de inglês de

> gramática (+) eu enchia o saco dos professores durante o intervalo (+) não deixava ter intervalo (+) ih:: no alojamento (+) eu ia pra internet caçar vídeo de aviação com fraseologia do App São Paulo (órgão de controle de tráfego aéreo) (+) fraseologia de Nova Iorque (+) e eu fui assim (+) assim eu fui indo […] (+) os áudios do livro eu botei no celular (+) dormia ouvindo (+) eh::: filme' (+) só legenda em inglês mesmo que eu não entendesse nada (+) série' também só legenda em inglês (+) (Grupo focal, comentário de uma aluna da 4ª série BCT, turma D).

Outro aluno da 4ª série, de uma turma considerada bastante proficiente (turma B), confessa ter começado a se preparar para o EPLIS fora da sala de aula somente às vésperas da prova, assistindo desenhos animados em inglês, mas afirma que não se envolveu em nenhuma prática focada no exame, conforme podemos verificar no trecho abaixo.

> mas nada tão voltado pro Eplis (+) foi mesmo bater o desespero MESMO foi uma semana antes (+) que eu tava assistindo (+) tava vendo desenho em inglês (+) porque bateu o desespero (+) no fim deu certo (Grupo focal, comentário de um aluno da 4ª série, turma B).

O instrutor João confirma que os alunos costumam rever o livro didático, porém poucos mencionam frequentar cursos de inglês geral ou ir à biblioteca em busca de gramáticas, o que condiz com o que foi reportado pelos alunos. No entanto, na sua opinião, são raros os casos de alunos que se mostram bastante comprometidos com o exame dedicando-se a seu estudo no tempo livre. A coordenadora e professora Pilar se recorda de um aluno da 3ª série em particular:

> então (+) agora (+) vou citar o exemplo desse aluno meu (+) que entrou no *site* (+) pegou a apostilinha (manual do candidato) (+) tirou cópia da apostila (+) do manual do EPLIS (+) na hora dos intervalos ele pede pra mim (+) professora, a senhora fazer uma parte do exame comigo" (+) ele me faz ler (+) ele pega lá a parte do:: vocês têm um simuladinho (+) aí ele pega lá (+) aí eu falo pra ele mhm pego alguma uma frase e pergunto pra ele *what's the situation* (+) e ele TÁ sim se preparando (+) e assim, (+) durante a aula é o tempo TODO participando […] mas são POUCOS alunos que fazem isso (Entrevista com Pilar, coordenadora pedagógica da 2a, 3a e 4a séries e professora na EEAR desde 2005).

Esses relatos apontam que, fora da sala de aula, o EPLIS exerce efeito retroativo para alguns alunos, porém com predomínio de práticas preparatórias de natureza geral (WATANABE, 2004a), ou seja, o exame faz com que certos alunos busquem maneiras de se aperfeiçoarem na língua, mas as características das atividades das quais participam não estão relacionadas diretamente com o teste. Por exemplo, os alunos buscam desenvolver a compreensão oral por meio de filmes e séries, ao invés de procurar por áudios de radiotelefonia disponíveis na internet ou no *site* da avaliação externa.

No trecho abaixo, um aluno da 3ª série argumenta que, se os alunos fossem conscientizados sobre o EPLIS logo no início do curso, eles teriam mais direcionamentos sobre como se prepararem, aumentando, assim, as chances de aprovação. Realmente, estudos como o de Xie e Andrews (2012) trazem evidências de que conhecer as demandas do teste é essencial para que os alunos possam direcionar e controlar melhor suas práticas preparatórias. Como preparação extraclasse, esse aluno menciona a leitura do manual do candidato, porém enfatiza que essa preocupação com o exame teve início apenas na 3ª série do curso, com o aumento do contato com o teste, devido à vinda dos instrutores de inglês aeronáutico para completar o quadro docente.

> no alojamento agora o pessoal imprime o manual do candidato e fica lá lendo como que faz (+) mas SÓ na terceira série (+) se a gente tivesse um foco maior na 2ª série (+) talvez algumas turmas tivessem mais sucesso (+) na 4ª série (Grupo focal, comentário de um aluno da 3ª série BCT, turma A).

Se por um lado foi mencionado o contato com o manual do candidato, por outro nenhum aluno reportou fazer uso dos recursos disponíveis no *site* do EPLIS, tais como amostras de tarefas do teste e vídeos ilustrativos da fase II. Um comentário ao questionário de um aluno da 4ª série do curso, da turma A, chama a atenção ao evidenciar o desconhecimento sobre a disponibilidade desses recursos: "Não tenho conhecimento sobre vídeos e amostras da fase II do exame disponíveis no *site*", o que pode estar ligado à falta de divulgação ou ao desinteresse dos envolvidos em procurar, por conta própria, meios para se atualizar a respeito do teste, uma vez que o ele alcança relevância menor comparado às provas escritas.

Por outro lado, o instrutor João afirma que exige que seus alunos entrem na página eletrônica do exame à procura de mais informações. Mas, não podemos atestar que todos os professores e instrutores ajam da mesma forma. Assim, suponho que o fato de os alunos não visitarem o *site* do EPLIS pode estar associado à relevância atribuída por eles ou à falta de conhecimento sobre a existência dessas ferramentas. Na sua fala, João também ressalta a importância de os alunos conhecerem a avaliação que irão realizar e terem direito ao mesmo tipo de informação que está disponível aos demais candidatos, práticas essas que têm sido associadas a um efeito retroativo positivo e a um aumento de sua intensidade. Assim, reporta:

> o exame tem que ser eh::: sem MITO algum (+) é ostensivo (+) eles têm que ter acesso ao exame (+) tão qual o controlador lá de fora (+) o já formado (+) ele tem que / como aqui a escola é muito limitada eles têm que correr atrás de alguma maneira (+) então, eu os oriento SIM a que baixem o manual do candidato (+) vejam como é a prova (+) que isso facilita a vida deles depois (Entrevista com João, instrutor na EEAR desde 2016).

Assim como o instrutor João e a professora e coordenadora Pilar, a professora Ana sustenta que a preparação extraclasse é mínima ao dizer que, apesar da preocupação externada pelos alunos em relação ao exame, não percebe nenhum tipo de dedicação extra: "fora da sala de aula (+) não vejo (+) não vejo mesmo (+) quando CHEGA a época do exame eles ficam preocupados (+) mas SE PREPARANDO mesmo ah::: eu não vejo (+) é difícil". Isso indica que o efeito retroativo fica mais restrito ao ambiente de sala de aula, devido, provavelmente, à baixa relevância que o EPLIS parece possuir no cenário escolar.

Quanto à relação entre o nível de proficiência do aluno e o seu comprometimento extraclasse, a coordenadora e professora Pilar e o instrutor João responderam negativamente. Para as professoras e o instrutor, os alunos que se engajam em atividades preparatórias fora da sala de aula são os mais conscientes da importância do inglês no controle de tráfego aéreo, porém não necessariamente os mais proficientes na língua. A professora Ana reforça que a dedicação varia de aluno para aluno e sugere que possa estar ligada a características pessoais como, por exemplo, à seriedade com que os alunos tratam a avaliação externa:

> é que é muito pessoal (+) depende de cada aluno né, (+) tem aluno que leva isso SUPER a sério (+) nossa, vem uma equipe aplicar um teste (+) então, eu vou levar a sério (+) vou mostrar o que eu sei (+) tem aluno do nível mais basiquinho que se preocupa (+) agora tem aluno que' (+) ah, esse não é o momento agora (+) eu vou deixar pra depois (+) eu vou fazer o que eu consigo (+) vou deixar pra depois (+) então, depende (+) depende (Entrevista com Ana, professora na EEAR desde 2013).

De fato, essas percepções encontram suporte em diversos outros estudos, nos quais características pessoais, tais como autoconfiança, motivação, autonomia foram apontadas como fatores que interagem na intensidade do efeito retroativo (BARTHOLOMEU, 2002; CORREIA, 2003; LANZONI, 2004; GREEN, 2007b).

Como principal motivo que os impede de se dedicarem mais ao estudo do inglês, os alunos apontaram a sobrecarga do curso BCT. No trecho abaixo, uma aluna da 2ª série, turma D, expõe que, devido ao grande número de assuntos a serem estudados durante o curso, cada assunto acaba adquirindo um período de dedicação, que se inicia com a disciplina e termina com a realização de uma prova que atesta o nível de domínio sobre o conteúdo ministrado. No caso do EPLIS, a preocupação alcança o nível mais elevado na 4ª série pela proximidade de sua realização.

> aqui na escola a gente tem uma característica que (+) faz parte de todos os alunos Bct (+) que é assim (+) a gente tem um monte de matéria (+) prova (+) esquece (+) aí vai ter uma nova matéria (+) por exemplo (+) esse ano (+) tive navegação de declaração (+) passei (+) esqueci (+) vai vir o Sirius (+) agora a preocupação maior é o Sirius (+)

> uma preocupação de cada vez (+) então eu acho que a preocupação do Eplis realmente só chega na 4ª série (+) (Grupo focal, comentário de uma aluna da 2ª série BCT, turma D).

A fala dessa aluna parece indicar que a cultura de avaliar[66] predominante no cenário escolar é ainda de carácter somativo, classificatório e de produto, pautada em uma visão psicométrica e positivista, com foco na avaliação *da* aprendizagem ao invés de na avaliação *para* a aprendizagem (INBAR-LOURIE, 2008b; SCARAMUCCI, 2006). Na seção seguinte, abordo a influência do EPLIS nas provas de inglês na EEAR para compreendermos melhor essa questão.

Portanto, apesar de o EPLIS ser considerado uma ferramenta importante para avaliar a proficiência no desempenho da profissão, sua relevância ainda é bastante limitada no contexto escolar. As avaliações de rendimento, por acarretarem consequências imediatas, mobilizam mais os alunos e os professores, enquanto o EPLIS pode ser colocado em modo de espera. Acredito que, por essa razão, são poucos os alunos que buscam atividades preparatórias extras visando à preparação para o exame, o que faz com que o efeito retroativo do EPLIS fique restrito à sala de aula.

5.2.4 As avaliações de rendimento na EEAR

Em contextos de alta-relevância, exames externos apresentam potencial para modificar as avaliações realizadas pelos professores em sala de aula também em termos de tópicos, habilidades e métodos utilizados (ALDERSON; WALL, 1993; WALL; HORAK, 2007). As avaliações escolares referentes às disciplinas de língua inglesa da EEAR foram consideradas de alta-relevância por alunos e professores, devido às consequências imediatas que seus resultados podem causar. Analisei, então, as características mais salientes dessas provas e a influência do EPLIS na elaboração de seus itens nas percepções de alunos, professores e instrutor.

As avaliações de inglês na EEAR são divididas em provas escritas e provas orais. As provas escritas são compostas por duas partes que avaliam a compreensão de textos e a compreensão oral por meio de 40 questões de múltipla escolha, sendo 20 questões para cada parte, com peso igual entre elas. Por série, são aplicadas duas provas escritas, uma na metade

[66] O termo "cultura de avaliar" é utilizado por Scaramucci (1997) para se referir ao conjunto de crenças, pressupostos e mitos, construídos ao longo dos anos, sobre a avaliação e sua importância para o ensino e aprendizagem.

do semestre e outra no final do semestre, a fim de atestar se os objetivos de aprendizagem foram atingidos. Os resultados obtidos nessas duas avaliações são somados e, para ser considerado aprovado na disciplina, o aluno deve ter, no mínimo, 70% de acerto. A prova oral é aplicada no final da série com o objetivo de oferecer um parecer ao aluno sobre seu desempenho oral. Porém, o resultado obtido nesse teste não é contabilizado na nota final da disciplina.

A partir dessa breve descrição das provas de inglês e cientes de que dada a sua alta relevância elas podem direcionar o ensino, já é possível tecer algumas considerações. Primeiramente, acredito que a leitura e a compreensão oral, por compartilharem o mesmo grau de importância nas avaliações escritas, tendem a ser mais enfatizadas em sala de aula. Além disso, pelo fato de a produção oral não gerar uma nota, ela se torna menos valorizada tanto por parte do professor e instrutor quanto por parte do aluno, que concentram seus esforços no desenvolvimento das habilidades que valem pontos (provas escritas), uma vez que as notas causam grande impacto na vida do aluno. Tanto a análise estatística das respostas dos alunos e suas falas no grupo focal apresentam evidências para essas suposições: a prática de leitura é vista como atividade bastante recorrente em sala de aula, assim como a compreensão oral, enquanto a produção oral fica em terceiro lugar.

Sobre o papel das avaliações de rendimento na prática do professor em sala de aula, o professor Eduardo é bastante claro ao dizer que as provas escritas determinam o que deve ser coberto, funcionando exatamente como um mecanismo direcionador do ensino. Por sua vez, o conteúdo avaliado na prova é direcionado pelo material didático utilizado. A fala de Eduardo corrobora a percepção de que a habilidades de leitura e a compreensão oral são mais exploradas em sala de aula. Assim, relata:

> por causa das avaliações escritas (+) a coordenação fala' (+) oh isso aqui você tem que trabalhar (+) esses aqui são os pontos que você NÃO pode deixar de trabalhar (+) então, nesse sentido sim (+) obrigado a dar aquela matéria que vai ser cobrada na avaliação /... / nesse sentido que eu falei sim né, (+) sim (+) ela:a (+) a avaliação funciona como um direcionamento (+) ela limita ou ela determina aqueles pontos do livro que eu tenho que cumprir (Entrevista com Eduardo, professor na EEAR desde 2010).

Os alunos, quando indagados sobre o impacto do EPLIS nas provas realizadas na escola, se limitaram a citar diferenças envolvendo o formato das questões de compreensão oral das provas escritas das 2ª e 3ª séries e o exame. Eles argumentaram que os áudios nas provas em sala de aula, embora contemplem conversas entre piloto e controlador assim como o EPLIS, costumam ser mais longos, os itens de múltipla escolha são elaborados na língua inglesa e existem diversas perguntas para um mesmo áudio, o que faz com que eles tenham que processar

muita informação de uma só vez. Na 4ª série, as provas passam a trazer questões de múltipla escolha elaboradas em português, se aproximando mais do formato utilizado no EPLIS, o que pode indicar efeito retroativo. No trecho abaixo, um aluno da 4ª série, da turma B, explicita essa diferença:

> a 1ª 2ª e 3ª série as provas são totalmente em inglês porém (+) o Eplis não é totalmente em inglês (+) no Eplis as perguntas são em inglês e as respostas são em português (+) e a gente só foi ter uma abordagem assim na 4ª série (+) eh:: como foi falado (+) eles colocam um texto muito grande (+) um áudio muito grande (+) pra muitas perguntas e JÁ no Eplis não é assim (+) no Eplis tem áudios de por exemplo 10 segundos e (+) é um áudio e uma questão (+) um áudio e uma questão (+) e são / (+) realmente acaba a gente só tendo o contato (+) acho que mais próximo do Eplis foi na 4ª série e AINDA assim não foi da mesma maneira que foi o Eplis (Grupo focal, comentário de um aluno da 4ª série BCT, turma B).

A chefe e professora Isabel confirma que a decisão de implementar questões de múltipla escolha em português na prova de compreensão oral da 4ª série foi tomada há, aproximadamente, três anos, a fim de familiarizar os alunos com a fase I do EPLIS, evidenciando efeito retroativo. Assim, relata: "agora a gente já inseriu (+) nesses três dois anos pra cá (+) foi inserido a primeira parte do *listening* em português né (+) que você faz a transcrição do inglês pro português". Os professores e instrutor também afirmaram que são privilegiados áudios de radiotelefonia em função do EPLIS, ainda que os livros didáticos das 3ª e 4ª séries contemplem áudios pertencentes a outros gêneros, tais como noticiários e entrevistas.

O fato de o nível de relevância das provas escritas na EEAR ser considerado muito elevado pelos alunos faz com que eles sejam orientados pela obtenção de notas. No trecho abaixo, uma aluna da 3ª série, da turma G, revela que ser aprovada nas provas do curso e aprender inglês são objetivos diferentes, e o primeiro se sobressai devido ao medo de ser excluída da EEAR:

> a pressão de ser desligado (+) meu Deus, será que eu vou ser conseguir" (+) será que eu não vou" (+) acho que a gente fica mais preocupado com a nota (+) séra que eu vou conseguir passar para não ser desligado" porque eu tenho que tá aqui (+) […] a gente tenta passar na prova só para não ser desligado (+) essa é a meta (+) não pra aprender inglês (+) a gente vê que é importante aprender inglês depois que os instrutores falam' (+) se você não aprender inglês você não vai conseguir controlar (Grupo focal, comentário de uma aluna da 3ª série do curso BCT, turma G).

Como explica Scaramucci (1997), a visão de que a aprendizagem e a avaliação são práticas distintas e separadas está calcada em uma cultura de avaliar cuja função é meramente de verificação e promoção e que despreza a dimensão educativa da avaliação no aprimoramento

do processo de ensino e aprendizagem, o que parece ser o caso da EEAR. A fala dessa aluna também sugere que um bom desempenho nas provas não configura necessariamente que houve aprendizagem, indicando que pode haver falhas no processo de desenvolvimento dessas avaliações.

O cenário de alta-relevância das provas de rendimento, somado a prováveis problemas na elaboração do teste, propicia o desenvolvimento de estratégias para lidar com as avaliações que são consideradas indesejáveis, as chamadas estratégias para responder testes (*testwiseness strategies*). Essas estratégias são definidas por Cohen (2006) como artifícios para se chegar à resposta correta, que exploram, por exemplo, o método utilizado na avaliação e as falhas na elaboração do item de prova. Por serem independentes do nível de proficiência do aluno, são consideradas práticas prejudiciais (MESSICK, 1982; SMITH, 1991; HALADYNA; DOWNING, 2004; COHEN, 2006), ao inflarem os resultados.

Sobre as provas de inglês, uma aluna da 4ª série, da turma D, menciona casos em que o item contém exatamente o que foi falado no áudio: "aqui é mais o chute (+) você ouvia a frase (+) você ouvia a frase igual (+) marcava a questão e acertava". Outra aluna da 1ª série, da turma G, cita a técnica de procurar por pistas lexicais: "com esse MONTE de áudio que eles passam (+) com 300 (exagerando) perguntas pra um áudio grande (incompreensível) eh:: aquela dica de você procurar a palavra-chave e encaixar na frasezinha". Essas duas estratégias podem fazer com que o aluno acerte a questão sem, necessariamente, compreender o que está sendo dito.

O professor Eduardo também considera que as provas escritas apresentam problemas: "as provas de *listening* e de *reading* que o aluno tem o texto ali na frente (+) eu tenho os meus *beefs* (reclamações) (+) *I have my beefs* (eu tenho minhas reclamações)". Eduardo cita, por exemplo, enunciados com linguagem muito rebuscada que confundem o aluno, textos retirados da internet sem passar por um processo de revisão ou adequação linguística e alternativas em que a palavra do áudio ou texto é substituída por outra similar, porém não adequada para aquele uso: "então, às vezes o professor vai elaborar uma questão (+) aí ele não quer colocar a palavra do texto (+) aí acha uma parecida (+) mas aí (+) você vai ler' (+) *it's English but not as I know it*" (Entrevista com Eduardo, professor na EEAR desde 2010).

Haladyna e Downing (2004) argumentam que esses óbices são relativamente comuns em exames feitos por professores, em razão das dificuldades desses em lidar com conceitos e práticas da avaliação. De fato, os professores de línguas no Brasil enfrentam problemas relacionados ao letramento em avaliação, como apontado por Scaramucci (2006), causados pela ausência de uma preocupação mais séria dos cursos de Letras e de formação de

professores em geral para com a área de avaliação, que deixam de considerar a avaliação numa perspectiva mais social e ainda a tratam como um conjunto de conceitos técnicos e especializados.

No caso do EPLIS, os alunos que já haviam realizado o exame consideram as questões da fase I mais bem elaboradas, impedindo o uso de estratégias para responder testes. No trecho abaixo, um aluno da 4ª série, da turma C, ressalta a qualidade do teste e afirma que seus itens exigem um entendimento mais profundo do áudio, o contrário do que ocorre nas provas em sala de aula.

> no Eplis às vezes até / tinha um áudio um pouco maior ali (+) áudios de por exemplo 40 45 segundos mas pra uma questão (+) aí a gente tinha que entender todo o contexto (+) as questões do Eplis eram MUITO BEM melhor elaboradas (+) realmente tinha que entender o que estava acontecendo (+) tinha caso por exemplo (+) de entender (+) quem é o controlador quem é o piloto (+) pra conseguir responder a questão (+) coisa que não acontecia nas provas do galpão (Grupo focal, comentário de um aluno da 4ª série BCT, Turma B).

A elaboração dos itens da fase I do exame em português é apontada por uma aluna da 4ª série como outra característica que dificulta o emprego de artifícios para a escolha da resposta correta. Na sua fala, ela confirma que os itens das provas da EEAR oferecem dicas para se chegar à resposta correta sem ser necessário compreender o que está sendo falado.

> eles não colocam o aluno pra pensar porque (+) quando você vai responder alguma coisa em português (+) que você ouviu em inglês (+) você vai ter que raciocinar aquilo você está ouvindo (+) você tem que raciocinar (+) qual o significado (+) em que tempo estava (+) estava' (+) está'(+) foi' (++) então (+) aqui não (+) aqui é mais o chute (+) você ouvia a frase (+) você ouvia a frase igual (+) marcava a questão e acertava (+) na 4ª série a maioria tem dificuldade em *listening* por isso (Grupo focal, comentário de uma aluna da 4ª série BCT, Turma D).

Contrariamente à percepção dessa aluna, a decisão dos elaboradores do EPLIS de confeccionar os itens da fase 1 em português se deu na tentativa de que o desempenho do candidato não fosse prejudicado por sua capacidade de leitura, evitando, assim, variância irrelevante ao construto. Dessa forma, a dificuldade do item estaria mais relacionada à qualidade de elaboração do que à mudança do código. No entanto, estudos podem ser conduzidos no sentido de investigar essa relação.

Infelizmente, por questões de segurança, não tive acesso às avaliações escritas que me possibilitasse analisar a visão de linguagem que embasa a elaboração desses itens e se ela condiz com um ensino voltado para o desenvolvimento da competência comunicativa. Nas entrevistas, Pilar, Ana, Eduardo e João asseguraram que as questões tanto de compreensão oral

quanto textual focam no significado e não na decodificação de itens isolados de vocabulário e de escolhas gramaticais, porém os relatos dos alunos sobre o uso de artifícios para se chegar à resposta correta parecem dizer o oposto.

Em relação às provas orais aplicadas no final de cada disciplina, é importante ressaltar aqui que elas possuem uma história bastante recente na EEAR. Foram implementadas em 2015, com o objetivo de oferecer um parecer ao aluno sobre sua produção oral. Até então, somente eram avaliadas na EEAR a capacidade de compreensão oral e textual por meio de provas escritas. Ainda que não seja gerada uma nota para o desempenho do aluno na prova oral, o que reduz sua relevância entre esses participantes uma vez que o ambiente escolar parece estar direcionado à obtenção de resultados, vejo essa mudança como um progresso alavancado por influência da entrada do EPLIS na escola, ocorrida no 2º semestre de 2014.

Analisei uma versão da prova oral aplicada aos alunos da 3ª série e a proposta da prova oral para a 4ª série. A prova oral da 3ª série está dividida em cinco partes que avaliam, separadamente, cinco dos critérios previstos na tabela de níveis da OACI, utilizada no EPLIS: pronúncia, estrutura, vocabulário, fluência e interação. Na primeira parte, são apresentadas ao aluno quatro palavras relacionadas ao trabalho que devem ser pronunciadas por ele com o objetivo de gerar uma nota para o critério "pronúncia". Na segunda parte, o aluno recebe quatro sentenças escritas que devem ser explicadas/parafraseadas, a fim de que o critério "vocabulário" seja avaliado. A terceira parte consiste em quatro situações (sentenças) para as quais os alunos devem aplicar regras gramaticais como, por exemplo, substituir a voz ativa pela passiva. Na parte 4, é apresentado ao aluno um diagrama para que seja criado um diálogo, simulando uma comunicação entre piloto e controlador, com o objetivo de avaliar sua capacidade de interagir na língua. Somente nessa parte, a prova é feita em dupla. A última parte consiste na leitura de um pequeno texto que deve ser resumido oralmente pelo aluno, a fim de que o professor tenha subsídios para avaliar sua fluência. Fui informada de que a prova oral aplicada à 2ª série só se diferencia do teste da 3ª série na parte 2. Ao invés de sentenças, são apresentadas aos alunos da 2ª série palavras escritas relacionadas ao controle de tráfego aéreo para serem parafraseadas.

Primeiramente, quanto às tarefas propostas, posso dizer que duas delas apresentam semelhança àquelas encontradas no EPLIS. A tarefa que consiste em explicar uma situação problemática de tráfego aéreo está presente no exame, porém, no EPLIS, o candidato pode utilizar ou não uma paráfrase para reportar a situação. Essa similaridade entre a prova oral e o EPLIS também é apontada por João como influência do exame externo.

> o Eplis tem entrevista na segunda fase (+) a prova oral tem situações na 3ª série que eles são MUITO semelhantes (+) como a segunda parte da entrevista do *what's the problem* (+) o aluno não tem a chance ainda de dizer uma sugestão pro problema (+) mas ele já tem a situação que ele tem que parafrasear (Entrevista com João, instrutor na EEAR desde 2016).

Outra correspondência refere-se à tarefa de recontar uma história a partir da leitura de um pequeno trecho. Embora a proposta do EPLIS seja um pouco diferente, pois não envolve um texto, mas, sim, uma figura com base na qual é pedido ao candidato que crie uma história, essas tarefas podem ser consideradas semelhantes por envolver um relato de eventos passados.

Um aluno da 4ª série também afirma que o EPLIS e a prova oral têm muitos pontos em comum, se diferenciando apenas em relação à criação de um diálogo (*roleplay*), atividade essa encontrada no teste oral da EEAR. Nessa tarefa, dois alunos, com base em um diagrama, se interagem como piloto e controlador visando à solução de um problema. No trecho abaixo, o aluno expõe sua dificuldade em lidar com essa proposta, o que me faz questionar se ela é realmente trabalhada em sala de aula pelos professores e instrutores e se esses sabem lidar com esse tipo de atividade.

> o teste oral que a gente faz é parecido porém não é idêntico (+) mesmo na 4ª série a gente faz o teste (+) só que acaba fugindo (+) por exemplo (+) tem um negócio lá que (+) até esqueci o nome daquilo (+) que parece lá fazer uma conversação, né (+) entre o controlador e o piloto [*roleplay*] (+) isso não tem no Eplis e esse negócio (+) a gente se enrolava todo pra falar (+) não saía nada (+) um ficava olhando pra cara do outro [(incompreensível) não sabia o que tava acontecendo (+) tentava encaixar a fraseologia mas no fim não saía nada (+) isso tinha no nosso exame oral (+) (Grupo oral, comentário de um aluno da 4ª série BCT, turma B).

Perguntada sobre a proposta de *role-play* na prova oral, Pilar justifica que, por meio dessa tarefa, os professores e instrutores são capazes de avaliar o desempenho do aluno em termos de interação. Além disso, ela pontua que não haveria outra forma de averiguar a confiança e a segurança do aluno na interação com o outro, no caso o piloto, sem que ele fosse exposto a uma tarefa similar a que ele irá desempenhar em sua profissão.

> é a comunicação (+) se a gente precisa que os nossos controladores tenham eh:: sejam seguros, né (+) tenham segurança pra lidar com a língua (+) para usar a língua naa:s comunicações radiotelefônicas (+) eu acho que fica difícil' você inferir essa interação né, (+) sem expô-lo, né (+) sem trazer uma situação emergencial ou algo que (+) que possa eh: (+) algo que realmente vá trazer alguma dificuldade pra ele na hora de produzir a língua, né (+) de usar a língua (+) mas isso é uma questão mais de logística (+) mais de (incompreensível) mas também de avaliação, né (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Realmente, esse tipo de tarefa, apesar de não estar contemplado no EPLIS, é considerado bastante representativo do uso da língua na comunicação aeronáutica, uma vez que simula uma interação real. A proposta de uma tarefa que difere daquelas encontradas no exame mostra que, embora existam evidências de efeito retroativo na prova oral, essa avaliação não se limita ao EPLIS. Por outro lado, o fato de os alunos terem apontado dificuldades em desenvolvê-la sugere que a atividade de *role-play* não é tão enfatizada em sala de aula, o que necessita ser averiguado.

Outra similaridade encontrada entre o EPLIS e a prova oral da 3ª série é o emprego da tabela de níveis da OACI para avaliar o desempenho do aluno em termos de interação e fluência. Para isso, o professor responsável por conduzir a entrevista com o aluno preenche uma ficha de avaliação, na qual constam excertos da descrição desses dois aspectos como, por exemplo, "o aluno apresenta respostas imediatas, informativas e apropriadas ao problema proposto" na avaliação da interação. Todavia, ainda que o uso dessa tabela possa ser considerado evidência de efeito retroativo, não é possível afirmar que os professores compreendem e sabem usar adequadamente seus critérios, dado que eles carecem de capacitação na área, como foi reportado por eles mesmos no estudo quantitativo, além de alguns ainda terem alegado pouco conhecimento sobre o exame e sobre a tabela de níveis.

Somado a isso, o fato de os critérios serem avaliados somente em partes específicas da prova ao invés de ao longo de toda a avaliação o distancia do EPLIS e gera preocupação. A pronúncia, por exemplo, é avaliada apenas na parte 1, quando o aluno é solicitado a ler em voz alta quatro palavras isoladas. Isso faz com que o conceito de pronúncia seja reduzido à reprodução de sons, desconsiderando aspectos como tonicidade, ritmo, sotaque e entonação da fala, presentes na tabela da OACI. O mesmo ocorre com a estrutura. São apresentadas quatro sentenças para que o aluno aplique determinadas regras gramaticais como em um teste de conhecimento. Não se avalia o domínio da estrutura pelo candidato no uso da língua, tampouco se os erros cometidos interferem ou não na comunicação, conforme previsto na tabela. A avaliação do vocabulário também fica restrita à paráfrase de quatro palavras ou sentenças e desconsidera se o aluno demonstra vocabulário suficiente e adequado para lidar com as demais situações presentes na avaliação.

Outra diferença na comparação entre a prova oral da EEAR e o EPLIS é a ausência de interação entre interlocutor (no caso, o professor) e candidato (aluno). Na prova oral, o aluno recebe fichas escritas (palavras a serem pronunciadas, situações a serem parafraseadas, texto a ser resumido), não havendo, portanto, uma interação com o interlocutor, o que impede que a compreensão oral do aluno seja avaliada. Somente na tarefa de criação do diálogo haveria a

possibilidade de avaliar a compreensão oral dos alunos integrada à produção oral. Contudo, em entrevista com a coordenadora e professora Pilar, ela explica que a compreensão oral do aluno, assim como a compreensão de textos escritos, já foi avaliada anteriormente pela prova escrita, não havendo a necessidade de reavaliá-la. Esse posicionamento sugere que a natureza colaborativa da compreensão oral, encontrada na fase II do EPLIS, não é considerada nas provas orais da EEAR. Em termos de efeito retroativo temos, portanto, evidências de que a compreensão oral da fase I do exame exerce influência sobre as provas escritas das 2ª, 3ª e 4ª séries, porém o mesmo não pode ser dito em relação à integração das habilidades de compreensão e produção orais propostas na fase II e as provas orais das 2ª e 3ª séries.

A prova oral da 4ª série é popularmente conhecida pelos alunos e professores como "projeto", pois consiste na elaboração, em grupo, de um pôster com colagens de imagens relacionadas à aviação e ao controle de tráfego aéreo. No dia da prova oral, cada grupo é orientado a descrever oralmente as figuras do pôster e contar uma história relacionando-as. Essa proposta é muito parecida com a parte 4 da fase II do EPLIS. A professora Ana assim a descreve:

> o projeto é assim"(+) a gente pede pra eles descreverem (+) eles montam um pôster (+) eles escolhem uma figura (+) e aí eles precisam DESCREVER aquele aquelas figuras (+) eles podem usar mais de uma figura no pôster (+) eles descrevem e depois contam uma história /.../ essa é a avaliação oral (+) que a gente usa como avaliação oral (+) que ainda não é somativa né (+) não vale nota (Entrevista com Ana, professora na EEAR desde 2013).

Fica evidente a influência do EPLIS, porém o que causa estranhamento é o fato de essa prova oral, em específico, contemplar somente uma parte da fase II do exame, o que pode fazer com que apenas esses aspectos do exame sejam enfatizados pelos professores em sala de aula, o que estreitaria os métodos de instrução. Gimenez (1999) já havia constatado a possibilidade de algumas características do teste se tornarem mais proeminentes do que outras. Outro problema reside no fato de a tarefa de contar uma história ser questionada em termos de representatividade do uso da língua no ambiente de trabalho pelo instrutor João. Por ser controlador de tráfego aéreo, João possui mais conhecimento técnico para julgar o teste em termos de autenticidade e afirma que essa tarefa não faz parte do domínio de uso do controlador, devendo ser substituída.

> acredito que deve ser retirada talvez a tarefa de contar a história (+) eu não vejo motivo eh:: operacional (+) no meu dia a dia (+) no meu órgão operacional (+) não tem porque contar uma história em inglês (+) eu nunca iria contar uma história em inglês pro piloto (Entrevista com o instrutor João, professor na EEAR desde 2016).

Essa discrepância percebida por João já havia sido reportada no processo de validação do EPLIS conduzida por Scaramucci et al (2010). Portanto, o fato de a prova oral da 4ª série abordar somente essas duas tarefas é bastante questionável, pois pode levar professores e instrutores a supervalorizarem uma tarefa que não é tão representativa do uso real da língua e a deixarem em segundo plano tarefas mais adequadas como, por exemplo, o gerenciamento do diálogo entre piloto e controlador, por não estarem contempladas no EPLIS nem no teste oral da 4ª série.

Sobre a visão fragmentada percebida na prova oral das 2ª e 3ª séries, Pilar justifica que foi a forma que a equipe encontrou de aumentar a confiabilidade do teste, caso o desempenho do aluno venha a ser considerado para a nota final na disciplina. Segundo Pilar, como as provas na EEAR carregam consequências graves, os alunos costumam questionar os resultados obtidos e, somente dessa maneira, seria possível contabilizar os acertos e erros. Ou seja, na tentativa de manter um alto nível de confiabilidade dos resultados, a validade dessas avaliações é comprometida. Fica evidente que, apesar do avanço com a inserção de uma prova oral, o cenário escolar, no que se refere ao ensino e à aprendizagem de inglês, ainda é dominado por um paradigma tradicional e estruturalista da avaliação no qual: a maior preocupação é para com a confiabilidade; o objetivo é verificar a aprendizagem por meios de acertos e erros para fins de promoção; a correção deve ser objetiva; o produto é privilegiado; as habilidades são avaliadas separadamente; e as provas são conduzidas como eventos únicos no meio e no final do semestre.

Durante a entrevista, sabendo que o aluno na EEAR é movido pela obtenção de resultados, discuti com Pilar a respeito da prova oral não ser ainda considerada como parte integrante da nota final do aluno. Na sua fala, ela destaca dois pontos: a impossibilidade de manter dois professores por sala e a ausência de aparelhos para a gravação da prova. Ela afirma que os alunos podem contestar o resultado obtido e, sem esses meios, a confiabilidade do teste oral poderia ser colocada em xeque e os professores e instrutores responsabilizados.

> fica MUITO difícil de eu colocar DOIS professores (+) e às vezes eu não tenho dois professores (+) dois avaliadores (+) então, a gente tá rodando a avaliação com um professor só (+) a gente não tem um meio pra gravar essa prova (+) então, eu acredito assim"(+) pra essa prova valer efetivamente (+) eu tenho que ter MEIOS pra eu poder me fiar na hora de mostrar pro aluno que o grau dele tá de acordo com aquilo que foi apontado na:a no:o *report* (+) lá no documento dele (+) e assim' a gente NÃO tem (+) a gente não tem um gravador (+) a gente não tem DOIS avaliadores (+) então, TUDO isso pode vir eh: contra nós (+) professores (+) instrutores (+) sabe, (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Outro ponto que dificulta essa tomada de decisão está relacionado à distribuição de pesos entre as provas de compreensão de textos, compreensão oral e produção oral. Existe o receio de que se os pesos não forem distribuídos adequadamente, o desempenho do aluno na prova oral pode fazer com que haja um número alto de reprovações, dado que essa é a área em que os alunos apresentam mais dificuldade. Assim, explica Pilar:

> tem uma outra questão (+) como fazer essa divisão" (+) então a gente tava pensando (+) eu tinha uma proposta de fazer de como / de quanto valeria a prova de áudio (+) quanto valeria a prova escrita (+) e quanto valeria a prova oral (+) entendeu" (+) fazer uma média aí pra gente poder chegar num momento em que a gente vai conseguir (+) eh: ter um equilíbrio e que a gente vai trazer na ponta (+) no final do curso da disciplina (+) não vai trazer uma redução muito drástica (de alunos aprovados) (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

De fato, não podemos desconsiderar as restrições de tempo e recursos sob as quais os professores e instrutores atuam e que afetam as práticas avaliativas em sala de aula, quanto mais a deficiência na área de letramento que envolve grande parte dos docentes no Brasil. Bachman e Palmer (2010, p. 8) argumentam que "não há um teste perfeito para todas as situações. Para qualquer situação, há apenas alternativas com seus pontos fortes e suas falhas".[67] Todavia, devemos refletir, principalmente, quando um teste carrega uma alta-relevância para alunos e usuários, se esse instrumento é o mais apropriado para a situação e para as necessidades do aprendiz. Caso a resposta não seja afirmativa, mudanças devem obrigatoriamente ser implementadas.

Sobre a dedicação dos alunos à prova oral mesmo sabendo que o desempenho não interfere na nota final, Pilar responde positivamente e justifica dizendo que os alunos veem esse teste como uma possibilidade de se prepararem para o EPLIS. Também é ressaltado, em sua fala, o papel do teste oral no controle das emoções.

> são (+) são (dedicados) (+) eu acho isso bem interessante porque (+) eh (+) impacto do EPLIS também (+) porque eles veem esse momento de interação com o professor como algo de muita importância /.../ eu falo' gente olha' vocês estão tendo a oportunidade de APRENDER acima de TUDO a lidar com as emoções (+) porque nem sempre é o seu professor que vai te avaliar (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

[67] No original: "There is no perfect test for all situations. In any situation there are only alternatives, each with strengths and weaknesses".

Pilar acrescenta que expor os alunos a um exame, com a presença de um professor ou instrutor que nem sempre lhes é familiar, faz com que eles desenvolvam estratégias para lidar com a carga emocional e, assim, se sintam mais confiantes e seguros para realizar o EPLIS.

> então assim, eles têm / são avaliados por vezes por professores que eles nunca tiveram aula (+) então é um momento que ele vai aprender (+) eu falo pra eles'(+) gente, vocês têm que aprender a lidar com essa situação de ser avaliado (+) vocês têm que avaliar qual a melhor forma (+) isso é um pouco de autoconhecimento (+) o que você vai fazer pra não ficar nervoso" (+) pra não deixar que isso influencie na sua fala (+) você vai movimentar as suas mãos" (+) você vai olhar em um ponto fixo na parede" (+) você vai eh:: desviar o olhar do avaliador" (+) então, você tem que criar as estratégias que vão te AJUDAR a te sentir a se sentir mais seguro na hora da avaliação (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Outro ponto mencionado por Pilar refere-se à relação entre a ansiedade e o desempenho do aluno no EPLIS. Apesar de raros, Pilar diz existir casos em que a ansiedade e o *stress* provocados pelo exame prejudicaram o desempenho de alunos que ela considerava que seriam aprovados: "eh: mas existem casos, (+) de alunos que o emocional realmente afeta, né (+) influencia no resultado final dele (+) tem gente que bloqueia, né". Nesse sentido, a prova oral tem também atuado positivamente, segundo Pilar.

De fato, familiarizar o aluno com a situação avaliativa é considerada uma estratégia adequada de preparação, pois ajuda a gerenciar sua ansiedade, reduzindo, assim, a variância irrelevante ao construto, o que propicia uma avaliação mais precisa do desempenho do aluno (MESSICK, 1982; HALADYNA et al, 1991) e essa prática deve ser promovida em sala de aula.

Assim como Pilar, a chefe e professora Isabel também afirma que os alunos encaram a prova oral com seriedade, porque estão preocupados com o EPLIS e desejam ser aprovados no exame, porém admite que, se o desempenho na avaliação escolar fosse considerado para fins de aprovação, os alunos poderiam demostrar ainda mais interesse. Assim, relata:

> se fosse cobrado como sendo né somativo (+) ele ia até dar mais importância na hora de falar (+) mas por causa do EPLIS (+) ele sabe que durante a carreira dele TODA ele vai ter que fazer esse exame /.../ mas ele também não deixa de dar importância porque ele sabe que o Eplis (+) vai vir um avaliador aqui (+) e ele QUER passar (+) ele quer ter nota 4 (+) ele quer passar pelo menos pra segunda fase (+) ele sabe que vai ter alguém questionando ele (Entrevista com Isabel, chefe da Subseção de Idiomas e professora na EEAR desde 2011).

Em suma, existem evidências de que o EPLIS influencia as avaliações na EEAR. A ênfase dada à inclusão de áudios de radiotelefonia (RT), somada à elaboração de itens em

português na prova escrita da 4ª série com o objetivo de familiarizar o aluno com o formato do teste indicam impacto da fase I do exame. A adoção do EPLIS na EEAR também parece ter motivado a elaboração de uma prova oral, no qual encontramos similaridades com a fase II do exame em termos de critérios avaliados e tipos de tarefas. Por outro lado, a integração das habilidades de compreensão e produção orais proposta nessa fase do exame não afetam essa avaliação. Para que os efeitos pretendidos sejam atendidos precisamos também capacitar os professores, oferecendo oportunidades reais de letramento em avaliação para que eles possam construir avaliações mais adequadas às necessidades dos alunos e informadas teoricamente, impedindo, por exemplo, que sejam empregadas estratégias para responder testes que inflam os resultados sem um ganho real na proficiência. Dada a relevância das provas escritas, também é necessário e urgente encontrar meios para que a validade dos exames na EEAR seja aumentada. Além disso, vemos também a necessidade de que seja atribuída nota às provas orais para que elas se tornem tão relevantes quanto as provas escritas já o são, o que elevaria o *status* da produção oral em sala de aula.

6.2.5 A preparação para o exame e sua natureza

Na análise estatística, foi encontrada discordância entre alunos, de um lado, e professores e instrutores, de outro, sobre a adequação da preparação para o EPLIS oferecida na EEAR. Enquanto 93% dos professores e instrutores afirmaram que as aulas de inglês na Escola preparam bem os alunos para o EPLIS, somente 43% dos alunos concordaram com essa alegação. Ademais, a média de respostas dos alunos da 4ª série a essa pergunta foi a mais baixa, o que nos levou a inferir que, por conhecerem melhor o exame, eles identificaram discrepâncias entre os objetivos do ensino e a avaliação externa. No entanto, os comentários ao questionário e as conversas sobre o assunto revelaram crenças importantes desses alunos sobre aprendizagem e avaliação, ligadas ao conceito tradicional de cultura de avaliar (SCARAMUCCI, 2006) que despreza a dimensão educativa da avaliação.

Quando indagados sobre como eles avaliam a preparação na EEAR, os alunos argumentaram que ela poderia ter sido mais direcionada ao EPLIS. No comentário de aluna da 2ª série, turma D, ela deixa transparecer que o desempenho no exame está associado à prática de itens similares em sala de aula. Na sua opinião, quanto mais focado no exame estiver o treinamento, maiores são as chances de aprovação. Assim, relata: “desde o início do curso as aulas deveriam conter mais exercícios semelhantes à avaliação do EPLIS para o aluno se acostumar com o que vai ser cobrado no dia da prova e assim obter um bom resultado”.

A visão de que a preparação no mesmo formato do exame é fundamental para a obtenção de bons resultados é também compartilhada pelos participantes do grupo focal. Na fala de uma aluna da 4ª série, da turma D, fica evidente que, na sua opinião, se os alunos fossem treinados a descrever uma figura desde o princípio, eles se sairiam melhor nessa tarefa: "então não deu tempo assim pra gente absorver (+) como poderia ter absorvido se começasse isso LÁ na primeira série (+) botar a gente para descrever".

A preparação focada, exclusivamente, em itens similares ao exame é considerada prejudicial para o desenvolvimento da competência comunicativa do aluno, pois restringe o currículo ao que está presente no teste, ignorando tudo o que não é avaliado (SMITH, 1991; POPHAM, 1991; HALADYNA; DOWNING, 2004). Além disso, por o exame avaliar somente uma parcela das habilidades contidas no domínio de uso do controlador de tráfego aéreo, a prática com itens dessa amostra não garante que o aluno tenha desenvolvido competência naquilo que não foi avaliado, o que pode ser agravado caso o construto esteja sub-representado no exame.

No entanto, os alunos não compreendem da mesma forma. Provavelmente, influenciados por experiências de avaliação, nas quais macetes e técnicas de memorização eram suficientes para se obter um bom desempenho em uma prova, eles acreditam que para serem aprovados no EPLIS precisam "treinar" à exaustão os modelos de questão encontrados no exame. Essa cultura de avaliar parece ser tão enraizada que faz com que os alunos se tornem impacientes com o uso de materiais que não estejam no formato da avaliação, como pontua Prodomou (1995).

Pressuponho, então, que existe certa pressão por parte dos alunos para que o ensino de inglês na EEAR seja totalmente direcionado ao exame, o chamado *teaching to the test*. A coordenadora e professora Pilar confirma que recebe muitas críticas de alunos que acreditam que as aulas de inglês na EEAR deveriam ser um treinamento para o EPLIS. Em resposta a eles, Pilar ressalta que o objetivo do curso de inglês na EEAR é a preparação para uma comunicação aérea segura, e que o EPLIS é apenas um instrumento que irá atestar se o aluno alcançou o desempenho desejado.

> eu procuro dizer pra eles que assim' (+) nós estamos preocupados com a comunicação aérea (+) então, nosso foco é a comunicação aérea e que o EPLIS será a pontinha lá (+) será uma consequência de todo de toda a performance dele em sala de aula (+) de tudo aquilo que ele foi eh:: (+) exposto durante o curso (+) então, isso eu costumo passar pros alunos (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

Pilar parece bastante consciente das consequências ruins que uma preparação em mesmo formato poderia causar e demonstra sua preocupação em alertar os professores a respeito. No trecho abaixo, ela usa o jargão militar "bizural" para se referir às dicas que os alunos gostariam que fossem dadas para serem aprovados no EPLIS e descreve que sucumbir às suas vontades seria negativo e um "perigo". Pilar reitera que a grande maioria dos professores acredita ser importante abordar o exame em sala de aula, porém sem o carácter de treinamento para o teste. Assim, relata:

> eu acho que o professor NÃO pode ficar (+) o que eu costumo falar' (+) não pode ficar atrelado às etapas do exame, né (+) então (+) assim, (+) a gente não pode CEDER ah:: às críticas, né (+) ceder aos apontamentos que os alunos nos trazem, (+) isso aí é: um PERIGO (+) porque nós não somos um curso que vai ensinar, vai dar o bizural de como fazer o EPLIS (+) a nossa importância é BEM maior do que essa, né (+) então isso seria algo negativo, né (+) se os professores enxergassem dessa forma que eu acredito que assim' (+) não (+) eu acredito que não acreditamos (+) a grande maioria acha que é importante (+) colocar algo (+) trazer algo (+) contribuir com algo que eles vejam mas não ficar TREINANDO assim para o exame (Entrevista com Pilar, coordenadora pedagógica da 2ª, 3ª e 4ª séries e professora na EEAR desde 2005).

O professor Eduardo também foi categórico ao afirmar que o EPLIS não restringe sua prática pedagógica:

> NEGATIVO (+) não (+) como se o EPLIS pudesse limitar" (+) não (+) eu acho que não (+) por exemplo (+) eu particularmente como eu gosto de pronúncia (+) apesar de não trabalhar tudo que eu gostaria de trabalhar (+) eu não penso assim' (+) vou parar aqui porque a partir daqui o EPLIS não cobra mais (+) eu vou até onde eu quero ir (Entrevista com Eduardo, professor na EEAR desde 2010).

No entanto, Pilar se referiu anteriormente à 4ª série como uma espécie de curso preparatório para o teste: "a gente treina para o Eplis veementemente (+) é a única série que está voltada exclusivamente para o exame", levantando a hipótese de que possa haver algum tipo de preparação voltada exclusivamente para o formato do exame nessa série.

Se por um lado a preparação em mesmo formato é considerada prejudicial por restringir o conteúdo e os métodos de instrução em sala de aula (POPHAM, 1991; HALADYNA et al, 1991; HALADYNA; DOWNING, 2004), por outro lado, sabemos que a familiarização dos alunos com a prova e seus itens é vista como pertinente e essencial no processo de efeito retroativo, conforme discutido no Capítulo II. Gebril e Eid (2017), por exemplo, pontuam que conhecer de antemão as características do teste a que serão submetidos oferece aos alunos suporte psicológico para lidar com situações que possam afetar negativamente o desempenho durante o teste, tais como a ansiedade, o *stress* e a insegurança.

Benefícios esses já mencionados pela coordenadora pedagógica e professora Pilar ao comentar sobre o propósito da implementação das provas orais para os alunos. Nesse sentido, a preparação considerada mais adequada seria aquela que promove a familiarização do aluno com as características do exame, oferecendo estratégias para lidar com diferentes tipos de itens, bem como para gerenciar o tempo e a ansiedade, porém sem se limitar aos formatos contemplados no teste.

Em resumo, com base nas percepções dos professores e alunos, podemos supor que nas 2ª e 3ª séries existe a intenção de alguns professores e instrutores de familiarizar os alunos com o formato do exame e os tipos de tarefas cobradas, porém sem se limitarem a isso, promovendo, por exemplo, o desenvolvimento da habilidade de leitura e de atividades não encontradas no teste como a criação de diálogos entre piloto e controlador. Na 4ª série, os relatos já apontam para uma preparação mais direcionada ao EPLIS, com ênfase na descrição de figuras e relato de histórias, o que seria considerado inadequado por estreitar demasiadamente o currículo. Os professores e instrutores parecem cientes de que uma preparação em mesmo formato gera prejuízo para a instrução, porém os alunos acreditam que praticar itens similares melhoraria o desempenho no exame. A natureza da preparação oferecida será discutida a seguir, na análise das observações em sala de aula.

5.3 As ações em sala de aula

Uma das críticas de Alderson e Wall (1993) aos primeiros estudos sobre efeito retroativo residia na carência de dados provenientes de observações em sala de aula que pudessem confirmar os relatos de professores e alunos sobre o impacto do exame. De fato, os dados coletados em sala de aula colaboram para que haja uma visão mais completa do fenômeno, pois possibilitam que as percepções e atitudes dos participantes sejam confrontadas com as suas ações. Para tanto, observei as aulas de três professores e um instrutor, participantes do estudo quantitativo, que haviam se voluntariado a participar dessa etapa da pesquisa e tinham disponibilidade para entrevistas individuais.

Em proporção, eles representam 25% dos que responderam ao questionário. Dos quatro observados, um é controlador de tráfego aéreo. A decisão por observar as aulas de, pelo menos, um instrutor de inglês aeronáutico deveu-se ao fato de esses profissionais terem passado a integrar o corpo docente da EEAR a partir de 2016. A proporção 1 instrutor para 3 professores observados coincide com a proporção de respondentes ao questionário (12 professores e 4 instrutores), ou seja, 25%.

Em relação à experiência, dois professores apresentavam mais de 6 anos de ensino de inglês aeronáutico e foram considerados mais experientes e os outros dois possuíam 5 anos ou menos de docência nessa área e, portanto, menos experientes em relação aos primeiros. Como explicado no Capítulo III, essa divisão foi pautada nas características dos dois grupos emergentes no estudo preliminar: o primeiro com até 2 anos de experiência, e o segundo com 3 anos ou mais. Considerando que, entre a aplicação do questionário preliminar e o questionário principal, se passaram quase 3 anos, o primeiro grupo estaria com até 5 anos de experiência no ensino de inglês aeronáutico e o segundo, 6 ou mais. Busquei, também, sempre que possível, observar as aulas dos professores e do instrutor em turmas mais proficientes e menos proficientes, uma vez que os resultados quantitativos mostraram que o fator proficiência parece interagir com o efeito do exame.

Para direcionamento dos aspectos a serem analisados, utilizei uma ficha de observação de aulas (Apêndice E). Dada a natureza do exame, esperava observar uma abordagem de ensino voltada para o uso da língua e com ênfase na integração das habilidades orais, conteúdos relacionados ao ambiente de trabalho do controlador, desenvolvimento do inglês comum para comunicação aeronáutica, das funções linguísticas e das estratégias necessárias para gerenciamento da comunicação aeronáutica, além de amplas oportunidades de prática da produção oral dos alunos.

A análise está organizada por professor ou instrutor observado, bem como por séries e turmas em que ele atuou no período da geração dos dados. Apresento, primeiramente, um retrato da aula assistida com base na ficha de observação, no qual especifico as atividades realizadas, sua duração, o tipo de interação dominante, as habilidades e aspectos linguísticos desenvolvidos, o conteúdo abordado, o tipo de material utilizado, bem como informações que identificam o professor, a classificação da turma, o número de alunos, a data da observação e os tempos de aula (cada tempo de aula equivale a 45 minutos). Em seguida, faço uma análise da aula, comparando-a com os efeitos pretendidos e as hipóteses levantadas no estudo quantitativo. Uso, também, os comentários surgidos nas entrevistas individuais com os participantes para dar respaldo à análise, triangulando, assim, as diversas fontes de dados.

5.3.1 A professora e coordenadora Pilar

A professora Pilar faz parte do corpo docente de língua inglesa da EEAR desde 2005. Possui graduação em Letras − Inglês e Mestrado em Linguística Aplicada. Participou ativamente da reestruturação das disciplinas de língua inglesa para o curso BCT, ocorrida em

2008, em cumprimento ao Plano de Implementação dos Requisitos de Proficiência em Inglês, como uma de suas mentoras, adotando, por completo, as mudanças e, por essa razão, pode ser considerada uma *adopter* na perspectiva de Burrows (2004). Por ser parte desse processo desde o início, possui 10 anos de experiência no ensino de inglês aeronáutico, sendo uma das professoras mais experientes nessa área na EEAR. Em sua entrevista, Pilar afirma ter adquirido bons conhecimentos sobre inglês aeronáutico durante esses anos e sobre o EPLIS, além de reconhecer a influência do exame em sua prática pedagógica. Pilar também participou da oficina de dois dias, oferecida pelo ICEA em 2015, sobre o exame, e do curso de formação de instrutores de inglês aeronáutico. No ano de 2016, Pilar atuou nas 2ª e 3ª séries do curso BCT, mas também costuma ministrar aulas para a 4ª série. Durante o período da geração dos dados, tive oportunidade de observar aulas da professora Pilar na 2ª série em três turmas (*A*lfa, *B*ravo e *G*olf). As turmas A e B podem ser consideradas mais proficientes e a turma G menos proficiente. Na 3ª série, também foi possível observar aulas em turmas de níveis de proficiência bastante distintos, turmas *B*ravo e *G*olf.

5.3.1.1 Aulas na 2ª série do curso BCT

O Quadro 17 resume a aula ministrada por Pilar na turma Golf. Os aspectos mais salientes dessa aula foram o predomínio da interação professor-classe, ausência de atividades em pares ou em grupos e poucas oportunidades de produção oral, o que contradiz os efeitos pretendidos do EPLIS em termos de abordagem de ensino. A única atividade que demandou dos alunos alguma produção oral foi a recriação de uma história. O fato de essa atividade ter sido proposta pela própria professora demonstra certa influência do exame, dado que na fase II do teste os alunos têm que contar uma história com base em uma imagem. No entanto, o tempo destinado para a sua realização foi bastante curto e o tipo de interação, em que a professora coloca perguntas para a toda a classe (professor-classe), fez com que poucos alunos participassem. Durante a correção, Pilar retomou a história, explicando vocabulário e pontos gramaticais a partir dos erros cometidos pelos alunos. Pilar colocou algumas perguntas à turma que, por várias vezes, foram respondidas por ela mesma. Apesar de ter abordado o vocabulário e os itens gramaticais de forma contextualizada e a partir do que foi produzido por alguns alunos, Pilar dominou toda a discussão.

O padrão professor-classe não foi alterado para as atividades de compreensão oral que se sucederam. A contextualização do áudio foi feita por Pilar, com a colocação ocasional de perguntas que foram respondidas por um grupo pequeno de alunos ou pela própria

professora. Também foi pedido aos alunos que lessem em voz alta todas as questões. A cada questão lida, Pilar oferecia uma explicação sobre o seu conteúdo ou prováveis respostas. A correção das atividades foi iniciada logo após cada áudio ter sido tocado por duas vezes seguidas, não permitindo que os alunos checassem suas respostas entre si, o que seria uma forma de promover alguma interação entre eles. Pilar coletou oralmente as respostas dos alunos, ofereceu mais explicações sobre vocabulário, por algumas vezes tocou novamente partes do áudio e corrigiu erros gramaticais cometidos por eles.

Quadro 17 – Aulas observadas na 2ª série turma G

Professora: Pilar	**Série: 2ª**	**Turma: G**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 22/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
4 min	Recriar uma história baseada na leitura prévia de um texto.	Alunos-classe	Avarias à aeronave / descrever uma sequência de eventos passados	Leitura e produção oral	Escrito / Livro, proposta criada pela professora
19 min	Explicar vocabulário e pontos gramaticais	Professor-classe	Avarias à aeronave	Vocabulário e estrutura	Escrito / Frases ditas pelos alunos
21 min	Marcar verdadeiro ou falso e escolher a resposta correta com base na escuta de uma conversa radiotelefônica	Professor-classe Individual (3 min)	Avarias à aeronave	Compreensão oral	Áudio / Livro, gênero: RT, proposta criada pela professora
36 min	Responder questões abertas com base no áudio, correção da atividade, explicação sobre vocabulário e correção gramatical	Individual (3min) Professor-classe (33min)	Pouso de emergência	Compreensão oral e Vocabulário	Escrito / Livro

O conteúdo abordado em sala de aula foi definido pelo material didático, não havendo, portanto, influência do EPLIS. No entanto, o tópico trabalhado é relevante para o controle de tráfego aéreo e está no domínio do exame, além disso o áudio utilizado pertence ao gênero comunicação aeronáutica. A função comunicativa trabalhada na primeira atividade "descrever uma sequência de eventos passados" também consta no Doc 9835 AN/453. Para a primeira atividade de compreensão oral, Pilar mudou o formato das questões, optando por modelos mais fechados (múltipla escolha e verdadeiro ou falso), o que pode estar relacionado com o baixo nível de proficiência dos alunos ou com o método de avaliação adotado na fase I do EPLIS (múltipla escolha). Essas decisões também revelaram que a professora não se limita ao livro didático, uma vez que propõe atividades em formatos diferentes daqueles encontrados no material, o que também foi constatado nas demais aulas assistidas. No entanto, não houve integração com outras habilidades, e a compreensão oral foi trabalhada isoladamente.

Nessa aula, não foi observada nenhuma menção explícita ao exame, seja pela professora seja pelos alunos. No entanto, Pilar citou a existência de uma lista que havia sido distribuída aos alunos no ano anterior que continha marcadores de discurso e frisou a importância de os alunos conhecerem e utilizarem esses conectores quando estiverem contando uma história, o que evidencia a preocupação da professora para com esse tipo de atividade. Ademais, nas atividades de compreensão oral, Pilar aconselhou os alunos a não tentarem entender tudo o que está sendo dito, mas focarem nas informações necessárias para responder às questões colocadas, o que seria considerada uma estratégia que pode ser aplicada à fase I do EPLIS.

Em resumo, foi possível observar exemplos do que poderia ser considerado efeito retroativo do exame em alguns tipos de atividades realizadas e nas estratégias mencionadas. O fato de essas atividades não estarem no mesmo formato do exame evidencia uma preparação de caráter mais variado. No entanto, houve poucas oportunidades de interação entre os alunos que propiciassem o uso da língua, principalmente, com foco na negociação do significado, e as habilidades orais não foram integradas. A produção oral dos alunos limitou-se a respostas curtas às questões propostas por Pilar ou que estavam no material didático, sendo que poucos participaram realmente. Essa constatação sustenta a percepção dos alunos de que a produção oral é menos trabalhada em turmas com baixa proficiência, sendo a opção de perguntas e respostas entre professor e alunos o mais frequente, o que restringe a participação de todos. Embora haja tarefas no EPLIS em que o interlocutor coloca perguntas ao candidato, espera-se deste um turno mais longo de fala, o que não foi observado em sala de aula, dificultando o estabelecimento de uma relação mais direta com o exame. A aula de Pilar foi dominada por explicações de vocabulário com algumas correções de gramática, indicando uma base mais estruturalista.

Na entrevista que se seguiu, perguntei a Pilar se ela era capaz de perceber alguma influência do exame nessa aula em específico, ela apontou o relato de histórias, o vocabulário ensinado, o emprego de marcadores do discurso, bem como o uso adequado do passado simples, o que, de fato, é avaliado no EPLIS, o que demonstra seus conhecimentos sobre o teste:

> eles não têm que CRIAR uma história (+) eles tem que recontar uma história (+) mas é algo que vai ajudá-los (+) é uma tarefa semelhante (+) se eles usarem os conectivos, isso já ajuda, né (+) na parte de fluência (+) se eles souberem usar empregar os verbos no passado de forma correta (+) eles vão tá pontuando na parte de estrutura, né (+) se eles usarem bem oo: *simple past* /.../ trabalhando os tempos verbais, né (+) trabalhando com o vocabulário específico (+) os marcadores de discurso (+) de certa forma isso vai trazer contribuições pra um bom desempenho dele (Entrevista com a professora Pilar, conduzida em 28 de março de 2017).

Quanto ao fato de os alunos não terem tido muitas oportunidades de prática voltada para a produção oral, Pilar justifica que a escolha por trabalhar com a classe inteira de uma só vez foi fruto da pressão do tempo, mas reconhece que essa decisão comprometeu a participação dos alunos. Assim, relata:

> eu procuro diversificar' (+) eu procuro diversificar (+) mas essa foi uma estratégia que eu usei diante do tempo que eu tinha (+) pra::a realizar a atividade (+) né'(+) pra poder dar conta do conteúdo que eu tinha que dar conta /.../ ficou prejudicado (+) NESSA atividade ficou prejudicado (Entrevista com a professora Pilar, conduzida em 28 de março de 2017).

Em relação às longas explicações sobre vocabulário e alguns itens gramaticais, Pilar fala sobre a sua preocupação em cobrir, em sala de aula, o conteúdo que será exigido na prova escrita, corroborando a hipótese de alta-relevância das avaliações de rendimento para o professor e para o aluno no contexto escolar, que acabam por ditar a ênfase dada em sala de aula.

> não, não (+) mas tem conceitos que SÃO muito importantes e que são cobrados na avaliação, né (+) então, não tem como (+) a gente cobra o que é *burn off* (+) o que que é *dump fuel* (+) a gente cobra o *suggest* da parte da gramática (+) *alternate airdrome* (+) que eles falam *alternative* (+) embora americano usa *alternative airdrome* (+) mas na documentação na nossa documentação a gente usa *alternate airdrome* (+) eles usam muito ao invés de *destination* eles usam *destiny* (Entrevista com a professora Pilar, conduzida em 28 de março de 2017).

Embora Pilar e os demais entrevistados tenham afirmado que as avaliações escritas enfatizam o significado da mensagem, a preocupação externada por Pilar para com o vocabulário e a gramática em sala de aula, no trecho anterior, sugere que as avaliações escritas podem ter uma visão de linguagem mais estruturalista.

Em relação aos alunos mais proficientes da 2ª série, foram observadas aulas nas turmas Alfa e Bravo. Os Quadros 18 e 19 reportam as observações feitas na turma Alfa nos dias 22 e 23 e no dia 29 de março, respectivamente. Nas aulas dos dias 22 e 23 (Quadro 18), foi possível perceber mais oportunidades de prática oral e maior interação entre os alunos. É importante ressaltar aqui que essa mudança não tem qualquer relação com a entrevista que sucede à observação, pois optei por assistir a um bloco de aulas em turmas diferentes antes de termos nosso primeiro encontro. Dessa forma, a primeira entrevista ocorreu somente no dia 28 de março.

Quadro 18 – Aulas observadas na 2ª série turma A – parte 1

Professora: Pilar	**Série: 2ª**	**Turma: A**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 22 e 23/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
10 min	Discutir as características principais da comunicação aeronáutica	Trabalho em pares	Comunicação aeronáutica	Produção oral	Escrito / Proposta criada pela professora
5 min	Correção da atividade, explicação de novos termos, extraindo informações dos alunos e reformulando o que foi dito	Professor-classe	Comunicação aeronáutica	Vocabulário e gramática	Correção oral
30 min	Explicar situações rotineiras de tráfego aéreo a partir da leitura de conversas piloto-controlador e identificar informações desnecessárias	Trabalho em grupo	Comunicação aeronáutica no controle de aproximação / descrever uma comunicação ocorrida previamente	Leitura e produção oral	Escrito / Livro
35 min	Correção da atividade, explicação das situações e explicação sobre vocabulário	Professor-classe	Controle de aproximação	Vocabulário	Correção oral

Duas atividades de produção oral foram desenvolvidas nessas aulas. A primeira tinha como objetivo a discussão em pares, e a segunda integrava leitura e produção oral em um trabalho em grupo, com o propósito comunicativo de descrever uma comunicação ocorrida previamente. O tópico abordado, comunicação aeronáutica, apesar de prescrito pelo livro didático, manteve relação com o exame, bem como a função desenvolvida na segunda atividade de produção oral. Essas duas atividades, discussão e explicação de uma situação de controle de tráfego aéreo foram propostas pela professora e, por serem semelhantes a tarefas da fase II do exame, podem ser interpretadas como evidência de efeito retroativo na sua prática pedagógica. Durante essas atividades, Pilar monitorou os grupos exercendo o papel de facilitadora, o que condiz com a abordagem de ensino comunicativo, recomendada pelo Doc 9835 AN/453 como um dos efeitos pretendidos de exames no contexto aeronáutico.

Quando perguntada sobre a contribuição das atividades desenvolvidas para o desempenho dos alunos no EPLIS, Pilar cita especificamente a atividade de explicar situações de tráfego aéreo prevista na parte 2 da fase II do exame.

> eu acho que na parte de:e eh:: (+) na entrevista oral (+) na parte dois (+) eu acho que o tempo todo (+) eles tentarem dizer qual a situação (+) o que tá acontecendo aqui, né (+) eles têm que fazer uma interpretação do que foi tratado em cada diálogo (+) você já tá fazendo u::m (+) explicando qual a situação (Entrevista com a professora Pilar, conduzida em 28 de março de 2017).

No entanto, da mesma forma que ocorreu com a turma menos proficiente, a correção das atividades demandou tempo considerável de instrução e foi dominada pela professora que explicou novamente as situações discutidas pelos alunos, leu em voz alta alguns trechos dos textos, colocou novas perguntas e, principalmente, explicou o significado de palavras e expressões contidas nos textos. Apesar de os alunos terem participado um pouco mais e produzido turnos mais longos, quando comparados com a turma menos proficiente, o tempo de fala da professora se sobrepôs ao tempo de fala dado aos alunos. Fica evidente a preocupação da professora com que os alunos mostrem clareza no entendimento do vocabulário específico. Ainda que Pilar tenha pedido aos alunos que explicassem esses termos, na maioria das vezes, a explicação é fornecida por ela própria, o que pode estar relacionado ao fato de os alunos nessa série possuírem ainda poucos conhecimentos sobre a área técnica.

Na entrevista, Pilar confirma que a correção tem também como objetivo verificar se os alunos aprenderam o vocabulário e relaciona essa prática à prova oral realizada na 2ª série, no qual o aluno é apresentado a uma palavra e solicitado a explicar oralmente o seu significado. Assim, reporta:

> é checar se eles REALMENTE sabem, entendeu" (+) por exemplo eles acham que uma *jet blast* (+) uma *jet blast* é a mesma coisa que uma *wake turbulence* (+) a diferença de um e de outro (+) porque a gente cobra deles na prova oral também, (+) que eles expliquem CADA palavrinha dessas novas que:e / os vocábulos que aparecem na lição (Entrevista com a professora Pilar, conduzida em 28 de março de 2017).

O EPLIS, como previsto na tabela de níveis da OACI, também avalia vocabulário, principalmente na parte 2 da fase II, quando o interlocutor lê uma situação e pede ao candidato que a explique. Porém, diferentemente da proposta da prova oral da 2ª série, os termos não são apresentados isoladamente, mas como parte de um contexto, mesmo sendo ele curto (sentença).

Outro fato que chama a atenção é a ausência de atividades que promovam o uso do vocabulário ensinado. Pilar explica o significado de novas palavras, porém não cria situações orais em que os alunos podem utilizá-las, o que faz com que o aluno seja capaz de identificar a palavra, porém não há evidências de que ele sabe como usá-la no discurso. Por fim, não houve nessa aula, tampouco nas demais, qualquer menção ao exame, evidenciando que o teste é pouco mencionado e discutido em sala de aula nas primeiras séries do curso.

Na aula assistida no dia 29 de março na turma A (Quadro 19), como podemos perceber, a professora Pilar propôs várias atividades semelhantes àquelas encontradas no EPLIS, porém sem fazer referência ao exame, tais como descrição de uma figura, explicação

de situações de tráfego aéreo e recriação de uma história. Nas atividades de compreensão oral, vimos, novamente, que as questões não seguem o mesmo formato do EPLIS (múltipla escolha), denotando uma preparação para o exame de natureza mais variada.

Quadro 19 − Aulas observadas na 2ª série turma A – parte 2

Professora: Pilar	**Série: 2ª**	**Turma: B**	**No. Alunos: 17**	**Tempos de aula: 03**	**Data: 29/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
6 min	Revisar o conteúdo de um texto e um áudio trabalhado na aula anterior	Professor-classe	Avarias à aeronave; / descrever uma sequência de ações passadas	Produção oral	Proposta criada pela professora
7 min	Descrever uma figura, e responder questões abertas com base em um áudio	Trabalho em pares (3min) Individual (4min)	Avarias à aeronave / prever eventos futuros	Produção oral e Compreensão oral	Visual e escrito / livro, proposta criada pela professora
6 min	Correção da atividade	Trabalho em pares	Avarias à aeronave	Leitura e produção oral	Oral
10 min	Prever eventos futuros, explicação e correção de vocabulário, ler em voz alta as perguntas	Professor-classe	Pouso de emergência	Vocabulário	Oral
3 min	Responder questões abertas com base em um áudio	Individual	Pouso de emergência	Compreensão oral	Áudio e escrito / Livro
16 min	Correção da atividade em pares	Trabalho em pares Professor − classe	Pouso de emergência	Produção oral e vocabulário	Oral
9 min	Sublinhar as ideias principais e recontar o final da história	Trabalho em pares	Pouso de emergência, / reportar uma situação ocorrida	Leitura e produção oral	Escrito / Livro, proposta criada pela professora
4 min	Correção da atividade	Alunos − classe	Pouso de emergência	Produção oral	Oral
6 min	Explicar situações de tráfego aéreo escritas no quadro	Alunos − classe Professor − classe	Pouso de emergência / dar uma explicação sobre um evento passado	Leitura, Produção oral e vocabulário	Escrito, proposta criada pela professora
9 min	Completar as lacunas de uma letra de música	Individual	Sentir-se fora de lugar	Vocabulário	Áudio e escrito, proposta criada pela professora
12 min	Correção, leitura em voz alta de versos da música pelos alunos, explicação de vocabulário	Professor − classe	Sentir-se fora de lugar	Vocabulário	Áudio, proposta criada pela professora
16 min	Revisar o conteúdo aprendido na aula, explicação de vocabulário	Alunos − classe Professor − classe	Avarias à aeronave e pouso de emergência	Vocabulário	Oral

Em termos de interação, houve mais oportunidades de trabalho em pares, o que evidencia que, nas turmas mais proficientes, a professora tende a ser menos controladora. Isso também foi percebido durante a correção das atividades. Diferentemente das aulas anteriores,

Pilar pediu aos alunos que discutissem suas respostas, antes de checá-las com a classe inteira, o que colaborou para minimizar o seu tempo de fala e aumentar a interação entre os alunos.

No entanto, em relação ao ensino de vocabulário, a professora se mostrou ainda como a principal fonte de informações. Embora o vocabulário seja tratado de forma contextualizada, os alunos não apresentam papel ativo no processo de compreensão, o que também prejudica a promoção da aprendizagem autônoma, outro princípio norteador da abordagem comunicativa. Ademais, não houve propostas de situações para uso do vocabulário, o que faz com que o aluno seja capaz de reconhecê-lo, porém não assegura que ele saiba usá-lo.

Na entrevista que sucedeu a essa aula, Pilar afirma ter muita preocupação com o conteúdo apresentado, principalmente, em relação a vocabulário e admite ser bastante controladora, dominando o tempo de fala em sala de aula. Nesse trecho, também é possível perceber o processo de reflexão no qual Pilar se engaja, impulsionado, provavelmente, pela entrevista. Pilar questiona a efetividade de sua aula quando, por exemplo, no final de uma lição, um aluno faz uma pergunta referente a um conceito que já havia sido explicado. Como disse no Capítulo III, essas entrevistas também são uma oportunidade de o professor refletir sobre sua prática pedagógica e ajustá-la se assim julgar conveniente.

> eu sou mesmo (controladora) (+) [o que você é mesmo"] controladora (risos) [na parte de vocabulário"] eu sou (+) eu acho que eu sou muito (+) eu acho que eu falo muito ainda (+) mas também o aluno não tem (+) às vezes ele não tem tempo de usar um dicionário (+) eu fico perturbada às vezes quando o aluno vai chegar lá no fim da lição (+) e ele chega com uma pérola pra mim (+) poxa vida (+) o que é *dump fuel*" /.../ porque aí você fala assim'(+) puxa vida (+) caramba (+) será que a minha prática não tá sendo efetiva" (Entrevista com a professora Pilar, conduzida em 31 de março de 2017).

Outro ponto interessante dessa aula foi a integração das habilidades de leitura e produção oral. O texto funciona como insumo, ponto de partida, para a promoção da interação entre os alunos. Essa escolha é bastante rica, pois demonstra que o exame não restringe a prática da professora. Além disso, confirma os resultados quantitativos de que essa habilidade é bastante recorrente em sala de aula, por influência do livro didático e das avaliações escritas. No entanto, em termos de efeitos pretendidos, esperava-se que houvesse também mais integração entre as habilidades de compreensão e produção orais.

Foram assistidos mais dois tempos de aula da professora Pilar, na turma Bravo, no dia 10 de abril de 2017, antes da condução da última entrevista. O Quadro 20 resume os eventos ocorridos. Podemos observar que a maneira de trabalhar vocabulário da professora foi

modificada em relação às demais aulas assistidas. Nas aulas anteriores, Pilar explicava o significado das novas palavras sem engajar os alunos nesse processo e sem criar situações de uso. Nessa aula, Pilar pediu aos alunos que listassem os verbos mais comuns para dar instruções a aeronaves que estão se movimentando no pátio do aeroporto. Em seguida, ela os orientou a criar diálogos utilizando esses verbos. Para fechar a atividade, Pilar solicitou que um representante de cada grupo lesse em voz alta o seu diálogo suprimindo o verbo de ação, a fim de que os demais alunos adivinhassem pelo contexto fornecido o verbo utilizado. Dessa maneira, a professora foi capaz de integrar escrita, compreensão oral e vocabulário em uma única atividade que refletiu o ambiente de trabalho do controlador.

Quadro 20 − Aulas observadas na 2ª série turma B

Professora: Pilar	**Série: 2ª**	**Turma: B**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 10/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
12 min	Explicar o significado de uma lista de palavras	Professor – classe (7min) Trabalho em pares (5min)	Movimentos no solo	Vocabulário e produção oral	Escrito, proposta criada pela professora
8 min	Listar os verbos utilizados para dar instruções aos pilotos no solo	Professor – classe (3min) Trabalho em grupos (5min)	Movimentos no solo, dar instruções	Vocabulário	Escrito, proposta criada pela professora
7 min	Criar diálogos curtos entre piloto e controlador em uma situação no solo, utilizando um dos verbos listados	Professor – classe (3min) Trabalho em grupos (4min)	Movimentos no solo, dar instruções	Escrita / produção oral e vocabulário	Escrito / Oral, proposta da professora
6 min	Ler em voz alta o diálogo criado suprimindo o verbo de ação e adivinhar o verbo suprimido	Aluno − Alunos	Movimentos no solo, dar instruções	Compreensão oral e vocabulário	Escrito / proposta da professora
7 min	Checagem dos verbos, explicação sobre vocabulário, correção de pronúncia	Professor − classe	Movimentos no solo	Vocabulário	Oral
30 min	Preencher lacunas de conversas entre piloto e controlador com informações do áudio; ler em voz alta as respostas	Individual (10 min) Professor − classe (20 min)	Movimentos no solo, dar instruções	Compreensão oral	Áudio / livro, proposta criada pela professora

A função linguística trabalhada pela professora nessa atividade, “dar instruções”, é uma das mais comuns no controle de tráfego aéreo. O EPLIS, todavia, não avalia essa função diretamente, o que corrobora a percepção de que o exame não limita a prática pedagógica de

Pilar, tampouco o livro didático, uma vez que essa atividade foi elaborada por ela própria, com o objetivo de preparar os alunos para as tarefas de compreensão oral que sucederam.

O áudio utilizado para a tarefa de compreensão oral pertence ao material, porém Pilar modifica o formato da atividade para lacunas a serem preenchidas com informações específicas. Esse tipo de tarefa assemelha-se às exigidas na fase I do exame, que também tem como objetivo a compreensão de informações específicas. Na entrevista, quando perguntada sobre o propósito dessa atividade, ela menciona que seria uma forma de familiarizar os alunos com o exame: "porque:e (+)(+) pra eles se acostumarem também né,", evidenciando efeito retroativo da fase I.

Em relação ao tempo de fala da professora, percebemos uma redução bastante expressiva, que pode ser resultado do processo de reflexão iniciado nas entrevistas. No trecho abaixo, extraído da última entrevista com Pilar, notamos mais um momento de intensa reflexão sobre a sua prática:

> mas eu sou altamente controladora (+) com eles então' (alunos menos proficientes) (+) na lousa o tempo todo (+) gente'(+) vamo aqui (+) vamo lá (+) pronúncia (+) vamo de novo (+) não sei, (+) enfim, (+) a gente aprende (+) a gente aprende que a gente tem que ser (+) que a gente tem que reduzir cada vez mais o *teacher talking time* (+) a gente precisa fazer isso (+) muitas vezes deixo eles falarem (+) mas também' a gente precisa cuidar da base (+) e assim' (+) se a gente não der a base como ele vai produzir" (Entrevista com a professora Pilar, conduzida em 12 de abril de 2017).

Nessa fala, Pilar revelou o conflito vivenciado entre oferecer oportunidades de produção oral para os alunos menos proficientes e consolidar a base linguística (conhecimentos de vocabulário e gramática), como se fossem práticas difíceis de serem conciliadas em sala de aula, o que confirmaria uma visão de ensino ainda de base estruturalista, na qual a proficiência linguística deve ser desenvolvida primeiramente e cujo papel do professor é transmitir conhecimentos aos alunos.

5.3.1.2 Aulas na 3ª série do curso BCT

Na 3ª série do curso, foram assistidas aulas nas turmas Bravo e Golf. Em relação à turma G, menos proficiente, o Quadro 21 apresenta os principais eventos ocorridos no dia 29 de março de 2017. Como podemos observar, houve duas atividades em que os alunos trabalharam em grupos. No entanto, pela baixa proficiência da turma, as atividades que tinham como objetivo engajar os alunos em uma produção oral tornaram-se práticas de escrita e a

interação foi dominada pelo uso do português. Em entrevista com a professora Pilar, ela explicou a dependência desses alunos da escrita como auxiliar para a produção oral.

> isso é uma outra característica dos grupos menos proficientes (+) eles PRECISAM desse TEMPO pra organizar as ideias no papel (+) eu até falei pra eles' (+) *it was a speaking activity* (+) mas eles tem essa necessidade de escrever e mostrar aquilo que eles estão escrevendo (+) porque eles não tem oh:: conhecimento SUFICIENTE da língua pra ter essa produção do *prompt* (+) *promptly* (Entrevista com a professora Pilar, conduzida em 31 de março de 2017).

Quadro 21 − Aulas observadas na 3ª série turma G – parte 1

Professora: Pilar	**Série: 3ª**	**Turma: G**	**No. Alunos: 17**	**Tempos de aula: 02**	**Data: 29/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
16 min	Explicar uma situação a bordo de um avião com base em vídeo sem áudio	Professor − classe (8 min) Trabalho em grupos (8 min)	Problemas com passageiros / descrever uma sequência de eventos passados	Produção oral / Escrita	Vídeo autêntico, proposta criada pela professora
16 min	Apresentação de algumas expressões de opinião, leitura em voz alta da situação, explicação de vocabulário e gramática	Alunos − classe (6min) Professor − classe (10min)	Problemas com passageiros	Vocabulário e gramática	Escrito
34 min	Sublinhar as ideias principais da história e resumi-la	Trabalho em grupo	Problemas com passageiros / descrever uma sequência de eventos passados	Leitura e escrita	Escrito / Livro, proposta criada pela professora
4 min	Recontar a história	Alunos − classe	Problemas com passageiros / descrever uma sequência de eventos passados	Produção oral / leitura em voz alta	Escrito / Oral, proposta criada pela professora

De fato, na última atividade em que os alunos deveriam recontar a história que eles haviam lido e resumido por escrito, uma boa parte deles apresentou uma produção coerente e organizada, sem a necessidade de ler seus resumos. Poucos foram os que os leram do início ao fim, em vez de apresentar com suas próprias palavras a narrativa, o que demonstra a importância de ser dada a oportunidade a esses alunos de organizarem suas ideias no papel antes de expressá-las oralmente. Por outro lado, devido à pressão do tempo, somente um grupo teve a oportunidade de recontar oralmente sua história, uma vez que a parte escrita tomou tempo considerável da instrução. Esse redirecionamento de atividades de produção oral em produção escrita, somado à pressão do tempo, corrobora a percepção dos alunos de que nas turmas menos proficientes a produção oral é mais limitada. Recordo que, na parte 4 da fase II do EPLIS, é permitido que o candidato faça anotações antes de iniciar sua história, porém o tempo dado é

bastante curto, 30 segundos, possibilitando a ele rascunhar somente ideias principais, diferente do que ocorre em sala de aula.

Por fim, os tipos de atividades propostas por Pilar se assemelharam ao EPLIS, confirmando que a professora oferece uma preparação para o exame, apesar de os alunos não terem essa percepção por não conhecerem a fundo suas características. Entretanto, não houve nenhuma menção ao EPLIS, mesmo a um semestre de sua realização. A meu ver, o esclarecimento sobre o teste deve ser feito desde o início do curso, a fim de que o seu nível de relevância aumente entre os alunos.

Na aula do dia 10 de abril (Quadro 22), percebemos, logo na primeira atividade, novamente a produção escrita como suporte para a produção oral, o que fez com que o tempo dedicado ao desenvolvimento da produção oral fosse reduzido. Por exemplo, os alunos levaram 14 minutos para prepararem seus diálogos escritos e somente dois deles tiveram a oportunidade de simulá-lo oralmente. Na entrevista, Pilar ressaltou, mais uma vez, a dependência dos alunos menos proficientes para com a escrita como habilidade auxiliar. Ressalto aqui que o desenvolvimento da escrita não é um objetivo de ensino a ser desenvolvido em sala de aula. Os alunos recorrem à escrita por iniciativa própria, e o professor não interfere, pois reconhece a necessidade dos alunos com baixa proficiência de se apoiarem em algo escrito para a produção oral.

> isso é uma característica dos alunos que têm menos proficiência (+) eles têm essa necessidade de colocar primeiro no papel (+) organizar as ideias (+) porque toda a vez que a gente manda fazer alguma coisa (+) eu não sei se com os outros eles estão fazendo isso (+) mas pelo menos comigo (+) eu tenho visto que isso tem sido MUITO recorrente (+) TUDO eles querem escrever (+) né (+) então, porque eles não se sentem seguros ainda com a língua' (Entrevista com a professora Pilar, conduzida em 12 de abril de 2017).

O gerenciamento do diálogo é uma das principais funções linguísticas no controle de tráfego aéreo e está listada no Doc 9835 AN/453, porém não é avaliado diretamente no EPLIS. O fato de essa atividade fazer parte do repertório da professora Pilar confirma que o exame não engessa sua prática pedagógica. Por outro lado, de todas as aulas assistidas, somente em uma foi observado esse tipo de atividade. Em entrevista, Pilar justifica:

> eu pedi pra todos os professores pra gente TENTAR focar no *roleplay* (+) porque no primeiro bloco a gente não conseguiu (+) a gente ficou muito focado só no recontar as histórias /.../ É (+) então, (+) e aí eu pedi (+) foi um pedido meu (+) gente'(+) vamos focar no *roleplay* (+) vamos tentar dar um *roleplay* (+) pelo menos pra cada grupo (Entrevista com a professora Pilar, conduzida em 12 de abril de 2017).

Nessa fala, Pilar comenta que a tarefa de recontar histórias é mais privilegiada em sala de aula pelos professores e instrutores comparada à simulação de diálogos. Pilar também reforça a importância de se trabalhar, em sala de aula, a simulação de diálogos, pelo menos, uma vez antes da prova oral, já que essa tarefa está nele contemplada e os alunos costumam apresentar dificuldade em sua execução. A falta de entendimento em como realizá-la foi também reportada pelos alunos no grupo focal. No entanto, abordar esse tipo de tarefa uma única vez me parece insuficiente para que os alunos desenvolvam a habilidade necessária para saber lidar com sua complexidade, uma vez que o aluno necessita entender as condições de produção que envolvem o diálogo, ou seja, quem fala com quem (aeronave e órgão de controle), qual o propósito comunicativo envolvido, qual a razão do estabelecimento da comunicação e como ela é feita.

Quadro 22 − Aulas observadas na 3ª série turma G – parte 2

Professora: Pilar	**Série: 3ª**	**Turma: G**	**No. Alunos: 17**	**Tempos de aula: 01**	**Data: 10/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e áreas linguísticas**	**Tipo de material**
14 min	Criar um diálogo entre piloto e controlador com base em um diagrama	Trabalho em pares	Problemas com passageiros / gerenciar o diálogo	Escrita / Produção oral	Visual e escrito, proposta criada pela professora
2 min	Simular um diálogo entre piloto e controlador com base em um diagrama	Alunos – classe	Problemas com passageiros	Vocabulário e gramática	Visual e escrito / livro
8 min	Explicação sobre os usos dos modais para expressar possibilidade e probabilidade	Professor − classe	Problemas com passageiros	Estrutura (is likely to, might or may)	Escrito
10 min	Analisar as situações e expressar a probabilidade de algo acontecer	Individual	Problemas com passageiros / prever um evento futuro	Escrita e gramática	Escrito / Livro

Ainda sobre a criação de diálogos orais, Pilar também argumenta a favor do uso dessa tarefa como meio de utilizar todo o vocabulário aprendido na lição. No trecho abaixo, ela mostra-se consciente de que nada adianta o aluno saber o significado de determinado termo sem ser capaz de usá-lo, e reconhece que esse tem sido um ponto fraco das aulas, como já explicitado.

> é que assim' (+) o que eu costumo falar (+) não adianta eu encher a lousa aqui (+) entendeu" (+) a gente construir um monte de coisa (+) saber o vocabulário e não saber usar'(+) então, pra todo final de lição você tem que ser capaz DE (+) usar aquelas palavrinhas aqueles jargões que você aprendeu (+) no final'(+) você vai REALMENTE ver verificar se você conseguiu aprender aquilo (+) né, (+) o que foi (+) o que foi trabalhado na lição (+) então, eu acho assim'(+) toda lição a gente peca

> por isso (+) acho que toda lição deveria que ter no final (+) quer dizer (+) no final ou então toda aula (+) não sei (+) enfim, (+) a gente teria que eh::: pelo menos fechar com a situação (+) se estudou *air rage* (+) então, a gente vai ter que pegar uma situação de emergência e ter uma *air rage* (+) eles vão ter que tentar fazer um diálogo (+) né (+) pra poder como como fechar (+) usando tudo o que eles aprenderam né, (Entrevista com a professora Pilar, conduzida em 12 de abril de 2017).

Adianto aqui que, embora a simulação de diálogos entre piloto e controlador seja cobrada na prova oral da EEAR e existam propostas dessa natureza nos livros didáticos das 2ª, 3ª e 4ª séries, essa tarefa foi observada em somente mais uma ocasião. Podemos atribuir essa pouca ocorrência ou ausência em sala de aula ao fato de a prova oral não alcançar um nível alto de relevância no contexto escolar, já que o desempenho obtido não é considerado para a nota final da disciplina. Outra causa decorre da possibilidade de os professores e instrutores terem assumido que, por não estar contemplado no EPLIS, esse tipo de tarefa não é importante. De qualquer forma, isso representaria um impacto negativo, uma vez que o desenvolvimento de estratégias comunicativas como, por exemplo, as estratégias de negociação de sentido para o gerenciamento de situações de tráfego aéreo, é considerado um dos efeitos pretendidos da política linguística da OACI.

A última atividade conduzida por Pilar na turma G, analisar situações a bordo da aeronave e fazer previsões, foi também observada por mim na turma B no dia 10 de abril de 2017 (Quadro 23).

Quadro 23 − Aulas observadas na 3ª série turma B – parte 2

Professora: Pilar	**Série: 3ª**	**Turma: B**	**No. Alunos: 16**	**Tempos de aula: 01**	**Data: 10/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
3 min	Explicação sobre os usos dos modais para expressar possibilidade e probabilidade	Professor − classe	Problemas com passageiros	Estrutura (is likely to, might or may)	Escrito, proposta criada pela professora
20 min	Analisar as situações e expressar a probabilidade de algo acontecer	Trabalho em pares	Problemas com passageiros / prever um evento futuro	Produção oral / Escrita e gramática	Escrito / Livro
6 min	Correção da atividade,	Alunos − classe Professor-classe	Problemas com passageiros	Leitura em voz alta/ gramática e vocabulário.	Oral

Enquanto, na turma G, os alunos trabalharam individualmente, por 10 minutos, em uma atividade de produção escrita, os alunos da turma B interagiram em pares, por 20 minutos, em inglês, integrando as habilidades de produção oral e escrita. Isso evidencia que o professor, em função da proficiência dos alunos, modifica o propósito (habilidade desenvolvida) e o

padrão de interação das atividades, o que faz com que haja mais oportunidades de uso da língua nas turmas mais proficientes.

Nas demais aulas observadas na turma B (Quadro 24), notamos ênfase nas habilidades de compreensão e produção orais. Diversos momentos de interação entre os alunos foram criados pela própria professora, demonstrando sua independência em relação ao livro didático. No entanto, o ensino de vocabulário é ainda bastante dominado por Pilar. Antes das atividades de compreensão oral, por exemplo, Pilar fez uma explanação sobre o funcionamento do sistema hidráulico de uma aeronave, introduzindo diversas palavras relacionadas às partes da aeronave e seus movimentos. Apesar de Pilar colocar algumas perguntas aos alunos, poucos participaram e, quando o fizeram, os turnos de fala eram bastante curtos.

Como dito anteriormente, Pilar é uma das professoras mais experientes da EEAR e participou ativamente da implementação do ensino de inglês para fins específicos no curso BCT. Ela possui níveis altos de conhecimentos sobre inglês aeronáutico e sobre o exame, o que pôde ser verificado nas aulas observadas e nas entrevistas. Em suas aulas nas 2ª e 3ª séries, ainda que o exame não tenha sido mencionado, há evidências de efeito retroativo, porém em graus diferenciados em função da proficiência dos alunos. Nas turmas mais proficientes, atividades de produção oral semelhantes às do exame são mais recorrentes, há mais interação entre os alunos e a professora adota uma postura menos controladora. Já, nas turmas menos proficientes, o impacto do teste na abordagem de ensino parece menor, com pouca interação entre os alunos, turnos de fala mais curtos e muita ênfase no ensino de vocabulário. O tempo de produção oral, nessas turmas, é sempre reduzido, devido à necessidade dos alunos de desenvolverem a atividade, primeiramente, na forma escrita.

Em termos de preparação para o exame, podemos dizer que Pilar oferece uma preparação bastante variada nas turmas mais proficientes: as atividades assemelham-se às do exame, porém não seguem o mesmo formato. Além disso, atividades de escrita e leitura foram trabalhadas como suporte para a produção oral, assim como o desenvolvimento de funções linguísticas não cobertas no exame, afastando a possibilidade de estreitamento do currículo como efeito negativo. Por outro lado, a integração entre as habilidades de compreensão e produção orais foi pouquíssimo recorrente. Normalmente, o que se vê é o desenvolvimento da compreensão oral não colaborativa, o que eu atribuo a um forte impacto da fase I do EPLIS. Somado a isso, foi observada apenas uma ocorrência de atividades que simulam a comunicação entre piloto e controlador na 3ª série, embora esse tipo de tarefa esteja contemplado no livro didático. Relaciono esse fato à deficiência do exame em não abordar esse tipo de tarefa tão típica do uso da língua no ambiente profissional.

Quadro 24 − Aulas observadas na 3ª série turma B – parte 1

Professora: Pilar	**Série: 3ª**	**Turma: B**	**No. Alunos: 15**	**Tempos de aula: 03**	**Data: 27/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
3 min	Revisão do assunto tratado na aula anterior, explicação sobre item gramatical	Professor − classe	Tecnologia	Estrutura	Oral
16 min	Discutir tipos de tecnologia aplicadas à aviação e ao controle de tráfego aéreo	Professor − classe (4 min) Trabalho em grupo (12 min)	Tecnologia	Produção oral	Visual, proposta criada pela professora
13 min	Fazer uma apresentação oral sobre tecnologias aplicada à aviação e ao controle de tráfego aéreo	Alunos − classe (8min) Professor − classe (5min)	Tecnologia	Produção oral	Oral, proposta criada pela professora
10 min	Tomar nota das principais diferenças entre os modelos Boeing e Airbus	Professor – classe (3min) Individual (7min)	Tecnologia	Compreensão oral	Escrito e Áudio / livro, Gênero: conversa entre comprador e vendedor
5 min	Correção da atividade, explicação do áudio, explicação de vocabulário e pontos gramaticais	Professor − classe (5min)	Tecnologia	Vocabulário, Estrutura	Oral
9 min	Completar lacunas com palavras do áudio	Individual (5 min) Professor − classe (4min)	Tecnologia	Vocabulário	Áudio e escrito / Livro
15 min	Explicação sobre sistema hidráulico, leitura em voz alta,	Professor − classe (15min)	Avarias à aeronave (sistema hidráulico)	Vocabulário	Visual, proposta criada pela professora
5 min	Circular as palavras corretas para completar o resumo de um áudio	Individual (3min) Professor − classe (2min)	Avarias à aeronave (sistema hidráulico)	Compreensão oral	Escrito, áudio / livro, Gênero: RT
9min	Explicar situações de tráfego aéreo retiradas do áudio	Trabalho em grupo (4min) Professor − classe (5min)	Avarias à aeronave / descrever um estado	Produção oral	Escrito, proposta criada pela professora

A leitura em voz alta foi constatada em sua prática pedagógica em diversos momentos. Pilar, por exemplo, lê em voz alta o enunciado dos exercícios, trechos do livro didático ou do material preparado por ela, com o objetivo de tecer explicações sobre vocabulário para os alunos ou engajá-los em uma atividade no padrão professor-classe. Em relação aos alunos, aqueles menos proficientes costumam ler em voz alta o que eles produziram na forma escrita, em vez de tentarem se expressar espontaneamente sem recorrer às anotações.

Acredito que, por essa razão, os alunos tenham avaliado no questionário a leitura em voz alta como uma atividade bastante presente em sala de aula.

5.3.2 O professor Eduardo

O professor Eduardo passou a integrar a equipe de professores de língua inglesa da EEAR no ano de 2010. É formado em Letras – Inglês e, na época da coleta de dados, estava finalizando seu curso de Mestrado em Linguística Aplicada. Atua no ensino de inglês há 22 anos, sendo os últimos sete dedicados ao ensino de inglês aeronáutico. Em relação aos seus conhecimentos em inglês aeronáutico, Eduardo relata que foram sendo construídos ao longo dos anos por meio da experiência e do contato com a coordenação e colegas de trabalho. Na opinião dele, esses conhecimentos ainda estão em processo. Em relação ao EPLIS, Eduardo reportou que, quando entrou na EEAR, em 2010, não recebeu nenhum treinamento específico, o que veio a ocorrer somente em 2015, quando participou integralmente da oficina oferecida pelo ICEA em conjunto com a EEAR. Quanto à influência do exame em sua prática pedagógica, o professor afirma que o exame direciona suas aulas, porém gostaria de conhecê-lo melhor. No ano de 2016, assim como Pilar, Eduardo atuou nas 2ª e 3ª séries do curso BCT, mas também costuma ministrar aulas para a 4ª série. Tive oportunidade de observar suas aulas nas turmas *C*harlie e *D*elta, da 2ª série, e nas turmas *D*elta, *E*co, *F*ox da 3ª série. Infelizmente, não foi possível assistir aulas em turmas de níveis de proficiência bastante distintos, como ocorreu com Pilar, em razão da dificuldade de conciliar o período de observação ao planejamento fixo de aulas. As turmas *C* e *D* são consideradas medianas em termos linguísticos, enquanto as turmas *E* e *F* são menos proficientes. Normalmente, os alunos são distribuídos em oito turmas, de *A*lfa a *G*olf.

5.2.3.1 As aulas na 2ª série do curso BCT

O Quadro 25 resume a aula observada na turma D no dia 28 de março. Essa aula foi marcada pela ênfase no vocabulário e na pronúncia, leitura em voz alta de trechos por parte dos alunos, baixa interação entre eles e poucas oportunidades de produção oral, exceto turnos curtos quando alguns alunos respondiam perguntas colocadas pelo professor. O padrão de interação dominante foi professor-classe, o que elevou o tempo de fala de Eduardo. Os alunos tiveram a oportunidade de interagir somente uma vez, quando o professor solicitou que eles escrevessem uma notícia de jornal. Durante a correção dessa atividade, os alunos foram

orientados a ler o que haviam escrito, em vez de reportarem com suas próprias palavras, o que fez com que não houvesse nenhuma atividade com foco no desenvolvimento da produção oral.

Quadro 25 − Aulas observadas na 2ª série turma D

Professor: Eduardo	**Série: 2ª**	**Turma: D**	**No. Alunos: 17**	**Tempos de aula: 03**	**Data: 28/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
37 min	Explicação de vocabulário e leitura em voz alta	Professor – classe (37min)	Problemas durante o pouso	Vocabulário e leitura em voz alta	Escrito / livro Proposta criada pelo professor
07 min	Completar sentenças com palavras destacadas no texto	Individual (7min)	Problemas durante o pouso	Vocabulário	Escrito / livro
19 min	Criar uma notícia de jornal	Professor – classe (9min) Trabalho em pares (10min)	Problemas durante o pouso / descrever eventos passados	Produção escrita e estrutura (*reported speech*)	Escrito, proposta criada pelo professor
22 min	Categorizar uma seleção de palavras de acordo com seu padrão de entonação (*word stress*)	Professor – classe (16min) Individual (6min)	Problemas com passageiros	Pronúncia	Escrito / livro
20 min	Categorizar os verbos regulares, presentes no texto, de acordo com a pronúncia do -ED	Professor – classe (20min)	Problemas com passageiros	Pronúncia	Escrito, proposta criada pelo professor

Esses aspectos contradizem os efeitos esperados do exame em termos de ênfase na integração das habilidades de compreensão e produção oral, na oferta de amplas oportunidades para a prática da produção oral e no ensino voltado para o uso da língua. Na primeira atividade, por exemplo, o professor Eduardo solicitou aos alunos que lessem em voz alta trechos de um texto proposto no material didático e de uma apresentação oral elaborada por ele. Após a leitura de cada trecho, o professor fazia ponderações sobre o vocabulário abordado e, algumas vezes, corrigia a pronúncia de certas palavras. Ainda que, durante esse momento, Eduardo tenha colocado perguntas aos alunos, as respostas eram curtas, além de não envolverem todos. Na maior parte do tempo, as explicações sobre vocabulário foram dadas pelo próprio professor. O modo como ele ensina o vocabulário assemelhou-se muito ao de Pilar, com uma postura mais centralizadora e domínio do turno de fala, além da leitura em voz alta como forma de engajar os alunos em uma atividade professor-classe.

Na entrevista, perguntei a Eduardo a razão que o levou a trabalhar o vocabulário dessa forma, ou seja, com os alunos lendo os trechos em voz alta e ele, como professor, explicando as palavras desconhecidas. Eduardo menciona, bem como Pilar, sua preocupação em assegurar que o vocabulário foi coberto em sala de aula e sua percepção de que os alunos

da 2ª série não são capazes de lidar com os termos técnicos da área de controle de tráfego aéreo, como eu havia assumido na análise da aula de Pilar.

> nunca é um motivo SÓ né, (+) talvez seja parte minha (+) o controle meu (+) de ter certeza que eu estou passando a informação (+) tipo, nossa, eu COBRI (+) digamos' doze palavras (+) nossa, então tá, todas as doze palavras foram explicadas pros alunos (+) talvez seja pelo fato de ser 2ª série (+) eu ACHAR que eles não vão conseguir explicar né (Entrevista com o professor Eduardo, conduzida no dia 30 de março de 2017).

Quando perguntei se havia alguma ligação entre essa aula e o EPLIS, Eduardo respondeu que apenas em relação ao vocabulário ensinado, no sentido de que o aluno poderia utilizar o que foi aprendido no exame. No entanto, na apresentação elaborada pelo professor para a apresentação do vocabulário, percebi que alguns exemplos não contemplavam o uso dos termos no contexto de controle de tráfego aéreo, mas, sim, em contextos gerais, o que pode ser indício de lacunas nos conhecimentos sobre a natureza da linguagem na comunicação aeronáutica. É possível, portanto, perceber que o ensino de vocabulário ganha muita importância em sala de aula, principalmente, na 2ª série do curso. Porém, parece se restringir ao reconhecimento do significado da palavra, uma vez que as oportunidades de uso são limitadas.

O Quadro 26 reporta os principais eventos da aula assistida no dia 31 de março, na 2ª série, na turma C. Nessa aula, as atividades de produção oral propostas pelo professor se tornaram práticas de escrita e, posteriormente, de leitura em voz alta diante da baixa proficiência dos alunos. Por exemplo, na atividade em que o professor solicitou aos alunos que descrevessem, em pares, seus desenhos, antes de apresentarem para toda a turma, grande parte dos alunos optou por escrever a descrição, o que culminou em uma baixa interação entre eles. Durante a apresentação oral, a maioria dos alunos leu em voz alta o que haviam escrito, mesmo tendo Eduardo solicitado que eles não recorressem à leitura. Dessa maneira, a percepção dos alunos de que há poucas oportunidades de produção em sala de aula, principalmente, nas turmas menos proficientes, também decorre da dificuldade de se expressarem oralmente, o que os tornam mais dependentes da escrita e da leitura em voz alta, como foi também observado na sala de aula de Pilar.

Em termos de interação, vimos que a aula foi dominada pelo trabalho individual. Houve apenas uma proposta de atividade em que os alunos interagiram em pares. Na entrevista, o professor Eduardo explica que, nas turmas menos proficientes, os alunos acabam conversando em português quando são colocados em pares ou em grupos. Por isso, ele prefere desenvolver

as atividades no padrão professor-classe ou aluno-classe. Já, nas turmas mais proficientes, atividades em pares ou grupos são mais fáceis de serem controladas, porque os alunos não necessitam recorrer à língua materna todo o tempo. De qualquer forma, Eduardo mostra-se consciente da importância de promover atividades mais interativas também nas turmas menos proficientes. Assim, relata:

> porque não quer dizer que só deveria fazer esse tipo de exercício com os proficientes (+) eu acho que os menos proficientes também merecem esse tipo de exercício (+) e outra' quando você dá chance deles conversarem entre si e não falar alto com o professor (+) começar a conversar com o professor em voz alta na frente de todo mundo né (+) é melhor quando eles têm a chance pra conversar com o coleguinha (+) vamos supor uma *describe the picture* (+) *describe the picture in pairs and then* né (Entrevista com o professor Eduardo, conduzida no dia 10 de abril de 2017).

Quadro 26 − Aulas observadas na 2ª série turma C

Professor: Eduardo	**Série: 2ª**	**Turma: C**	**No. Alunos: 20**	**Tempos de aula: 03**	**Data: 31/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
20 min	Desenhar um *layout* de um aeroporto, utilizando o vocabulário aprendido na aula anterior	Individual	Instalações no aeroporto		Visual, proposta criada pelo professor
4 min	Apresentação de expressões para a descrição de lugares: *This part of the airport is called / named / known as ...* *This is where...*	Professor – classe	Instalações no aeroporto /	Expressões para descrever lugares no aeroporto	Escrito, proposta criada pelo professor
8 min	Descrever oralmente o desenho do aeroporto para o colega	Trabalho em pares	Instalações no aeroporto / descrever um estado	Produção oral e escrita	Oral, proposta criada pelo próprio professor
41 min	Apresentar oralmente o desenho do aeroporto	Aluno – classe (individual)	Instalações no aeroporto / descrever um estado	Produção oral e leitura em voz alta	Oral / Escrito, proposta criada pelo professor

Outro ponto interessante da aula de Eduardo foi o fato de as atividades terem sido propostas por ele, desvencilhando-se do livro didático, assim como ocorreu com Pilar. Perguntei a Eduardo qual era a sua intenção ao pedir que os alunos fizessem uma apresentação oral sobre o desenho do aeroporto. O professor responde que seria o desenvolvimento da habilidade de produção oral e acrescenta:

> ajuda na prova oral (+) ajuda no Eplis (+) ajuda (+) *hopefully for the rest of their lives* (+) tinha que ser o foco principal nosso né (+) assim, deixar os alunos falarem (+) o ideal seria eles explicarem tudo (+) eles falarem (+) eles debaterem (+) mas aí, é uma situação utópica né, (+) não aconteceria (+) talvez em uma sala onde os alunos

> falassem idiomas diferentes (+) tivessem eh::: a 1ª língua diferente (Entrevista com o professor Eduardo, conduzida no dia 10 de abril de 2017).

Nesse trecho, temos indícios de que o EPLIS, apesar de não exercer influência sobre os tópicos abordados, impacta a prática do professor Eduardo ao fazer com que ele proponha atividades diferentes daquelas encontradas no material e mais semelhantes às do exame, como, por exemplo, a descrição de imagens. Além disso, Eduardo ensina aos alunos expressões que podem ser empregadas por eles nessa tarefa. Quando indagado sobre qual a contribuição dessas expressões, o professor é direto: "no *describe the picture* né, (+) pensando numa situação na carreira (+) será" (+) se eles precisarem descrever alguma coisa (+) ou se o piloto falar com eles descrevendo alguma coisa né". Também é possível notar, na fala de Eduardo, o processo reflexivo vivenciado por ele, em que ele argumenta que o objetivo principal deveria ser promover oportunidades para que os alunos pudessem interagir na língua alvo, mas justifica que as características do contexto impedem que isso ocorra como, por exemplo, a baixa proficiência de parte dos alunos.

Foi assistida mais uma aula do professor Eduardo na 2ª série, na turma C, no dia 11 de abril (Quadro 27). Em comparação com a primeira aula observada, podemos perceber que houve ganho real nas oportunidades de produção oral. Na primeira atividade, o professor integrou a leitura à produção oral ao solicitar que os alunos identificassem o percurso tomado pela aeronave no texto. Durante sua execução, Eduardo monitorou os alunos, oferecendo ajuda e os encorajando a interagirem em inglês. Em seguida, pediu que alguns alunos recontassem o que havia acontecido com a aeronave utilizando o diagrama desenhado por ele no quadro. Ao perceber que os alunos cometeram muitos erros no uso do passado simples, Eduardo pediu a eles que listassem, oralmente, todos os verbos utilizados para recontar a história e relembrou a forma do passado simples de cada verbo. No entanto, ressalvo que muitos erros cometidos pelos alunos no uso de estruturas gramaticais básicas, durante suas apresentações, não foram tratados pelo professor, mesmo sendo considerados erros que comprometem o entendimento do ouvinte como, por exemplo, o uso do pronome possessivo *your* em vez de *his*, *her* ou *its* e do verbo *have* no lugar de *there to be*.

A ausência de correção de erros considerados globais, ou seja, que prejudicam o entendimento, foi também observada na prática de outros professores e instrutor. Atribuo essa postura à percepção de que na abordagem comunicativa os aspectos gramaticais são aprendidos sem a interferência ou controle do professor. De fato, a versão forte da abordagem comunicativa prega que a competência gramatical se desenvolve por si só, a partir da exposição a amostras compreensíveis e autênticas da língua (SOUZA, 2005; VIEIRA-ABRAHÃO, 2015). Contudo,

em sua versão fraca, a abordagem comunicativa reconhece que alguns aspectos da língua não são adquiridos caso não haja uma instrução formal (LONG, 1991). Na escala de níveis da OACI, em termos de domínio da estrutura, a presença de erros que prejudicam o entendimento da informação é uma das características que descrevem o nível pré-operacional (nível 3). Assim sendo, acredito que os erros globais devem ser tratados pelos professores e instrutores em sala de aula, principalmente, quando notada a sua persistência.

Quadro 27 − Aulas observadas na 2ª série turma C

Professor: Eduardo	**Série: 2ª**	**Turma: C**	**No. Alunos: 17**	**Tempos de aula: 02**	**Data: 11/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
15 min	Identificar o percurso da aeronave no aeroporto	Trabalho em pares	Incidentes no solo / descrever eventos passados	Leitura e produção oral (integradas)	Visual / Livro
20 min	Recontar o incidente	Alunos − classe	Incidentes no solo / descrever eventos passados	Produção oral	Oral, proposta criada pelo professor
7 min	Correção da atividade, explicação sobre o passado simples	Professor − classe	Incidentes no solo	Estrutura	Oral, proposta criada pelo professor
5 min	Relacionar figuras com palavras ou expressões	Individual	Veículos no aeroporto	Vocabulário	Visual, escrito / livro
13 min	Correção da atividade, descrever as figuras	Alunos − classe Professor − classe	Veículos no aeroporto / descrever eventos em progresso	Vocabulário e produção oral	Oral, proposta criada pelo professor

O modo como Eduardo abordou o ensino de vocabulário também foi distinto da primeira aula observada, quando solicitou aos alunos que lessem em voz alta trechos do texto que foram permeados de explicações de vocabulário. Dessa vez, o professor pediu que os alunos relacionassem figuras e palavras ou expressões e incentivou o uso do dicionário. Durante a correção, em vez de os alunos apenas fornecerem a resposta, Eduardo pediu a eles que descrevessem a imagem e fez perguntas para que pudessem elaborar mais o discurso. Isso proporcionou, além de turnos mais longos de fala, a prática do vocabulário aprendido dentro de um contexto de uso, ou seja, descrever aos pilotos o que se passava no aeroporto.

Em nenhuma das aulas assistidas do professor Eduardo na 2ª série, houve menção direta ao exame, embora haja certa influência do exame nas propostas das atividades elaboradas pelo professor. O mesmo ocorreu com a professora Pilar, o que deixa evidente que os professores preferem não discutir o exame nas séries iniciais, provavelmente, para não causar pressão ou ansiedade nos alunos, como foi reportado por Pilar durante as entrevistas. No

entanto, a meu ver, essa decisão diminui a relevância do exame entre os alunos e pode retardar os efeitos pretendidos.

Outro ponto das aulas observadas de Eduardo foi a ausência de atividades de compreensão oral. Questionei o professor sobre isso na entrevista e ele alegou que essa habilidade costuma ser a mais desenvolvida em sala de aula, apesar de acreditar que deveria ser dada mais ênfase na produção oral: "mais trabalhada" (+) como eu falei deveria ser *speaking* (+) mas não é o *speaking* (+) eu acho que é mais *listening* (+) os professores falam mais né". Então, por uma questão de planejamento, as aulas de Eduardo observadas por mim não incluíram o desenvolvimento da compreensão oral.

Também não houve muita diversidade nos propósitos comunicativos das atividades desenvolvidas nas aulas de Eduardo, restringindo-se a descrição de estados, eventos em progresso e eventos passados. Embora essas funções constem no Doc 9835 AN/453 e estejam contempladas no EPLIS, o que demonstra efeito retroativo, esperava um repertório mais extenso.

Na última entrevista conduzida, Eduardo falou sobre a contribuição da observação de aula e da entrevista para o seu processo de reflexão e crescimento profissional e lamenta o papel meramente administrativo da coordenação pedagógica do curso. Nesta fala, podemos depreender a importância do acompanhamento pedagógico para que mudanças, de fato, ocorram, ou seja, para que os efeitos pretendidos do exame encontrem solo favorável para florescerem.

> eu tô achando muito legal você assistir as aulas (+) a gente precisava de uma coordenação assim, /.../ eu tô aqui há sete anos e não fiz curso NENHUM (+) fiz o mestrado né, (+) tô fazendo o mestrado tô terminando (+) mas assim' (+) aqui dentro não me é oferecido NADA (+) assim'(+) não fiz o ctp10 (curso de inglês aeronáutico) e o ctp11 (curso de formação de instrutores de inglês aeronáutico) (+) e nem fiz treinamentos aqui (+) digo assim'(+) lógico, vocês (Icea) vieram aqui (+) vocês deram treinamento sobre o Eplis e tal (+) mas não é um acompanhamento tipo em sala de aula (+) específico para técnicas de ensino e aprendizagem, (+) essas coisas assim (Entrevista com o professor Eduardo, conduzida no dia 12 de abril).

5.2.3.2 As aulas na 3ª série do curso BCT

A primeira observação realizada na 3ª série ocorreu no dia 22 de março na turma F. O Quadro 28 resume os principais eventos ocorridos. A turma F apresenta baixo nível de proficiência e, segundo o professor Eduardo, é um grupo pouco participativo por causa da dificuldade de se expressarem oralmente na língua. No entanto, diferentemente da primeira aula

assistida na 2ª série, o professor criou mais oportunidades para a prática da produção oral e os alunos interagiram por mais tempo.

Quadro 28 − Aulas observadas na 3ª série turma F

Professor: Eduardo	**Série: 3ª**	**Turma: F**	**No. Alunos: 17**	**Tempos de aula: 02**	**Data: 22/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
25min	Recontar uma história baseada em um texto	Trabalho em grupo (13min) Alunos − classe (8min)	Problemas com passageiros / descrever eventos passados	Leitura e produção oral / escrita	Escrito / Livro, proposta criada pelo professor
10 min	Responder, oralmente, questões colocadas pelo professor	Professor − classe	Problemas com passageiros	Produção oral	Vídeos, proposta criada pelo professor
8 min	Responder, por escrito, perguntas sobre os textos lidos	Individual (8min)	Problemas com passageiros	Leitura	Escrito / Livro
10 min	Correção da atividade, checagem de pronúncia, vocabulário e estrutura	Professor − classe	Problemas com passageiros	Pronúncia, vocabulário e estrutura	Oral

Na primeira atividade em grupo, o professor atuou como facilitador, encorajando os alunos a utilizarem a língua inglesa e a checarem no dicionário o significado de palavras desconhecidas, com o propósito comunicativo de descrever eventos passados, função essa encontrada no Doc 9835 AN/453. Ainda que alguns alunos tenham recorrido à escrita e ao uso do português, posso dizer que grande parte interagiu oralmente em inglês. Na correção, o professor solicitou que um aluno de cada grupo recontasse a história lida para toda a turma. Sua intervenção ocorreu após cada apresentação, retomando os principais pontos da história e corrigindo erros de pronúncia e vocabulário. Erros de estrutura, como dito anteriormente, costumam ser pouco tratados pelo professor.

Na segunda atividade proposta, o professor preferiu o modelo de interação professor-classe, colocando perguntas para todo o grupo. Contudo, poucos alunos participaram, e os turnos de fala foram bastante curtos, o que fez com que o professor dominasse o tempo de fala. Confirmamos que, nas turmas menos proficientes, é necessário oferecer um tempo para os alunos organizarem seu discurso, seja na forma escrita ou na interação oral com outro colega. Caso esse tempo não seja concedido, a participação fica centralizada em poucos alunos, provavelmente, aqueles com um nível um pouco maior de proficiência ou que são mais desinibidos.

Na entrevista, quando indagado sobre o objetivo de sua aula, Eduardo explica que pretendia desenvolver a produção oral dos alunos na atividade de contar/recontar histórias,

sugerindo provável impacto do exame, ainda que ele não tenha mencionado o exame em sala de aula. Nesse trecho, Eduardo também discorre sobre o seu objetivo pessoal de reduzir o tempo de fala, maximizando, assim, as oportunidades de uso da língua pelos alunos, o que constitui um dos princípios da abordagem comunicativa.

> de *speaking* né, (+) eu queria que o meu *teacher's time* fosse pequeno pra dar tempo de eles falarem /.../ no Eplis eles TÊM que falar né, (+) eles têm que *retell a story no eplis* (+) não' *TELL a story* né, (+) ah::: *describe the picture and tell a story* (+) é um minuto pra cada um que eles têm" /.../ seria isso'(+) de dar (+) de tentar melhorar as habilidades de *speaking* (+) as habilidades de *speaking* deles (+) de contar uma história (+) de construir frases (Entrevista com o professor Eduardo, conduzida em 30 de março).

Assisti mais uma aula do professor Eduardo na 3ª série, na turma D, no dia 28 de março, antes da condução da primeira entrevista. O Quadro 29 retrata os principais eventos. Nessa aula, a interação retornou a ser dominada pelo padrão professor-classe, com praticamente nenhuma interação entre os alunos. O professor Eduardo iniciou dizendo aos alunos que uma das tarefas presentes no EPLIS é a descrição de uma figura e os aconselhou a explorarem, primeiramente, o contexto da imagem. Eduardo, então, mostrou uma figura do livro didático e colocou perguntas para a turma envolvendo o entorno da imagem. Essa atividade foi conduzida no padrão professor-classe, com a participação de poucos alunos, e o turno de fala foi dominado pelo professor. Essa foi a primeira vez que Eduardo mencionou em sala de aula um dos aspectos do exame, porém não discutiu detalhes.

A escolha do professor em realizar esse tipo de atividade em sala de aula pode ser interpretada como impacto do exame. Durante a entrevista, o professor revelou que costuma trabalhar a descrição com mais frequência na 3ª série do que na 2ª série e afirmou que, quando atua na 4ª série, esse tipo de atividade está presente diariamente em sua prática, indicando que a frequência com que essa atividade é desenvolvida aumenta com a proximidade do exame. Quando indagado sobre a razão, Eduardo afirma ser por influência do EPLIS e, também, pelo fato de o livro didático da 4ª série propor essa tarefa. As atividades de relatar uma história e descrever uma imagem parecem ser muito recorrentes na prática do professor, o que pode indicar que essas duas tarefas são mais salientes em termos de efeito retroativo. Por outro lado, isso pode ser prejudicial por reduzir os tipos de atividades propostas e as funções comunicativas desenvolvidas em sala de aula.

> Pesquisadora: isso (descrição de imagem) é algo que você sempre faz?
> Eduardo: sim (+) eu sempre começo pelo contexto (+) quando eu vou dar uma atividade de *describe the picture* (+) eu sempre falo'(+) primeiro que eu penso assim

> / (+) eu sempre falo pra eles assim'(+) você tem que me ajudar como se você estivesse descrevendo a foto pra uma pessoa no telefone (+) então, você não pode começar assim'(+) *the aircraft is* (+) por que" (+) que *aircraft*" (+) onde tá essa *aircraft*" (+) tem uma *aircraft*" (+) então, você tem que começar pelo contexto (+) essa é uma foto tirada no aeroporto /.../
> Pesquisadora: essa tarefa de *describing the picture* foi feita com a 3ª série (+) você faz essa tarefa em outras séries"
> Eduardo: mais na 3ª do que na 2ª (+) eu acho /.../ na 4ª série todas as unidades (+) praticamente todas as aulas (+) todos os dias tem um *describing the picture* na 4ª série
> Pesquisadora: Por quê?
> Eduardo: no caso do Eplis né, (+) e porque o livro vem com uma foto (+) primeira coisa que tem na unidade (+) e tem atividades de *describing the picture* no livro (+)

Quadro 29 − Aulas observadas na 3ª série turma D

Professor: Eduardo	**Série: 3ª**	**Turma: D**	**No. Alunos: 17**	**Tempos de aula: 02**	**Data: 28/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
4 min	Descrever uma figura	Professor − classe	Problemas com passageiros / descrever ações em progresso	Produção oral	Visual, proposta criada pelo professor
6 min	Responder, oralmente, questões colocadas pelo professor	Professor − classe	Problemas com passageiros / emitir opinião	Produção oral	Vídeos, proposta criada pelo professor
8 min	Apresentação de vocabulário	Professor − classe	Problemas com passageiros	Vocabulário e leitura em voz alta	Escrito, proposta criada pelo professor
6 min	Simular uma situação a bordo do avião	Aluno − classe	Problemas com passageiros	Produção oral	Oral, proposta criada pelo professor

O padrão de interação professor-classe foi mantido para a segunda atividade. Eduardo mostrou aos alunos dois vídeos autênticos sobre problemas com passageiros a bordo de aeronaves. Colocou, então, algumas perguntas abertas para a turma. Porém, os alunos não se engajaram muito. Na terceira atividade, o professor apresentou uma série de textos curtos, retirados da internet, sobre famosos que causaram problemas durante o voo. Solicitou aos alunos que os lessem em voz alta. Após a leitura de cada trecho, Eduardo comentava o conteúdo, explicava vocabulário desconhecido e colocava algumas perguntas à turma, dominando o tempo de fala, enquanto os alunos produziam turnos bastante curtos. Por fim, Eduardo pediu aos alunos que simulassem uma situação a bordo do avião envolvendo um passageiro problemático, dois comissários de voo e um policial. Embora, tenha sido uma atividade com foco na integração das habilidades orais, ela não se assemelha ao que ocorre em uma situação autêntica, na qual, por exemplo, o piloto deve relatar ao órgão de controle a situação em que passageiro o passageiro se encontra. Além da inadequação da proposta,

somente três alunos participaram em razão limitação do tempo. O restante apenas assistiu, sem nenhum envolvimento.

A última observação na 3ª série ocorreu no dia 11 de abril na turma Eco. O Quadro 30 resume os eventos ocorridos. Essa mesma aula já havia sido assistida por mim no dia 22 de março na turma Fox pelo fato de haver entre professores e instrutores um rodízio, o que faz com que a mesma lição seja dada por eles em três ou quatro turmas diferentes.

Quadro 30 – Aulas observadas na 3ª série turma E

Professor: Eduardo	**Série: 3ª**	**Turma: E**	**No. Alunos: 15**	**Tempos de aula: 02**	**Data: 11/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
8 min	Ler e discutir uma história baseada em um texto	Individual /grupo (8min)	Problemas com passageiros / descrever eventos passados	Leitura e produção oral	Escrito / livro, proposta criada pelo professor
8 min	Responder, oralmente, questões colocadas pelo professor	Professor − classe	Problemas com passageiros / emitir opinião	Produção oral	Vídeos, proposta criada pelo professor
20 min	Recontar a história previamente discutida	Alunos – classe	Problemas com passageiros / descrever eventos passados	Produção oral e compreensão oral	Oral, proposta criada pelo professor
5 min	Responder, oralmente, questões colocadas pelo professor	Professor − classe	Problemas com passageiros / descrever eventos passados	Produção oral	Oral, proposta criada pelo professor
13 min	Responder, por escrito, perguntas sobre os textos apresentados	Individual	Problemas com passageiros	Leitura	Escrito / livro
10 min	Correção da atividade, checagem de pronúncia, vocabulário e estrutura	Professor − classe	Problemas com passageiros	Pronúncia, vocabulário e estrutura	Oral

Primeiramente, o professor dividiu os alunos em grupos e atribuiu a cada grupo um texto que deveria ser lido por eles e discutido. Durante essa atividade, o professor monitorou o trabalho dos grupos, sanando dúvidas sobre o vocabulário presente nos textos. Ainda que alguns alunos tenham usado português, a maioria interagiu em inglês, sem a necessidade de recorrer à escrita. Em seguida, o professor mostrou alguns vídeos de situações semelhantes às dos textos e colocou perguntas aos alunos no padrão professor-classe, o que levou à baixa participação. O fato de esse tipo de atividade não engajar os alunos a produzir turnos mais longos de fala, principalmente, os menos proficientes, já havia sido discutido com o professor nas entrevistas, demonstrando que reconstruir a prática pedagógica não é uma tarefa fácil, tampouco simples.

Em seguida, Eduardo solicitou aos grupos que recontassem a história que haviam lido no início da aula. Diferentemente da aula na turma F, nessa atividade, cada aluno foi incentivado a contar uma parte da história, sem recorrer ao texto, e a explicar o significado de palavras julgadas desconhecidas, o que fez com que todos participassem ativamente. Após cada apresentação, Eduardo pedia a um outro grupo para explicar a história que havia sido contada, a fim de checar se os demais haviam compreendido. Dessa forma, o professor conseguiu integrar as habilidades de compreensão oral e produção oral em uma única atividade, e os alunos produziram turnos de fala mais longos.

Eduardo, então, apresentou algumas fotos de pessoas famosas que apresentaram mal comportamento a bordo de aeronaves. Algumas perguntas foram colocadas aos alunos que, por sua vez, não se engajaram muito, dado o padrão de interação professor-classe. Eduardo finalizou a atividade e pediu que a turma respondesse a dois exercícios do livro didático que checavam a compreensão dos textos lidos anteriormente. Durante a correção, o professor Eduardo verificou o significado das palavras aprendidas na lição, demonstrando sua preocupação para com o ensino de vocabulário.

Em suma, na primeira aula observada na 2ª série do curso, o professor Eduardo mostrou-se bastante controlador, principalmente, em relação ao ensino de vocabulário. O padrão de interação dominante foi o de professor-classe com poucas oportunidades de interação entre os alunos. Os alunos limitaram-se a ler em voz alta trechos de textos extraídos do livro didático ou da internet, enquanto o professor explicava o significado de palavras específicas. A partir da segunda aula assistida, o professor criou mais oportunidades de trabalho em pares ou em grupos. Os alunos interagiram mais livremente e produziram turnos de fala mais longos, o que condiz com os efeitos pretendidos do exame, podendo também ser resultado da entrevista conduzida. Apesar de o exame não ter sido mencionado nessa série, a ênfase dada pelo professor à descrição de imagens e ao relato de histórias sugere possível efeito retroativo. No entanto, outras tarefas do exame como, por exemplo, explicar uma situação problemática de controle de tráfego aéreo, sugerir uma solução para o problema e emitir opinião sobre determinados assuntos não foram observadas em nenhuma de suas aulas, o que pode indicar que o impacto do teste se concentra em apenas alguns de seus aspectos.

Na 3ª série do curso, o professor atuou um pouco mais como facilitador, mesmo em turmas menos proficientes, como foi o caso das turmas Eco e Fox. Houve mais oportunidades de produção oral livre, durante as quais Eduardo incentivou os alunos a falarem em inglês e sanou dúvidas referentes a vocabulário e estrutura. No entanto, ainda houve diversos momentos em que o professor recorreu ao padrão de interação professor-classe, com leitura em voz alta

por parte dos alunos, principalmente, quando sua intenção era ensinar vocabulário. Em termos de integração de habilidades, vimos que é mais comum integrar a leitura à produção oral. Já a integração entre compreensão e produção orais é mais rara e ocorre, geralmente, quando o professor coloca perguntas a serem respondidas pelos alunos, porém poucos costumam participar. Constatamos, também, que o professor Eduardo privilegia as tarefas de descrição de imagens e criação de histórias mais na 3ª série do que na 2ª, fazendo com que os propósitos comunicativos girem em torno de descrever estados, eventos em progresso e eventos passados. Também não foram observadas atividades que focassem no gerenciamento da comunicação aeronáutica e em estratégias para lidar com mal-entendidos. Como apontado anteriormente, esse tipo de tarefa é abordado tanto no material didático quanto na prova oral e sua ausência em sala de aula pode estar relacionada ao fato de o EPLIS não a contemplar, exercendo, assim, um possível efeito negativo.

5.3.3 O instrutor João

O instrutor João é controlador de tráfego aéreo há seis anos e atua há cinco anos como instrutor de inglês aeronáutico. Na EEAR, desde 2016, tem ministrado aulas de inglês para os alunos das 2ª e 3ª séries, como instrutor convidado. Em termos de capacitação, João realizou o curso de formação de instrutores de inglês aeronáutico, oferecido pelo ICEA, em 2012. João submeteu-se ao EPLIS como candidato em 2011 e obteve nível 5 (avançado) de proficiência. Alegou, em sua entrevista, que vem construindo conhecimentos sobre o EPLIS por meio das informações vinculadas na página do exame na *web* e citou o manual, o *quiz* disponibilizado para a fase I e os vídeos que exemplificam a fase II como fontes de conhecimento sobre o teste. No ano de 2017, João atuou nas 2ª e 3ª séries do curso. Durante o período da observação, João ministrou aulas nas turmas *G*olf e *E*co, da 2ª série, e nas turmas *G*olf, *E*co e *F*ox, da 3ª série. A turma *G* de ambas as séries é considerada a menos proficiente. Considerando que João é um instrutor de inglês aeronáutico, como observamos no estudo quantitativo, esse grupo julgou mais positivamente os níveis de conhecimentos que possuem sobre o exame, comparados aos professores, bem como a frequência com que são procurados pelos alunos para esclarecer dúvidas sobre o teste. A presença de leitura em voz alta também foi reportada como menos frequente entre os instrutores, bem como entre os docentes menos experientes. Na análise das discussões ocorridas no grupo focal, os alunos afirmaram que os instrutores discutem e trabalham com maior frequência tarefas relacionadas ao EPLIS, embora

a análise estatística das respostas dos professores e instrutores ao questionário tenha mostrado pouca diferença entre esses dois profissionais nesses aspectos.

5.3.3.1 As aulas na 2ª série do curso BCT

A primeira aula observada do instrutor João, na 2ª série, ocorreu na turma G, no dia 23 de março. O Quadro 31 retrata os principais eventos ocorridos. Os aspectos mais salientes foram a ênfase no ensino de vocabulário, padrão de interação professor-classe, nenhuma oportunidade de trabalho em grupo ou em pares e produção de turnos curtos de fala pelos alunos, distanciando-se, assim, dos efeitos pretendidos.

Quadro 31 − Aulas observadas na 2ª série turma G

Instrutor: João	**Série: 2ª**	**Turma: G**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 23/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
18 min	Sequenciar as etapas dentro de cada fase do voo e relacioná-las ao órgão de controle de tráfego aéreo	Instrutor − classe	Fases do voo	Vocabulário + estrutura *is related to*	Escrito, proposta criada pelo instrutor
7 min	Criar frases utilizando as fases do voo	Individual	Fases do voo	Vocabulário + estrutura (ex.: *Cruising is related to ACC facility*)	Escrito, proposta criada pelo instrutor
5 min	Adivinhar as fases do voo, utilizando uma carta de aeródromo	Instrutor − classe	Fases do voo	Vocabulário	Visual, proposta criada pelo instrutor
8 min	Apresentar os riscos envolvidos nas diferentes fases do voo	Instrutor − classe	Riscos em voo	Vocabulário	Visual e áudio autêntico, proposta criada pelo professor
10 min	Nomear os riscos envolvidos nas diferentes fases do voo	Instrutor − classe	Riscos em voo	Vocabulário e pronúncia	Visual, proposta criada pelo professor

João deu início à sua aula chamando um aluno por vez à lousa para que colocasse a sequência correta das etapas de voo. Durante essa atividade, o instrutor checou o significado de diversas expressões como, por exemplo (*start up*, *line up*, *roll for take off*), bem como a pronúncia e forma escrita. Os alunos também foram orientados a relacionar a etapa do voo ao órgão de controle responsável, utilizando a expressão *is related to*, por exemplo, *start up is related to tower*. A produção linguística dos alunos ficou, portanto, restrita às palavras abordadas e a essa sentença modelo. Em seguida, João pediu aos alunos que escrevessem essas

frases e, para fins de correção, que as lessem em voz alta. Por fim, um aluno foi chamado à frente da sala para que simulasse um voo, utilizando uma carta de um aeroporto e um avião de brinquedo. Enquanto o aluno fazia os movimentos, o instrutor colocava perguntas à sala com o objetivo de revisar o vocabulário ensinado. As respostas dos alunos eram curtas, normalmente, formadas por uma ou duas palavras.

Para apresentar o tópico Riscos em voo, João solicitou dos alunos os problemas mais comuns em cada fase do voo. Os alunos citaram alguns e o instrutor apresentou o termo *baloons*. Em seguida, ilustrou a importância do tópico colocando um áudio autêntico em que um controlador brasileiro reporta em inglês às aeronaves sob seu controle balões nas proximidades, cujas legendas estavam em português. Durante o áudio, João explicou as expressões utilizadas pelo controlador para comunicar aos pilotos a informação. Expor os alunos a situações reais é uma forma de contribuir para o aumento da consciência sobre a importância da língua inglesa no ambiente de trabalho. Isso já havia sido relatado pela professora Ana durante a entrevista, quando afirmou que a vinda dos instrutores de tráfego aéreo colaborou bastante nesse sentido.

Em seguida, o instrutor apresentou algumas imagens à turma com o objetivo de ensinar palavras relacionadas a riscos em voo. A produção dos alunos ficou, novamente, restrita à repetição dessas palavras e não houve a oportunidade de colocá-las em uso. Na entrevista, João admitiu que, nas turmas menos proficientes, por causa do tempo maior dedicado à apresentação do vocabulário, as atividades de produção oral costumam ficar comprometidas, o que também ocorreu nas aulas de Pilar e Eduardo.

Outro ponto abordado na entrevista foi em relação à ênfase dada no ensino de expressões (*chunks*) como, por exemplo *is related to*. O instrutor João argumentou que foi a forma que encontrou de fazer com que os alunos com baixíssima proficiência tivessem condições de produzir algo em inglês na fase II do EPLIS, caso viessem a ser aprovados na fase I. Nas palavras de João, se não fosse ensinado esse tipo de expressão para os alunos, eles não seriam capazes de interagir durante o exame: "junto com o vocabulário ele tem que ter um *chunk* (+) pra ele poder formar frase (+) senão ele não vai falar (+) aí de novo a dificuldade maior' (+) sargento, eu não sei falar nada". João acrescentou que, nessas turmas, a ênfase é mais no ensino de vocabulário e *chunks*, permeado de muita repetição (*drilling*), indicando uma base mais estruturalista.

Aqui cabe uma ressalva importante sobre o nível de proficiência dos alunos. Embora, o concurso a que eles se submeteram para ingresso na EEAR tenha como exigência o nível intermediário em inglês, a prova de entrada avalia apenas a compreensão de textos

escritos. Portanto, a ausência de uma prova que avalie as habilidades de compreensão oral e produção oral, primordiais no serviço de tráfego aéreo, permite que alunos com pouca ou nenhuma proficiência nessas habilidades tenham acesso ao curso. O sistema educacional da EEAR, por sua vez, não faz distinção entre os alunos, a não ser agrupá-los por níveis de proficiência. Assim sendo, o material didático e as avaliações estão acima da capacidade dos alunos que apresentam baixa proficiência. Os professores, portanto, tentam adaptar as atividades, o que faz com que a produção oral livre, ou seja, com turnos mais longos de fala, seja pouco recorrente em sala de aula.

Foi também assistida mais uma aula do instrutor João na turma E da 2ª série (Quadro 32). Essa aula teve também como foco o ensino de vocabulário. Em termos de interação, foi criada oportunidade de os alunos trabalharem em grupos na segunda atividade. As demais atividades permaneceram no padrão professor-classe, com o instrutor colocando perguntas aos alunos para coletar informações, detendo, assim, o maior tempo de fala. As respostas dos alunos a essas perguntas eram, na maior parte das vezes, bem curtas.

Quadro 32 − Aulas observadas na 2ª série turma E

Instrutor: João	**Série: 2ª**	**Turma: E**	**No. Alunos: 19**	**Tempos de aula: 02**	**Data: 30/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
11 min	Identificar instalações aeroportuárias em uma carta de aeródromo	Instrutor − classe	Instalações aeroportuárias	Vocabulário	Visual, proposta criada pelo instrutor
26 min	Relacionar imagens, palavras e definições	Trabalho em grupo	Instalações aeroportuárias	Vocabulário	Escrito, livro didático, proposta criada pelo instrutor
30 min	Apresentar as instalações aeroportuárias	Alunos − classe	Instalações aeroportuárias	Expressões enfáticas (*let me stress that / I'd like to emphasize / It's important to remember* + leitura em voz alta	Visual, proposta criada pelo instrutor

Durante a atividade em grupo, a maioria dos alunos se esforçou para interagir em inglês, apesar de o português estar ainda bastante presente. O instrutor monitorou o trabalho dos grupos, explicando vocabulário e checando pronúncia. Antes da correção da atividade, João apresentou três expressões de ênfase para que fossem usadas por eles na apresentação das instalações, o que ocorreu em seguida. As expressões eram: *let me stress that*, *I'd like to emphasize*, *it's important to remember*. Os alunos, de fato, utilizaram essas expressões, ao

serem chamados à frente da sala para explicarem o significado das palavras. Eles iniciavam sua fala com uma dessas expressões e, em seguida, liam em voz alta a definição do termo.

Em uma das entrevistas, conversamos novamente sobre a utilidade dessas expressões. João reforçou que o seu uso está voltado intencionalmente para o EPLIS, como forma de ajudar os alunos menos proficientes a se expressarem na fase II, evitando, assim, longos períodos de silêncio, o que poderia gerar grande frustração. Esse tipo de estratégia de uso da língua ensinada por João é classificado por Cohen (1998) como estratégia de proteção, que consiste em criar a ilusão de que o sujeito domina o material linguístico (KOBAYASHI, 2016). Assim, reporta João:

> ele se sente confiante (+) ele sente que ele consegue falar inglês (+) porque ele acaba decorando expressões simples (+) geralmente três a quatro expressões (+) e ele consegue perceber que com essas três ou quatro expressões ele consegue definir qualquer coisa (+) e isso também mostra / demonstra que ele consegue ser informativo em uma entrevista (+) e aí a minha TENDÊNCIA que você perguntou inicialmente'(+) você prepara / a sua aula é mais tendenciosa pro exame ou pra prova" (+) é mais pro exame (+) no momento que eu passo isso eu mostro que ele é capaz (+) eu mostro que ele consegue MESMO não sendo tão proficiente passar por uma entrevista e falar alguma coisa (Entrevista com o instrutor João, conduzida em 11 de abril de 2017).

Nesse trecho, também podemos perceber que o instrutor João relaciona o uso dessas expressões à questão de ser informativo. A informatividade é um dos aspectos analisados dentro do critério "interação" da tabela de níveis da OACI, juntamente com a adequação e a imediaticidade da resposta do candidato, o que indica possível impacto do exame na prática do instrutor. Por outro lado, João parece reduzir esse aspecto à quantidade de informação dita, ou seja, para ser informativo o candidato deve produzir bastante. Na entrevista, perguntei em qual parte do exame essas expressões poderiam ser utilizadas. Assim, João responde:

> TODAS (+) todas (+) exceto a história (+) todas (+) até mesmo pra ele começar a descrever a figura (+) ele pode começar (+) *I would like to emphasize that the picture shows* (+) na:a segunda fase (+) na primeira partezinha (+) que tem as perguntas de ordem pessoal (+) ele pode começar com *I would like to emphasize* (+) eu acredito (+) você pode até me corrigir (+) que o exame é livre (+) e QUANTO MAIS dados o avaliador tiver (+) MELHOR pro avaliador poder avaliar (Entrevista com o instrutor João, conduzida em 29 de março de 2017).

Logo em seguida, questionei-o sobre o uso dessas expressões no ambiente de trabalho, João diz enfaticamente que elas não têm nenhum uso prático: "Esse *chunk* não ajuda em NADA". Essa percepção de João nos remete às funções comunicativas avaliadas no exame e ao domínio de uso do controlador de tráfego aéreo. Scaramucci et al (2010) já haviam

pontuado problemas relativos à autenticidade das tarefas presentes na fase II do teste, principalmente, nas partes 1 e 3, que se distanciam da natureza da interação entre piloto e controlador. Outra crítica feita por Scaramucci et al (2010) reside no fato de que as tarefas dessa fase não propiciam uma interação colaborativa genuína entre candidato e interlocutor, característica principal da comunicação aeronáutica. Esses óbices encontrados no exame parecem estar afetando negativamente o ensino, ao afastá-lo da natureza da linguagem da comunicação aeronáutica. Isso foi percebido, por exemplo, na pouca ou nenhuma ênfase dada em sala de aula em atividades que propõem o gerenciamento do diálogo entre piloto e controlador e no uso de estratégias comunicativas para a negociação do sentido.

Outra questão interessante suscitada na entrevista com João diz respeito à lista de palavras ensinadas aos alunos nessa aula. João demonstra seu descontentamento ao dizer que alguns daqueles termos não são utilizados no contexto brasileiro de tráfego aéreo. Perguntei, então, qual seria a razão de trabalhá-los e João justificou que eram assuntos previstos no currículo escolar e cobrados na prova teórica da EEAR: "PUD (+) pud do curso (currículo escolar) e saber que na prova deles eles vão ser cobrados". Essa fala exemplifica e confirma o poder exercido pelas provas escritas no direcionamento do que é abordado em sala de aula, além de demonstrar que essas avaliações precisam estar mais alinhadas às necessidades dos alunos. Dado que o livro didático compõe o currículo escolar, é imprescindível que seja feita uma adequação do material às características da comunicação aeronáutica no controle de tráfego brasileiro.

5.3.3.2 As aulas na 3ª série do curso BCT

Foram observadas aulas do instrutor João em três turmas da 3ª série: *E*co, *F*ox e *G*olf. O Quadro 33 retrata os principais eventos das aulas assistidas na turma G, nos dias 27 e 28 de março. Essa turma é considerada a menos proficiente dentre as turmas da 3ª série, e os alunos apresentam muita dificuldade em se expressar oralmente, o que ficou evidente diante do uso excessivo do português por parte deles nas atividades em grupo. Na entrevista que sucedeu à observação, João reportou que, além de possuírem baixa proficiência, esses alunos mostram-se bastante imaturos e pouco conscientes da importância do inglês na profissão que escolheram: "MUITO português (+) é a turma que não está consciente do Eplis (+) ESSE exemplo (+) esse grupo (+) você tem um dado aí que ELES não estão conscientes".

Quadro 33 − Aulas observadas na 3ª série turma G

Instrutor: João	**Série: 3ª**	**Turma: G**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 27e 28/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
20 min	Preparar um pôster sobre as vantagens e desvantagens do CPDLC	Trabalho em grupo	Tecnologias aplicadas ao controle de tráfego aéreo / emitir opinião	Escrita + produção oral	Escrito, proposta criada pelo instrutor
13 min	Apresentar, oralmente, as vantagens e desvantagens do CPDLC	Alunos − classe	Tecnologias aplicadas ao controle de tráfego aéreo / emitir opinião	Produção oral + expressões enfáticas	Oral, proposta criada pelo instrutor
13 min	Apresentação de expressões que indicam propósito	Instrutor – classe	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito: *so that, in order to, for, to*)	Escrito, proposta criada pelo instrutor
14 min	Completar sentenças utilizando as expressões de propósito	Individual (7min) Instrutor-classe (7min)	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito)	Escrito / livro
12 min	Relacionar palavras e definições	Trabalho em grupo	Tecnologias aplicadas à aviação	Vocabulário	Escrito, proposta criada pelo instrutor
12 min	Correção da atividade, explicação de vocabulário e pronúncia	Instrutor-classe	Tecnologias aplicadas à aviação	Vocabulário e pronúncia	Oral
7 min	Completar sentenças com as palavras dadas	Individual	Tecnologias aplicadas à aviação	Vocabulário	Escrito / livro
5 min	Selecionar a resposta correta de acordo com o áudio	Individual (3min) Em pares (2min)	Tecnologias aplicadas à aviação	Compreensão oral	Áudio /livro, proposta criada pelo instrutor
5 min	Correção da atividade, explicação das respostas	Instrutor-classe	Tecnologias aplicadas à aviação		Oral

Em relação às atividades desenvolvidas na aula, a primeira teve como objetivo engajar os alunos na confecção de um pôster sobre as vantagens e desvantagens do CPDLC (*controller-pilot datalink communication*), comunicação via mensagens de texto, no controle de tráfego aéreo. Durante sua realização, houve uso excessivo do português pela maioria dos alunos. No entanto, essa atividade foi essencial para que a apresentação do pôster pudesse ocorrer, já que os alunos menos proficientes necessitam organizar, por escrito, sua fala. Sendo assim, alguns, de fato, conseguiram produzir turnos mais longos, sem a necessidade de recorrerem às suas anotações, e utilizaram as expressões de ênfase ensinadas pelo professor. Porém, a maioria leu o que havia sido escrito no pôster. No fim de cada apresentação, o instrutor corrigiu alguns erros de pronúncia e estrutura cometidos.

Quando perguntado sobre o propósito dessa atividade, João explicou que esse tópico é bastante recorrente na parte 3 da fase II do EPLIS, quando o candidato é solicitado a expressar sua opinião sobre diversos assuntos ligados ao contexto profissional, e, por essa razão, costuma enfatizá-lo em sala de aula, evidenciando influência do exame na ênfase dada a determinado conteúdo e à função de emitir opinião presente no teste.

> eu costumo dizer para os alunos MAIS proficientes também (+) de que esses são tópicos importantes (+) se tá ali no livro são tópicos importantes (+) ele tem algum motivo pra estar ali (+) e os alunos que estão mais conscientes eles já sacam na hora' (+) o sargento deve tá falando isso por causa do Eplis (+) o Eplis pergunta sobre tecnologia né, (Entrevista com João, instrutor na EEAR desde 2016).

Essa fala de João também respalda a percepção de que a abordagem de ensinar do instrutor varia de acordo com a proficiência dos alunos. João diz que, com os alunos mais proficientes, ele ressalta mais a importância dos tópicos abordados em sala de aula. Esses alunos, por estarem mais conscientes sobre o EPLIS, costumam compreender a associação feita pelo instrutor rapidamente.

Retornando à aula, o ensino das expressões de propósito ficou centralizado no instrutor que, por sua vez, explicou as regras gramaticais de uso dessas expressões. Em seguida, os alunos fizeram dois exercícios escritos do livro didático sobre esse assunto. Não houve oportunidades para a prática oral dessas expressões, embora houvesse uma proposta no livro didático com esse objetivo. Essa forma de tratar aspectos gramaticais é considerada estruturalista, pois não estimula o aprendiz a descobrir as regularidades linguísticas de forma intuitiva (VIEIRA-ABRAHÃO, 2015) e na interação com o outro.

Em relação à preparação que antecedeu às atividades de compreensão oral, o instrutor pediu aos alunos que, em grupos, relacionassem palavras e definições e, em seguida, completassem sentenças com essas palavras e checassem suas respostas em pares, o que fez com que houvesse mais interação entre eles. Para a compreensão oral, João elaborou um exercício muito semelhante ao encontrado na fase I do EPLIS com questões de múltipla escolha em português. Na entrevista, quando perguntado sobre o motivo de ter elaborado o exercício nesse formato, João responde:

> já deixando nos moldes da primeira fase (Eplis) e da 4ª série (+) a 4ª série encontra as atividades dessa maneira (+) como eles já são da 3ª série pra 4ª série e eles são um grupo menos proficiente (+) eu não quero deixá-los MENOS eh: alertas do que os mais proficientes que já estão correndo atrás (+) então, eu já trouxe pra eles (+) pelo menos uma vez na aula (+) eles já viram alguma coisa que (+) o Eplis vai correr assim (+) lembra que eu não falei (+) não citei em nenhum momento (+) mas eles já sabem

> (+) alguns já sabem ali (+) e até uma menina comentou'(+) por que o Eplis tem que ser assim" (Entrevista com o instrutor João, conduzida no dia 11 de abril de 2017).

O fato do instrutor João ter antecipado um tipo de exercício que seria visto na 4ª série do curso confirma as percepções dos alunos de que os instrutores de inglês aeronáutico e alguns professores, por conhecerem melhor o exame, costumam trabalhar tarefas semelhantes nas séries anteriores, ainda que não citem diretamente o EPLIS em sala de aula, tendo o mesmo ocorrido com a professora Pilar.

Nesse trecho, o instrutor João também evidencia mais uma vez diferenças no modo como alunos menos e mais proficientes lidam com o exame, confirmando os resultados quantitativos que sugeriram que os mais proficientes procuram saber mais sobre o teste, engajam-se em práticas preparatórias com maior frequência e reconhecem a importância da aprovação. Para compensar esse desequilíbrio, João tenta trabalhar nas turmas menos proficientes aspectos do exame que já vêm sendo explorados em turmas mais avançadas. Embora não tenha sido possível observar aulas do instrutor João em turmas com níveis mais elevados de proficiência, seus comentários revelam diferenças no modo como aborda o exame em turmas menos e mais proficientes.

Grande parte da unidade ministrada na turma G foi assistida por mim na turma E no dia 30 de março. O Quadro 34 recorda os principais eventos. Os alunos da turma E são um pouco mais proficientes do que os da turma G, o que fez com que interagissem mais em inglês nas atividades em grupo, com uso ocasional do português. Durante a apresentação oral, grande parte dos alunos se esforçou para não ler suas anotações e produziu turnos de fala um pouco mais longos e mais coesos. E, assim como ocorreu com o professor Eduardo, João não tratou pontualmente os erros gramaticais, considerados globais, no uso de estruturas básicas cometidos pelos alunos.

Outra diferença entre as turmas foi a redução do tempo gasto na atividade em que o instrutor explicou o uso de expressões de propósito. Enquanto, na turma G, a apresentação tomou 13 minutos, na turma E, foram dedicados a ela apenas 5 minutos, diminuindo, assim, o tempo de fala do professor em sala de aula e proporcionando mais tempo para que outras atividades pudessem ser desenvolvidas. Contudo, a apresentação ainda foi centralizada no instrutor. A correção da atividade sobre o uso dessas expressões envolveu um pouco mais os alunos, no sentido de que eram eles os responsáveis por justificar cada resposta especificando a regra gramatical, o que aumentou o tempo de fala deles e contribuiu para promover consciência linguística. Os alunos também participaram mais da correção do exercício de relacionar palavras/expressões e suas definições, ao terem que explicar com suas próprias

palavras os termos. Porém, não foram feitas atividades que permitissem o uso do novo vocabulário de forma mais contextualizada.

Quadro 34 − Aulas observadas na 3ª série turma E

Instrutor: João	**Série: 3ª**	**Turma: G**	**No. Alunos: 16**	**Tempos de aula: 02**	**Data: 30/03/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
8 min	Preparar uma apresentação sobre as vantagens e desvantagens do sagitário e do CPLC	Trabalho em grupo	Tecnologias aplicadas ao controle de tráfego aéreo / emitir opinião	Escrita e produção oral	Oral, proposta criada pelo instrutor
12 min	Apresentar, oralmente, as vantagens e desvantagens do CPDLC	Alunos − classe	Tecnologias aplicadas ao controle de tráfego aéreo / emitir opinião	Produção oral + expressões enfáticas (*let me stress that / I'd like to emphasize / It's important to remember*	Proposta criada pelo instrutor
5 min	Apresentação de expressões que indicam propósito	Instrutor − classe	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito: *so that, in order to, for, to*)	Escrito, proposta criada pelo instrutor
13 min	Completar sentenças utilizando as expressões de propósito e correção da atividade	Individual (7min) Alunos-classe (11 min)	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito)	Escrito / livro
10 min	Relacionar palavras e definições	Trabalho em grupo	Tecnologias aplicadas à aviação	Vocabulário	Escrito, proposta criada pelo instrutor
11 min	Correção da atividade, explicação de vocabulário e pronúncia	Instrutor − classe	Tecnologias aplicadas à aviação	Vocabulário e pronúncia	Oral

Na entrevista, João menciona outra evidência de que a proficiência dos alunos afeta de certa forma o seu modo de ensinar. O instrutor revela que, nas turmas menos proficientes, tende a ser mais controlador, dominando o turno de fala em uma interação do tipo professor-classe: "por isso que você deve ter notado nas suas observações (+) que os menos proficientes (+) pelo menos COMIGO (+) eu controlo mais". João também sugere que os alunos menos proficientes são mais indisciplinados, o que aumenta a necessidade de o professor exercer mais controle sobre os eventos em sala de aula. Assim, relata:

> eu iria um pouquinho além (+) as turmas que são mais proficientes GERALMENTE (+) elas são mais disciplinadas (+) então, eu consigo confiar que o aluno vai tá fazendo a atividade (+) o aluno menos proficiente me parece que ele é MENOS disciplinado (+) ele já percebe'(+) eu já não sei nada mesmo (+) então, não vou fazer a atividade (+) vou FINGIR que tô fazendo (+) porque eu tenho dificuldade (+) vou ficar aqui

> enrolando (+) o tempo vai passar e eu vou embora (Entrevista com o instrutor João, conduzida no dia 11 de abril de 2017).

Uma aula semelhante foi ministrada pelo instrutor João na turma F no dia 12 de abril de 2017. Em termos de proficiência, essa turma localiza-se entre a turma E a G. O Quadro 35 resume as atividades realizadas. O tempo de fala do professor durante a apresentação sobre as expressões de propósito voltou a ser mais longo, assim como ocorreu na turma G. A correção dessa atividade também foi mais extensa, com o instrutor explicando novamente o tópico e checando o significado de palavras presentes no exercício. Na atividade em grupo, em que os alunos deveriam preparar uma apresentação sobre as vantagens e desvantagens do CPDLC, a maior parte dos alunos se esforçou para se comunicar em inglês, porém houve muito uso do português.

Quadro 35 − Aulas observadas na 3ª série turma F

Instrutor: João	**Série: 3ª**	**Turma: F**	**No. Alunos: 15**	**Tempos de aula: 02**	**Data: 12/04/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
12 min	Apresentação de expressões que indicam propósito	Instrutor – classe	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito: *so that, in order to, for, to*)	Escrito, proposta criada pelo instrutor
26 min	Completar sentenças utilizando as expressões de propósito e correção da atividade	Individual (9min) Instrutor – classe (23min)	Tecnologias aplicadas ao controle de tráfego aéreo	Estrutura (expressões de propósito: *so that, in order to, for, to*)	Escrito / livro
22min	Preparar uma apresentação sobre as vantagens e desvantagens do sagitário e do CPLC	Trabalho em grupo	Tecnologias aplicadas ao controle de tráfego aéreo / emitir opinião	Escrita + produção oral	Proposta criada pelo instrutor

Em suma, apesar de o instrutor João não ter mencionado ou discutido diretamente o EPLIS em sala de aula, constatamos evidências que sugerem impacto do exame em sua prática pedagógica. Na 2ª série, por exemplo, o instrutor ensinou estratégias que ajudam os alunos, principalmente, os menos proficientes, a se expressarem melhor na fase II do teste, quando eles são solicitados a emitir opinião sobre determinado assunto. No entanto, nas turmas menos proficientes, as aulas foram bastante centradas no instrutor, com ênfase no ensino de vocabulário, pouca interação entre os alunos, turnos de fala demasiadamente curtos. A leitura em voz alta também estava muito presente na sua prática pedagógica, com os alunos recorrendo

às anotações para a realização de atividades de produção oral. O instrutor João afirmou que tende a ser mais controlador nessas turmas devido à baixa proficiência e à indisciplina dos alunos. Realmente, na 3ª série, constatamos que, os alunos da turma G, considerada a menos proficiente, falam muito português durante as atividades em grupo e, por muitas vezes, os assuntos não são relativos à aula. Outro indício de influência do exame na prática do instrutor envolveu uma atividade de compreensão oral realizada nessa mesma turma. Embora o áudio trabalhado não tenha sido do gênero aeronáutico, o professor elaborou um exercício de múltipla escolha em português. Esse formato coincide com o encontrado na fase I do EPLIS. Na entrevista, João afirmou que esse exercício foi feito na tentativa de familiarizar os alunos com o exame e mencionou que, nas turmas mais proficientes, esse aspecto já vinha sendo explorado, o que respalda a hipótese de que, nessas turmas, aspectos do exame costumam ser mais discutidos e trabalhados.

5.3.4 A professora Ana

A professora Ana é formada em Letras – Inglês e possui 15 anos de experiência docente. Desde 2013, Ana atua na EEAR como professora de inglês aeronáutico para o curso BCT, ministrando aulas, principalmente, nas 1ª e 4ª séries. Na primeira entrevista que tivemos, Ana demonstrou seu interesse em saber mais sobre inglês aeronáutico, a fim de lidar melhor com as atividades propostas no material didático e com as dúvidas dos alunos. Quando questionada sobre como foi seu processo de construção desses conhecimentos, Ana menciona ter participado de aulas em português de controle de tráfego aéreo, porém ressalta que ele se deu, praticamente, por meio do estudo individual do livro didático. Em momentos de dúvida, ela diz ter recorrido à coordenação do curso para orientação e suporte. Quanto ao EPLIS, Ana explicou que seus conhecimentos foram sendo construídos por meio da leitura do manual do candidato disponibilizado pelo ICEA e por meio do contato com a coordenação pedagógica e com professores que já tinham atuado na 4ª série. Além disso, por razões profissionais, Ana não conseguiu acompanhar integralmente a oficina sobre o EPLIS, ocorrida na EEAR, em 2015. No ano de 2017, Ana atuou nas 1ª e 4ª séries do curso. Não foram feitas observações de aulas na 1ª série pelo fato de o currículo dessa série estar passando por reconstruções para se adequar melhor às necessidades dos alunos. Em relação à 4ª série, foi possível observar aulas da professora Ana em quatro turmas, *A*lfa, *C*harlie, *E*co e *F*ox, possibilitando uma visão mais detalhada sobre como a professora atua em grupos de níveis distintos de proficiência.

5.3.4.1 As aulas na 4ª série do curso BCT

O Quadro 36 resume os eventos ocorridos na aula assistida na turma E no dia 06 de julho de 2017. Como era o primeiro contato da professora Ana com essa turma, ela se apresentou e pediu que cada aluno fizesse o mesmo, colocando suas expectativas em relação à disciplina. Todos os alunos mencionaram que gostariam de ser aprovados no EPLIS e a professora Ana afirmou que iria ajudá-los a atingir esse objetivo. Essa foi a primeira vez que presenciei os alunos se referindo ao exame em sala de aula, o que evidencia que nessa série eles procuram conhecer melhor o teste e suas características e conferem importância ao exame.

Durante a aula, foi possível constatar evidências de efeito retroativo do exame na prática da professora, começando pela explicação sobre como descrever uma imagem. Ana apresentou aos alunos algumas locuções prepositivas para serem utilizadas nessa tarefa, tais como *on the top*, *at the bottom*, *on the right or left side*, *in the center*, *at the background*. Em seguida, os instruiu a procurarem responder às seguintes perguntas: *What is the scene? / Where is it? / What are the major features / Who or What is in the picture? / What are they doing? / What is happening?* Essas estratégias são classificadas por Cohen (2006) como estratégias de uso e quando aplicadas a uma situação avaliativa tornam-se estratégias para a realização de um teste. Elas são consideradas eficazes e adequadas, pois focam no desenvolvimento linguístico, além de ajudarem os alunos a lidar melhor com as demandas da tarefa.

Ana, então, explicou aos alunos que, no EPLIS, não é permitido ao candidato fazer qualquer anotação escrita durante a descrição da imagem e os aconselhou a se esforçarem para se expressar oralmente. Em grupos, a professora pediu aos alunos que descrevessem uma imagem do livro didático por um período razoável de tempo, o que promoveu bastante interação entre eles. Durante a correção, ao notar que os alunos estavam utilizando expressões de sequência, Ana reforçou que, nesse tipo de tarefa, soa mais natural o uso de locuções prepositivas, enquanto que as expressões de sequência são mais comuns na criação de histórias. Na entrevista, ao ser perguntada sobre o quão útil seriam as estratégias ensinadas para a descrição de imagens, Ana revelou que os alunos menos proficientes demonstram bastante interesse em utilizá-las, diferentemente dos alunos mais proficientes que, segundo ela, não necessitam tanto dessas ferramentas. Assim, relata:

> eu acho que eles gostam (+) eles assim'(+) entendem isso como uma ferramenta legal (+) pra ser usada antes da descrição (+) é diferente da turma mais proficiente (+) os alunos que sabem mais eles nem (+) não sei (+) acho que já tá no automático (+) então, eles não se importam muito (Entrevista com a professora Ana, conduzida no dia 11 de julho de 2017).

Quadro 36 − Aulas observadas na 4ª série turma E

Professora: Ana	**Série: 4ª**	**Turma: E**	**No. Alunos: 13**	**Tempos de aula: 03**	**Data: 06/07/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
10 min	Apresentar-se e falar sobre suas expectativas	Aluno – classe	Apresentação pessoal	Produção oral	Proposta criada pela professora
8 min	Responder, oralmente, questões colocadas pelo professor sobre acidentes envolvendo fogo no motor da aeronave	Professor – classe	Fogo no motor da aeronave / emitir opinião	Produção oral	Audio-visual, proposta criada pela professora
11 min	Apresentação de vocabulário relacionado a fogo	Professor – classe	Fogo	Vocabulário	Escrito / livro didático, proposta criada pela professora
10 min	Relacionar palavras e definições Correção da atividade e explicação das respostas	Individual (5min) Professor – classe (5min)	Fogo	Vocabulário	Escrito / livro
6 min	Explicação sobre como descrever uma imagem	Professor − classe		Uso de estratégias	Oral, proposta criada pela professora
11 min	Descrever uma imagem	Trabalho em grupos	Fogo na aeronave / descrever uma ação em progresso	Produção oral	Visual / livro
15 min	Apresentar, oralmente, a descrição da imagem, Correção da atividade	Alunos- classe Professor − classe	Fogo na aeronave / descrever uma ação em progresso	Produção oral	Visual /livro
11 min	Comparar a descrição feita pelos grupos com a descrição do áudio	Professor − classe	Fogo na aeronave	Compreensão oral / leitura e Produção oral	Audio-visual, proposta criada pela professora
10 min	Ordenar as imagens de acordo com o áudio e relacionar perguntas e veículos	Individual	Corpo de bombeiros no aeroporto	Compreensão oral	Audio-visual / livro
5 min	Sublinhar a informação correta	Individual	Fogo a bordo	Compreensão oral	Áudio e escrito / livro

Conversamos, então, sobre as expressões de sequência, tais como *first of all*, *so*, *after that*, utilizadas pelos alunos durante a descrição da imagem. Ana explica que costuma esclarecer para os alunos o que é descrever uma imagem e o que é fazer um relato de uma história. Segundo ela, os alunos acabam misturando as duas tarefas. Nas suas palavras:

> contando a história (+) e ali eles estavam só descrevendo (+) é (+) então, eu gosto de diferenciar (+) ajudá-los a:a separar as partes né, (+) primeiro você descreve (+) aí você tem algumas palavras que vão te ajudar (+) alguns termos que vão te ajudar (+) na história esses aí são muito úteis (expressões de sequência) (+) mas na descrição eles acabam misturando muito quando estão descrevendo (+) aí depois vai contar a história acha que'(+) nossa, mas eu já falei isso (+) porque ele colocou lá na frente na hora de descrever né, (+) não precisa (+) eu tentei ajudar nesse sentido (Entrevista com a professora Ana, conduzida no dia 11 de julho de 2017).

O fato de a professora Ana trabalhar com os alunos a diferença entre essas duas tarefas em termos de demandas linguísticas reforça que existe uma preocupação com a preparação para o exame, indicando efeito retroativo, uma vez que são tarefas presentes no teste. Por outro lado, não foram explorados outros propósitos comunicativos em sala de aula, senão o de descrever estados e ações e de emitir opinião, o que sugere estreitamento do currículo.

Nessa aula, a professora Ana também focou no desenvolvimento da compreensão oral dos alunos. Três áudios de gêneros diferentes foram trabalhados, sendo que o último contemplava uma comunicação aeronáutica. Em termos de atividades, a professora utilizou as propostas contidas no livro didático, com exceção de quando pediu aos alunos que comparassem a descrição que eles haviam feito com a descrição do áudio. Nesse momento, presumi que Ana integraria a habilidade de compreensão oral à produção oral, o que seria considerado um tipo de efeito pretendido. No entanto, Ana decidiu projetar a transcrição do áudio, transformando a tarefa em leitura e produção oral.

Integrar a leitura à produção oral parece ser uma característica das aulas na EEAR, ao passo que a integração de habilidades orais é menos recorrente, limitada à colocação de perguntas à classe. Isso fica mais evidente quando Ana optou por não desenvolver com os alunos uma atividade do livro didático em que era solicitado a eles que criassem um diálogo entre um piloto e um controlador sobre uma suspeita de fogo a bordo da aeronave. Na entrevista, quando perguntei sobre a razão de não ter sido trabalhada essa atividade, Ana explicou que não vê muita importância, já que não é uma tarefa presente no exame nem na prova oral da 4ª série, em que os alunos têm que descrever figuras coladas em um pôster e contar uma história envolvendo-as. Nas suas palavras:

> não é muito explorada porque justamente a gente não vê muita importância dessa atividade (+) no Eplis (+) e até pra prova que eles vão fazer aqui também (+) então assim' (+) quando dá tempo (+) você até consegue / você FAZ (+) quando dá tempo eu faço (Entrevista com a professora Ana, conduzida no dia 11 de julho de 2017).

Como podemos perceber, a tarefa de gerenciamento do diálogo entre piloto e controlador, mesmo sendo considerada a mais representativa da comunicação aeronáutica – uma vez que permite ao aluno reproduzir o uso da língua em situações semelhantes às autênticas e a utilizar estratégias de negociação do sentido, é deixada em segundo plano, provavelmente, porque nem o EPLIS ou a prova oral da 4ª série, tampouco as avaliações escritas, a contemplam. Isso pode ser entendido como um efeito negativo que, por sua vez, precisa ser remediado.

Além da tarefa de descrição de imagem, Ana propôs mais uma tarefa de produção oral no início de sua aula com o propósito de emitir opinião. Todavia, o modelo de interação escolhido por ela, professor-classe, não propiciou a participação ativa de todos os alunos e somente alguns responderam às perguntas colocadas por ela. A apresentação de vocabulário também foi dominada por Ana, que pediu aos alunos que lessem em voz alta as definições das palavras projetadas no quadro. Essas características foram percebidas na prática de todos os professores e instrutor observados, o que distancia dos efeitos pretendidos em termos de abordagem de ensino.

O Quadro 37 retrata a aula assistida na turma C no dia 11 de julho de 2017. Essa aula foi conduzida de forma bastante semelhante à anterior. A professora iniciou com um vídeo e colocou algumas perguntas para serem respondidas oralmente pelos alunos. Em seguida, apresentou o vocabulário e solicitou que eles fizessem os exercícios do livro didático. Passou, então, para a etapa da descrição da imagem e propôs todas as atividades com foco na compreensão oral, conforme previsto no material. Embora Ana siga os exercícios prescritos, atividades extras foram inseridas por ela. Na parte de compreensão oral, a professora elaborou um exercício de múltipla escolha com as alternativas em português, aproximando-o do da fase I do EPLIS. Na entrevista, Ana confirma que os professores e instrutores, a pedido da coordenação do curso, têm trabalhado esse tipo de exercício em sala de aula a fim de familiarizar os alunos com o EPLIS e com a avaliação escrita da 4ª série. Assim, relata:

> essa atividade todo o instrutor tá aplicando pra acostumar o aluno com essa coisa de ouvir em inglês e ter que responder em português (+) porque na / isso acontece na 4ª série no Eplis quando eles fazem (+) parece que é assim né"(+) é assim (+) e aí a gente começou a preparar a prova da 4ª série também nesse layout (+) porque até a 2ª e 3ª a prova vem toda em inglês (Entrevista com a professora Ana, conduzida no dia 25 de julho de 2017).

Outra atividade proposta pela professora foi a criação de uma história com base na imagem que ilustra a unidade, em vez de solicitar aos alunos que fizessem o *role-play* previsto no material. Sobre isso, Ana admitiu que enfatiza mais em sala de aula a descrição de figura e o relato de histórias por causa do EPLIS: "é que a gente foca mais ali no:o *describing and telling the story* (+) por conta do Eplis (+) porque a gente sabe que no Eplis não tem esse *roleplay* (+) mas aí (+) pra *real life* né". Nessa fala, Ana pontua que o fato de o EPLIS não contemplar uma tarefa em que o candidato tem que se engajar em um diálogo afeta sua prática pedagógica, fazendo com que ela prefira focar nos aspectos exigidos pelo exame, mesmo sabendo que esse tipo de tarefa seria mais adequado ao ambiente de trabalho do futuro controlador.

Quadro 37 − Aulas observadas na 4ª série turma C

Professora: Ana	**Série: 4ª**	**Turma: C**	**No. Alunos: 14**	**Tempos de aula: 03**	**Data: 11/07/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
5 min	Responder, oralmente, perguntas colocadas pelo professor sobre um vídeo	Professor – classe	Condições meteorológicas adversas no pouso / descrever eventos em progresso	Produção oral /	Audio-visual, proposta criada pela professora
10 min	Apresentação de vocabulário relacionado a ventos	Professor – classe	Condições meteorológicas adversas no pouso	Vocabulário	Audio-visual, proposta criada pela professora
12 min	Relacionar palavras e definições Correção da atividade, explicação de novo vocabulário.	Individual / pares (6min) Professor – classe (6min)	Condições meteorológicas adversas	Vocabulário	Escrito / livro
15 min	Descrever uma imagem respondendo perguntas sobre ela	Trabalho em grupo (8min) Alunos – classe (7min)	Condições meteorológicas adversas no pouso / descrever uma ação em progresso	Produção oral /	Visual / livro
4 min	Comparar as respostas à descrição do áudio	Professor – classe	Condições meteorológicas adversas no voo	Compreensão oral / produção oral	Áudio, proposta criada pela professora
20 min	Selecionar as informações citadas no áudio e responder questões abertas Correção, explicação das respostas e vocabulário	Individual (8min) Professor – classe (12min)	Condições meteorológicas adversas no voo	Compreensão oral	Áudio e escrito / livro, Gênero: entrevista com um piloto
2 min	Responder questões de múltipla escolha em português sobre o áudio	Professor – classe	Condições meteorológicas adversas no voo	Compreensão oral (leitura) / Produção oral	Áudio e escrito, proposta criada pela professora
6 min	Listar os problemas meteorológicos enfrentados por pilotos durante um voo	Trabalho em grupo (3min) Alunos – classe (3min)	Condições meteorológicas adversas no voo	Vocabulário	Oral, proposta criada pela professora
5 min	Relacionar aeronaves e problemas enfrentados Relacionar trechos para formar sentenças do áudio	Individual (8min) Professor – classe (5min)	Condições meteorológicas adversas no voo	Compreensão oral	Escrito / livro, Gênero: RT
4 min	Completar um texto com as palavras	Individual / Pares (4min)	Condições meteorológicas adversas no voo	Vocabulário	Escrito / livro
17 min	Criar uma história com base em uma imagem	Trabalho em grupos	Condições meteorológicas adversas no voo / descrever eventos passados	Produção oral e escrita / leitura em voz alta	Visual, proposta criada pela professora

Perguntei, então, a Ana se ela costuma abordar outras tarefas do EPLIS como, por exemplo, a explicação de uma situação problemática de trafego aéreo e a proposta de uma solução para resolvê-la. Ana admitiu que, apesar de ter lido o manual do candidato, ainda não

conhece tão bem o teste e acrescentou que segue as recomendações que giram em torno de desenvolver as habilidades dos alunos em descrever imagens e criar histórias. Por fim, Ana comentou que seria bastante interessante incluir em suas aulas atividades dessa natureza e que estaria disposta se a coordenação, assim, recomendasse.

> eu só recebo as orientações por escrito de como tem que ser (+) ATÉ li o manual do Eplis (+) então, vem da orientação pedagógica /.../ se (coordenação) pegasse essa essa parte e trouxesse como sugestão pra gente inserir nas aulas (+) ficaria bem legal né, (+) trazer umas frases problemáticas (Entrevista com a professora Ana, conduzida no dia 25 de julho de 2017).

A fala de Ana, juntamente com o que foi observado em sua prática pedagógica, corrobora a percepção de que ainda existem lacunas a serem preenchidas em relação aos conhecimentos dos professores sobre o EPLIS e sobre a natureza da linguagem. Essas deficiências, somadas às limitações ainda existentes no teste quanto às funções comunicativas avaliadas, acabam enfraquecendo os efeitos pretendidos ao fazer com que somente alguns aspectos sejam privilegiados em detrimento de outros.

A terceira observação ocorreu no dia 25 de julho de 2017 na turma F. O Quadro 38 resume os eventos ocorridos. Essa aula iniciou-se com os alunos descrevendo uma imagem do livro que eles haviam trabalhado em grupo na aula anterior. Em seguida, Ana pediu a eles que comparassem sua descrição com a transcrição do áudio, integrando a leitura à produção oral, como havia feito com a turma E. Vemos que, nas turmas menos proficientes, Ana costuma alinhar a produção oral à leitura, ao passo que, nas turmas mais proficientes, no caso a turma C, é possível integrar as habilidades de compreensão e produção oral, sem passar pela leitura.

Na parte dedicada à compreensão oral, Ana novamente trouxe para a sala de aula um exercício de múltipla escolha com alternativas em português para ser respondido com base em um áudio de radiotelefonia, muito semelhante aos itens da fase I do exame. Nesse momento, Ana explicou aos alunos que havia feito a adaptação para tornar a tarefa similar ao EPLIS, com o objetivo de familiarizá-los com o exame. Essa foi a segunda vez que ela mencionou o teste nessa aula. A primeira ocorreu durante a descrição da imagem, quando Ana aconselhou os alunos a utilizarem estratégias para gerenciamento do tempo, uma vez que no EPLIS eles tinham apenas um minuto para desenvolverem essa tarefa.

Diferentemente das aulas anteriores, Ana decidiu propor aos alunos a criação de um diálogo entre piloto e controlador prevista no livro didático. Porém, fez questão de deixar claro que o foco da 4ª série seria na descrição de figuras e na criação de histórias e que eles estavam fazendo aquele tipo de exercício somente porque havia tempo disponível. No entanto, o padrão

de interação com que Ana conduziu o exercício, professor-classe, fez com que pouco tempo fosse dedicado à tarefa e não propiciou a participação de todos. Essa atitude de Ana mostra que a parte 4 da fase II do EPLIS apresenta efeito retroativo mais forte sobre sua prática pedagógica, levando-a dedicar mais tempo e esforço às tarefas de descrição de imagens e relato de histórias, respaldando os resultados de Gimenez (1999) quanto ao fato de alguns aspectos serem mais salientes do que outros.

Quadro 38 − Aulas observadas na 4ª série turma F

Professora: Ana	**Série: 4ª**	**Turma: F**	**No. Alunos: 14**	**Tempos de aula: 02**	**Data: 25/07/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
9 min	Descrever uma imagem a partir de perguntas	Alunos – classe	Incidentes durante a decolagem / descrever eventos em progresso	Produção oral	Visual, proposta criada pela professora
5 min	Comparar as respostas à transcrição do áudio	Professor – classe	Incidentes durante a decolagem	Compreensão oral (leitura) / Produção oral	Escrito, proposta criada pela professora
2 min	Responder questões com base na imagem	Professor – classe	Incidentes durante a decolagem / descrever eventos em progresso	Produção oral	Visual e escrito / livro
7 min	Sublinhar as informações corretas de acordo com o áudio	Individual − pares	Incidentes durante a decolagem	Compreensão oral	Escrito / livro Gênero: comunicado à imprensa
6 min	Completar as lacunas com informações do áudio	Individual − pares (5min) Professor − classe (10min)	Incidentes durante a decolagem	Compreensão oral	Áudio e escrito / livro
5 min	Selecionar a alternativa correta	Individual	Incidentes durante a decolagem	Compreensão oral	Áudio e escrito / livro, proposta criada pela professora, Gênero: RT
2 min	Relacionar trechos para formar sentenças	Professor – classe	Incidentes durante a decolagem	Compreensão oral	Áudio e escrito / livro, Gênero: RT
11 min	Criar um diálogo entre piloto e controlador	Professor – classe	Incidentes durante a decolagem – gerenciar o diálogo	Produção oral	Visual, Escrito / livro
9 min	Responder, oralmente, perguntas escritas	Trabalho em pares (3min) Alunos – classe (6min)	Incidentes durante a decolagem – emitir opinião	Produção oral	Escrito / livro

Outra característica da prática de Ana que chamou bastante a atenção foi o uso da tradução, algo que não foi observado com frequência nas aulas dos outros docentes. Quando questionada pelos alunos sobre o significado de certas palavras ou expressões, por inúmeras vezes, Ana preferiu traduzir os termos para o português do que explicá-los em inglês, o que ocorreu, com maior incidência na turma F, cujos alunos são menos proficientes. Vale pontuar

aqui que Ana, assim como os demais docentes, apresenta bom nível de proficiência em inglês, descartando a hipótese de que o uso de português de sua parte estaria ligado a problemas de proficiência. Na entrevista, Ana atribuiu essa característica tanto à baixa proficiência dos alunos quanto ao fato de as provas da 4ª série e o EPLIS utilizarem a transição do inglês para o português. Segundo a professora, a tradução facilita o entendimento e colabora com o desempenho nas provas. Assim, Ana descreve:

> dependendo do nível dos alunos também (+) essa turma que você assistiu hoje foi a eco e a golf (+) né, (+) fox e golf (+) então, eles eles parecem que (+) NECESSITAM né, dessa parte um pouquinho (+) não, não que (+) se não fosse cobrado assim na prova (+) eu acho que daria se a gente levasse todo em inglês (+) eu acho que dá (+) não é impossível não,
>
> você fala assim' (+) eu não vou ficar eh: substituindo inglês por inglês porque na prova ele vai ter o português (+) então, eu necessito / (+) eu traduzo né, (+) passo pro português (Entrevista com a professora Ana, conduzida no dia 25 de julho de 2017).

De fato, a mudança de código, do português para o inglês e do inglês para o português, é uma característica da comunicação aeronáutica no espaço aéreo brasileiro. Contudo, o controlador deve estar preparado para explicar, em inglês, o significado de um termo ou expressão, caso o piloto demonstre problemas no entendimento. Nas aulas observadas da professora Ana, não foram presenciadas atividades que tivessem como propósito o desenvolvimento da capacidade dos alunos de parafrasear, mesmo sendo avaliada pelo EPLIS na parte 2 da fase II.

A última aula assistida foi na turma A no dia 27 de julho de 2017, considerada a mais proficiente da 4ª série. O Quadro 39 retrata os principais eventos. A aula foi conduzida no mesmo padrão das anteriores. Constatamos, então, que Ana costuma iniciar suas aulas com um vídeo que ilustra o tópico da aula, colocando algumas questões a serem respondidas pelos alunos, no padrão professor-classe. Em seguida, Ana aborda o vocabulário trazido na unidade. Nas turmas menos proficientes, o vocabulário é, primeiramente, apresentado por Ana. Já nas turmas mais proficientes como a Alfa, Ana tende a delegar aos próprios alunos essa tarefa. Passa-se, então, à descrição da imagem, compara-se a descrição feita pelos alunos com a do áudio, baseando-se, geralmente, na sua transcrição. Em seguida, Ana, dá início às atividades de compreensão oral. Para isso, utiliza os exercícios previstos no livro didático e insere um exercício de múltipla escolha em português para familiarizar os alunos com formato da fase I do EPLIS. Para finalizar a aula, Ana costuma pedir aos alunos para criarem uma história com base na imagem descrita no início da aula.

Quadro 39 − Aulas observadas na 4ª série turma A

Professora: Ana	**Série: 4ª**	**Turma: A**	**No. Alunos: 15**	**Tempos de aula: 03**	**Data: 27/07/2017**
Duração (minutos)	**Tipo de atividade**	**Padrão de interação**	**Conteúdo (tópicos e funções)**	**Habilidades e aspectos linguísticos**	**Tipo de material**
5 min	Responder perguntas colocadas pelo professor sobre um vídeo	Professor – classe	Colisões no solo / emitir opinião	Produção oral	Vídeos, proposta criada pela professora
12 min	Relacionar palavras e definições Correção da atividade, explicação de novo vocabulário.	Individual / pares (5min) Professor – classe (7min)	Colisões no solo	Vocabulário	Escrito / livro didático
15 min	Responder perguntas colocadas pelo professor sobre algumas imagens	Professor – classe	Colisões no solo / descrever eventos em progresso	Produção oral	Imagens, proposta criada pela professora
8 min	Descrever uma imagem	Trabalho em pares	Colisões no solo / descrever eventos em progresso	Produção oral	Proposta criada pela professora
7 min	Responder perguntas colocadas pelo professor sobre a imagem	Alunos – classe	Colisões no solo / descrever eventos em progresso	Produção oral	Escrito / livro
4 min	Comparar as respostas à descrição do áudio	Professor – classe	Colisões no solo	Compreensão oral (leitura) / Produção oral	Proposta criada pela professora
4 min	Responder questões com base no áudio	Individual	Colisões no solo	Compreensão oral	Escrito, áudio / livro, Gênero: conversa entre controladores
5 min	Selecionar a alternativa correta	Individual	Colisões no solo	Compreensão oral	Escrito, proposta criada pela professora
2 min	Relacionar aeronaves a eventos	Individual	Colisões no solo	Compreensão oral	Escrito, áudio / livro, Gênero: RT
3 min	Ler em voz alta trechos escritos	Alunos – classe	Melhorias na comunicação aeronáutica	Leitura em voz alta	Escrito, proposta criada pela professora
10 min	Criar uma história com base em imagens	Trabalho em grupos	Colisões no solo / descrever eventos passados	Produção oral	Visual, proposta criada pela professora
15 min	Contar a história em um minuto e meio	Alunos – classe (10min) Professor – classe (5min)	Colisões no solo / descrever eventos passados	Produção oral	Proposta criada pela professora
4 min	Completar um texto	Individual	Colisões no solo	Vocabulário	Escrito / livro

Nessa turma, no entanto, algo um pouco diferente sucedeu. Ana trouxe para a sala de aula imagens reais envolvendo acidentes e incidentes aeronáuticos para serem trabalhadas pelos alunos, semelhante ao o que ocorre na fase II do EPLIS, em que, geralmente, é apresentada ao candidato uma foto de um evento verdadeiro. Na correção da atividade de relato de histórias dessas imagens, Ana explicou à turma que no EPLIS é concedido um minuto e meio para a realização dessa tarefa e, por essa razão, ele iria marcar o tempo de forma que eles adequassem sua história ao tempo disponível. Isso demonstra mais uma vez que Ana tem se mobilizado na preparação dos alunos para o exame. Na entrevista, ela discorre sobre isso:

> marquei (o tempo) (+) marquei (+) o meu relógio é meio ruim mas eu marquei (+) só pra eles terem uma noção (+) porque é uma coisa que eles se preocupam (+) no teste (+) na hora do Eplis (+) é um minuto e meio (+) será que eu tô falando em um minuto e meio" (+) então, eu tento / eu marco pra eles sim (Entrevista com a professora Ana, conduzida no dia 27 de julho de 2017).

De modo geral, a preparação que Ana oferece em relação à compreensão oral pode ser considerada de natureza variada por incluir diversos formatos e gêneros. Embora haja, claramente, uma influência forte do material didático, que dita a maioria das atividades trabalhadas por Ana, a inclusão de exercícios de múltipla escolha em português demonstra evidência de efeito retroativo. No entanto, quanto à produção oral, a ênfase dada às tarefas de descrição de imagens e criação de histórias, juntamente com a supressão de tarefas de produção oral diferentes a essas, apontam para uma preparação em mesmo formato, o que deveria ser evitado.

Outro problema diz respeito a alguns aspectos do exame serem ignorados ou pouco explorados por Ana como, por exemplo, explicar uma situação de tráfego aéreo e oferecer uma solução. Podemos dizer, então, que, em termos de produção oral, há um estreitamento do currículo da 4ª série que gira em torno de duas funções apenas (descrição de imagens e relato de histórias), o que pode estar relacionado às poucas oportunidades de capacitação oferecida aos professores e instrutores em relação ao exame.

Essas tarefas são, geralmente, feitas em pares ou em grupos, promovendo, assim, a interação entre os alunos e aumentando seu tempo de fala. No entanto, esse tipo de interação diverge daquela encontrada no contexto profissional, pois não envolve negociação de sentido para a solução de problemas, o que pode estar associado a um efeito negativo. Sobre isso, pouquíssimas oportunidades de integração das habilidades de compreensão oral e produção oral foram observadas. O gerenciamento do diálogo entre piloto e controlador e o ensino de estratégias para evitar mal-entendidos foram pouco trabalhados, a que Ana atribuiu, exatamente, ao fato de o EPLIS não contemplar essas tarefas. Outro tipo de exercício proposto pela professora para a prática de produção oral é a colocação de perguntas para serem discutidas oralmente. Porém, o padrão professor-classe, normalmente escolhido para conduzir esse tipo de atividade, não envolve todos os alunos, resultando em pouca participação.

Na comparação entre a 4ª série e as demais (2ª e 3ª séries), percebemos que na 4ª série o exame é mais mencionado e alguns de seus aspectos são discutidos com os alunos, o que não ocorre nas demais séries observadas, corroborando os resultados quantitativos e as percepções dos alunos no grupo focal e dos professores na entrevista. Os alunos da 4ª série, consequentemente, parecem demonstram conhecer melhor o exame. Além disso, tarefas

similares às do teste são mais recorrentes, tais como de múltipla escolha em português para desenvolvimento da compreensão oral, descrição de imagens e criação de histórias. Também são observadas mais oportunidades de produção oral na 4ª série. Essas constatações corroboram as suposições de que o efeito retroativo do exame se intensifica com a proximidade de sua realização, tornando-se mais explícito próximo à data de sua realização.

Quanto ao nível de proficiência dos alunos, percebemos algumas diferenças na abordagem de ensinar de Ana. Nas turmas menos proficientes, o ensino de vocabulário é mais centralizado na professora, a tradução e a leitura em voz alta estão mais presentes, alguns alunos tendem a interagir em português durante as atividades em grupo ou em pares, e as propostas de produção oral requerem o auxílio da escrita, fazendo com que o tempo destinado a essas atividades seja reduzido. Esses resultados também confirmam as hipóteses suscitadas no estudo quantitativo e os comentários de alunos e professores de que nas turmas mais proficientes o impacto do teste é maior, com mais oportunidades de interação e mais atividades de produção oral.

5.3.5 Diferenças na prática pedagógica dos professores e instrutor

No estudo quantitativo, notamos algumas diferenças expressivas no modo como professores e instrutores e profissionais mais e menos experientes percebem a influência do exame no contexto escolar e, mais especificamente, na sua prática pedagógica. Os instrutores reportaram níveis mais altos de conhecimentos sobre o exame, assim como aqueles com mais experiência no ensino de inglês aeronáutico. Já, as respostas dos professores e daqueles considerados menos experientes variaram um pouco mais, indicando inconsistências em relação à consolidação dos conhecimentos sobre o exame e sobre inglês aeronáutico nesses grupos. Nas entrevistas com os professores e instrutor observados, o professor Eduardo, considerado mais experiente, e a professora Ana, menos experiente, afirmaram que gostariam de conhecer melhor o exame, evidenciando que ainda existem falhas na capacitação quanto ao exame e ao inglês aeronáutico. Em sala de aula, tanto Ana quanto Eduardo exploraram algumas funções comunicativas presentes no teste em detrimento de outras, o que pode estar relacionado a deficiências nesses conhecimentos. Ana, por exemplo, em uma das entrevistas, quando perguntada sobre a razão de não abordar outras tarefas do exame em sala de aula, revela que não conhecia o teste muito a fundo, mas estaria disposta a implementar outros tipos de atividades caso recebesse orientação para isso. Esses óbices necessitam ser remediados para que o repertório de funções comunicativas trabalhadas em sala de aula seja expandido.

Sobre a abordagem do exame em sala de aula, os alunos disseram que os instrutores discutem aspectos do exame com maior frequência e são mais procurados pelos alunos para solucionar dúvidas sobre o teste. Visão essa também compartilhada pelo instrutor João que, por sua vez, explicou que os alunos costumam procurá-lo por causa do EPLIS, devido à percepção de que, por ser controlador e ter realizado o exame, ele teria condições de falar mais profundamente a seu respeito. No entanto, nas aulas assistidas, notamos que o exame foi pouquíssimo comentado nas 2ª e 3ª séries do curso, seja pelos professores ou pelo instrutor observados. Na realidade, somente o professor Eduardo mencionou o EPLIS em uma turma da 3ª série, evidenciando que o exame é, geralmente, pouco explorado nas séries que antecedem a 4ª. Já, na 4ª série, a professora Ana citou o teste várias vezes e explicou algumas de suas características. Como expus, acredito que a decisão de tratar o EPLIS de forma mais clara somente no último semestre retarda e diminui a intensidade de seu efeito e deve ser repensada.

Os professores e instrutores mais experientes também reportaram, no questionário, sofrerem menos influência do exame em sua prática pedagógica e no conteúdo abordado do que aqueles considerados menos experientes. Nas observações de aulas, encontrei evidências de efeito retroativo na prática de ambos os grupos, não sendo possível apontar qual é mais ou menos influenciado. A professora e coordenadora Pilar, por exemplo, é considerada a mais experiente entre os professores e suas aulas são bastante afetadas pelo exame, principalmente, nas turmas mais proficientes, nas quais é mais facilitadora, promove mais oportunidades de interação entre os alunos, varia mais as atividades de produção oral e explora diversas funções comunicativas. Em termos de abordagem de ensino, vimos que os professores e o instrutor, de modo geral, são mais estruturalistas nas turmas menos proficientes, principalmente, no que diz respeito ao ensino de vocabulário, demonstrando um papel bastante controlador e domínio do turno de fala, o que não condiz com os efeitos pretendidos e evidencia que o teste por si só não é capaz de provocar as mudanças esperadas.

Outra suposição levantada foi a de que os professores e aqueles profissionais com mais experiência no ensino de inglês aeronáutico trabalham a leitura em voz alta em sala de aula com mais frequência do que os instrutores e os menos experientes. Nas aulas assistidas, a leitura em voz alta pelos alunos estava presente na prática de todos os professores e instrutor, principalmente, nas turmas menos proficientes, com o propósito de correção da atividade em andamento. Os alunos menos proficientes, por exemplo, recorriam à leitura em voz alta de suas anotações em vez de tentarem se expressar espontaneamente. Também observei situações em que era solicitado aos alunos que lessem em voz alta trechos escritos, com o objetivo de focar a atenção de todos no vocabulário e de engajá-los em uma interação professor-classe nas

práticas de Eduardo, João e Ana, mesmo sendo o primeiro mais experiente. Essas constatações não me permitem afirmar que existem diferenças no modo como professores e instrutores utilizam a leitura em voz alta em sala de aula.

Quanto às atividades de produção oral, os professores responderam abordar com mais frequência do que os instrutores as tarefas de descrição de figuras, explicação de situações de tráfego aéreo e oferta de soluções. Nas observações , a descrição de imagens foi prática bastante comum nas aulas dos professores, bem como a criação de histórias. Porém, as tarefas de explicação e oferta de sugestão para situações de conflito foi observada somente na prática da professora e coordenadora Pilar, que busca variar bastante as atividades de produção oral em comparação com os demais. No caso do instrutor João, nenhuma dessas tarefas foi trabalhada com os alunos nas aulas assistidas. João optou por enfatizar bastante a apresentação oral de assuntos relacionados à aviação, com o propósito de explorar a função comunicativa de "emitir opinião" também contemplada no EPLIS, o que não significa que ele não aborde essas tarefas em sala de aula, sendo necessárias mais observações de aulas para afirmar algo dessa natureza.

5.4 Resumo do capítulo

Neste capítulo, apresentei os resultados do estudo qualitativo sobre o efeito retroativo do EPLIS nas percepções, atitudes e ações de professores, instrutores e alunos, com o objetivo de explicar, expandir, refinar e corroborar os resultados alcançados na fase quantitativa. Constatamos que, embora alunos e professores tenham afirmado no grupo focal e nas entrevistas que o EPLIS é mencionado em sala de aula, somente na 4ª série o exame é citado com maior frequência e alguns de seus aspectos discutidos, porém sem muito aprofundamento. Isso corrobora a hipótese de que a 4ª série possui níveis mais altos de conhecimentos sobre o teste. Nas turmas mais proficientes de todas as séries observadas, foram percebidas mais oportunidades de produção oral com propósitos comunicativos semelhantes aos do exame, demonstrando que o seu efeito é mais intenso nesses grupos, o que confirma os resultados quantitativos e as percepções de alunos, professores, instrutor. Também foi constatado que impacto do teste é mais explícito na 4ª série do curso. Nessa série, a ênfase nas habilidades de compreensão oral e produção oral é maior e duas tarefas presentes no exame foram bastante exploradas em sala de aula: a descrição de figuras e a criação de histórias. Por outro lado, verificamos que as avaliações de rendimento são mais relevantes do que o EPLIS devido às consequências que seus resultados acarretam, afetando bastante a prática do professor,

principalmente, em relação ao conteúdo ministrado. Ainda sobre a 4ª série, em relação à natureza da preparação para o EPLIS, notamos que, em termos de produção oral, a preparação tende a ser mais focada nas tarefas da parte 4 da fase II, estreitando as possibilidades de outras atividades serem realizadas em sala de aula. Outro ponto bastante importante foi a constatação de efeito negativo no contexto escolar. O fato de o exame não privilegiar o gerenciamento da comunicação entre piloto e controlador fez com que atividades dessa natureza fossem ignoradas ou pouco trabalhadas em sala de aula pelos professores e instrutor observados, mesmo sendo elas consideradas mais representativas das situações reais de uso da língua no contexto profissional.

CAPÍTULO VI

CONCLUSÕES E CONTRIBUIÇÕES DO ESTUDO

6.1 Considerações iniciais

Neste capítulo, retomo, primeiramente, as perguntas de pesquisa que guiaram este estudo e apresento os resultados obtidos. Proponho, então, um diagrama para explicar o efeito retroativo do EPLIS no cenário investigado. Em seguida, discorro sobre as contribuições teóricas, metodológicas e pedagógicas da pesquisa, relacionando-as com resultados de estudos anteriores, e discuto as implicações para a (re)elaboração e implementação de políticas linguísticas. Abordo, por fim, as limitações do estudo e as dificuldades encontradas, oferecendo sugestões e direcionamento para pesquisas futuras.

6.2 Perguntas de pesquisa revisitadas

Esta pesquisa teve como objetivo principal investigar o efeito retroativo do EPLIS nas percepções, atitudes e ações de professores de inglês e alunos do Curso de Formação em Controle de Tráfego Aéreo da EEAR. Em resposta às três perguntas de pesquisa, apresento um resumo das evidências geradas pela triangulação dos dados obtidos de diversas fontes. As duas primeiras perguntas tiveram como foco os efeitos do EPLIS nas percepções e atitudes de professores e alunos, respectivamente. A terceira pergunta, por sua vez, abordou o impacto do exame nas ações em sala de aula.

6.2.1 Pergunta de pesquisa 1

Qual é o efeito retroativo do EPLIS nas percepções e atitudes dos professores de língua inglesa do Curso de Formação em Controle de Tráfego Aéreo?

O estudo quantitativo sobre as percepções e atitudes de professores e instrutores indicou que o EPLIS exerce influência sobre a abordagem de ensino em sala de aula, os tópicos e habilidades abordados, os tipos de tarefas desenvolvidos e as avaliações realizadas. Essa influência parece ter sido intensificada após a implementação do exame no contexto escolar e

das benfeitorias em sua divulgação, o que foi confirmado pela análise qualitativa. Os professores e instrutor afirmaram que, se não fosse pela introdução do EPLIS na escola, eles não considerariam desenvolver ou enfatizar certas tarefas com os alunos, o que caracteriza, por excelência, o efeito retroativo. Eles também apontaram que o impacto é mais intenso na 4ª série, quando os alunos estão próximos da realização do teste. Além disso, reportaram que as mudanças no sistema avaliativo da EEAR ocorridas foram decorrentes do teste, tais como a predileção por áudios de radiotelefonia e a elaboração de itens de múltipla escolha em português para avaliar a compreensão oral dos alunos da 4ª série, bem como o desenvolvimento de provas orais para todas as séries.

Em relação às habilidades, foram encontrados indícios nos dados numéricos de que atividades de compreensão de textos eram mais recorrentes em sala de aula do que atividades com foco na produção oral, na perspectiva dos alunos, o que não era esperado em termos de efeitos pretendidos, dado que o EPLIS não avalia leitura. Supomos, então, que essa percepção poderia estar ligada à influência dos livros didáticos ou das avaliações escritas. No entanto, o estudo qualitativo não confirmou essa alegação. Professores e coordenadora declararam que as habilidades orais são as mais exploradas em sala de aula em razão do EPLIS e da natureza da comunicação aeronáutica.

Em relação às provas orais das 2ª e 3ª séries, no estudo qualitativo, os participantes pontuaram diversos aspectos de convergência entre elas e o EPLIS. Além de haver correspondência nos tópicos avaliados e em algumas tarefas, os professores reportaram utilizar a tabela de níveis de proficiência da OACI, denotando provável efeito retroativo. Contudo, diferentemente do EPLIS, essas avaliações não promovem a integração das habilidades de compreensão e produção orais. Somado a isso, o fato de os critérios serem avaliados isoladamente e em partes distintas da prova (pronúncia, estrutura, vocabulário, fluência e interação) sugerem problemas de entendimento da natureza da linguagem e do exame em si. A prova oral da 4ª série também avalia a produção oral dos alunos, porém em apenas duas tarefas relacionadas à parte 4 da fase II do EPLIS. Presumimos que, se por um lado, essa sobreposição de tarefas atesta efeito retroativo, por outro lado, a chance de os professores e instrutores enfatizarem somente esses aspectos em sala de aula pode ser aumentada, o que representaria um estreitamento do currículo e dos métodos de instrução.

A análise estatística também revelou diferenças entre professores e instrutores e profissionais com mais e menos experiência no ensino de inglês aeronáutico, em relação à influência do EPLIS em suas percepções e atitudes. Os instrutores alegaram conhecer melhor o exame e serem procurados com mais frequência pelos alunos para sanar dúvidas do que os

professores. Estes, por sua vez, apresentaram maior variação em suas respostas, indicando que ainda existem lacunas a serem preenchidas em relação aos conhecimentos que possuem sobre o teste. Também foi constatado que quanto mais experiente na área de ensino de inglês aeronáutico, mais conhecimentos professores e instrutores afirmam ter sobre o exame.

O estudo qualitativo confirmou que, para os alunos, os instrutores de inglês aeronáutico conhecem melhor o exame e a natureza da linguagem avaliada e, por essa razão, tendem a discutir e a explorar mais seus aspectos em sala de aula, sugerindo que a intensidade do efeito retroativo varia de acordo com os níveis de conhecimentos das demandas do exame. Em relação aos professores, os alunos disseram que alguns se igualam aos instrutores de inglês aeronáutico, porém outros parecem ainda possuir conhecimentos limitados sobre o que é o exame e suas características. Isso pode ser explicado pela escassez de oportunidades de capacitação a respeito do EPLIS oferecidas a esses profissionais, o que foi constatado nas respostas ao questionário e confirmado nas entrevistas.

Os resultados quantitativos e qualitativos mostraram que o EPLIS exerce impacto positivo entre professores e instrutores. Eles o veem como um instrumento de motivação na aprendizagem de língua inglesa e, sobretudo, como fator importante na elevação do nível de proficiência dos alunos, favorecendo a ocorrência de efeito retroativo. Os professores apresentam atitudes mais positivas em relação ao teste do que os instrutores ao avaliarem sua autenticidade e a confiabilidade de seus resultados. No entanto, uma parcela considerável de professores preferiu não opinar, o que pode indicar problemas no entendimento desses participantes quanto à natureza da linguagem na comunicação aeronáutica ou sobre o exame em si, conforme já reportado por eles. Quanto ao efeito negativo, na percepção dos professores e instrutores, os alunos parecem ser os mais afetados em termos de ansiedade e pressão, porém não atribuem níveis elevados, o que pode indicar que o exame não carrega alta relevância para esses participantes.

Em suma, professores e instrutores disseram que, de modo geral, o efeito retroativo do teste nas suas aulas tem aumentado desde sua introdução no cenário escolar. Eles reportaram que o exame influencia os temas, habilidades, tarefas e avaliações em sala de aula, principalmente, na 4ª série do curso. Esse impacto parece ser maior para aqueles que atribuem níveis mais altos aos conhecimentos que possuem sobre o teste. Contudo, as carências relatadas em relação ao entendimento sobre o exame e a natureza da linguagem pelos professores parecem interferir no processo de efeito retroativo.

6.2.2 Pergunta de pesquisa 2

Qual é o efeito retroativo do EPLIS nas percepções e atitudes dos alunos do curso?

A análise quantitativa das percepções e atitudes dos alunos revelou que esses participantes percebem o efeito retroativo nas atividades em sala de aula e extraclasse aumentando progressivamente em função da proximidade do teste. Na 1ª série do curso, o impacto do exame no currículo foi considerado fraco ou até mesmo nulo. Os alunos explicaram que os tópicos trabalhados nessa série não mantêm relação com o controle de tráfego aéreo e, embora haja ênfase no desenvolvimento da compreensão oral, os áudios apresentados não condizem com as situações de uso da língua no ambiente de trabalho, o que foi confirmado pelos professores nas entrevistas.

Também foi constatado, nos dados quantitativos e qualitativos, que atividades de produção oral semelhantes às do exame são mais exploradas na 4ª série. Segundo os alunos, nas 2ª e 3ª séries, atividades dessa natureza começam a ser um pouco mais frequentes, principalmente, nas aulas de instrutores e professores que conhecem melhor o EPLIS, respaldando também a percepção de que o efeito retroativo é mais intenso na prática pedagógica daqueles que possuem mais conhecimentos sobre suas demandas. Já na 4ª série, os alunos pontuaram que o trabalho com essas tarefas é intensificado, o que foi corroborado pela coordenadora ao afirmar que nessa série é feita uma maior familiarização dos alunos para com o EPLIS.

Os dados numéricos também mostraram que os alunos percebem que a compreensão oral de conversas de radiotelefonia é bastante enfatizada em sala de aula, assim como a compreensão de textos escritos relacionados à aviação e ao controle de tráfego aéreo, coincidindo com os dados estatísticos dos professores e instrutores, porém em contraposição ao que foi relatado pela coordenadora. A percepção dos alunos de que a produção oral é pouco desenvolvida nas séries iniciais vai de encontro aos efeitos pretendidos, como foi pontuado anteriormente. Na fase qualitativa da pesquisa, descobrimos que as avaliações de rendimento escritas realizadas em sala de aula são consideradas mais relevantes do que o EPLIS pelas consequências graves a que seus resultados estão sujeitos, tais como desligamento do curso e baixa prioridade na escolha de vagas. Essa constatação pode explicar o fato a leitura ser vista como bastante recorrente em sala de aula, ao lado da compreensão oral, dado que elas avaliam essas duas habilidades na mesma proporção. Além disso, os alunos disseram que o formato dessas provas propicia o uso de estratégias para responder testes, consideradas impróprias pelo

fato de inflarem os resultados sem a devida correspondência em termos de proficiência, indicando problemas ligados ao letramento em avaliação por parte de seus elaboradores.

Sobre as atividades preparatórias fora da sala de aula, o estudo quantitativo apontou que a proximidade do EPLIS interfere na frequência com os alunos da 4ª série buscam se engajar em tarefas no mesmo formato do exame. Contudo, de modo geral, essa frequência foi considerada baixa, o que pode também estar relacionado ao fato de o exame apresentar relevância mais baixa em relação às provas escritas e, consequentemente, efeito retroativo menor nas atividades extraclasse, hipótese essa corroborada pela análise qualitativa das percepções de alunos e professores.

A proximidade do exame também pareceu interferir nas atitudes dos alunos para com o exame. Os resultados quantitativos indicaram que os alunos da 4ª série são os que atribuem níveis mais baixos à significância do EPLIS e ao alinhamento entre ensino e avaliação externa. A diminuição no nível de relevância do exame pode estar relacionada ao fato de as avaliações escolares implicarem consequências mais graves para esses alunos. Quanto ao alinhamento entre o ensino e o exame, os comentários dos alunos revelaram crenças sobre o que é se preparar para um exame. Para os alunos, a preparação deve estar no mesmo formato do teste. No estudo qualitativo, a coordenadora confirmou que os alunos costumam pontuar que as aulas deveriam seguir o modelo adotado no exame, o que, na visão deles, elevaria os resultados. Essa percepção dos alunos sugere que a preparação para o EPLIS na EEAR recebe um carácter mais variado, o que reduziria a possibilidade de estreitamento do currículo e dos métodos de instrução como efeito negativo.

Os conhecimentos que os alunos possuem sobre o EPLIS também variaram em função de sua proximidade. Os resultados estatísticos mostraram que os alunos da 1ª série são aqueles que menos sabem sobre o exame, ao passo que os da 4ª série já apresentam níveis mais altos de conhecimentos, inclusive específicos. Na análise qualitativa, os alunos confirmaram que, embora muitos professores mencionem o teste desde o início do curso, detalhes são mais bem explorados somente na 4ª série do curso, evidenciando, novamente, que o efeito é mais intenso nessa série.

A proficiência declarada pelos alunos é outro fator que parece intermediar o efeito retroativo do EPLIS. O estudo quantitativo apontou que os alunos considerados mais proficientes apresentam mais conhecimentos sobre o exame e se engajam com mais frequência em atividades preparatórias dentro e fora da sala de aula. Isso sugere que os professores tendem a mencionar e a discutir aspectos concernentes ao EPLIS com maior frequência em turmas mais proficientes, o que também foi confirmado pelos alunos no grupo focal. Na fase qualitativa, os

professores e instrutor observados reportaram que, de fato, nessas turmas, é possível atribuir mais atividades de produção oral de natureza livre similares àquelas encontradas no teste, bem como desenvolver estratégias orientadas para a realização dessas atividades, aumentando, assim, o efeito retroativo para esses alunos.

Somado a isso, a análise quantitativa forneceu indícios de que os alunos mais proficientes se sentem mais motivados a estudar inglês, apresentam atitudes mais positivas, percebem melhor o alinhamento entre ensino e avaliação externa e atribuem mais importância à aprovação. No estudo qualitativo, os professores e instrutor entrevistados afirmaram que veem diferenças nas atitudes dos alunos para com o exame em função do nível de proficiência. Para eles, os alunos mais proficientes demonstram mais confiança e motivação, enquanto os menos proficientes parecem mais apáticos. Isso pode significar que o nível de dificuldade do teste é visto pelos alunos como desafiador, provocando um efeito retroativo mais intenso para os mais proficientes que, ao perceberem chances reais de aprovação, se mobilizam em busca de mais informações e maneiras de se prepararem.

Encontramos, também, na fase quantitativa da pesquisa, diferenças significativas no modo como alunos com mais ou menos conhecimentos sobre o exame se preparam fora do ambiente de sala de aula. Os resultados qualitativos indicaram que, nas percepções dos alunos, saber mais sobre o exame é essencial para a conscientização de sua importância e para o engajamento em atividades preparatórias extraclasse. Nesse sentido, os alunos da 4ª série, por possuírem mais informações sobre o teste, procuram desenvolver atividades preparatórias focadas com maior frequência, indicando que o efeito é maior para aqueles que conhecem suas demandas.

Em suma, os alunos percebem o efeito retroativo no contexto escolar em função da proximidade, nível de proficiência e conhecimentos construídos sobre o exame. O efeito retroativo nos tópicos, habilidades e tarefas em sala de aula é considerado fraco na 1ª série, tornando-se mais evidente nas 2ª e 3ª séries, principalmente, nas práticas dos professores e instrutores que detêm mais conhecimentos sobre o exame. Na 4ª série, o impacto do teste alcança seu ápice, influenciando as habilidades, tarefas e avaliações desenvolvidas em sala de aula. Em termos de relevância, os alunos pontuaram que os resultados das provas escritas que eles realizam na escola são mais importantes e, por isso, inferimos que esses testes também exercem influência sobre as aulas, juntamente com o EPLIS. Os alunos mais proficientes atribuem uma intensidade maior ao impacto do teste nas atividades realizadas dentro e fora do ambiente escolar. Eles se sentem mais motivados, apresentam atitudes mais positivas para com o exame externo e atribuem níveis mais altos de relevância a ele. A intensidade do efeito

também parece ser maior para aqueles que consideram conhecer melhor as demandas do teste. Eles reportaram que realizam mais atividades de produção oral em sala de aula e se dedicam mais à preparação extraclasse para o teste.

6.2.3 Pergunta de pesquisa 3

Qual é o efeito retroativo do EPLIS nas ações de alunos e professores em sala de aula?

A análise das aulas observadas confirmou que, de fato, o EPLIS exerce influência sobre os eventos que ocorrem nas aulas das 2ª, 3ª e 4ª séries, porém alguns fatores tais como a proximidade, a proficiência dos alunos, os conhecimentos sobre o teste e a alta relevância das provas de rendimento escritas interferem no processo de efeito retroativo. Primeiramente, confirmamos que o impacto do EPLIS é mais explícito e intenso na 4ª série do curso, quando se mencionam e se discutem aspectos do teste mais explicitamente e frequentemente em sala de aula. Nessa série, as habilidades de compreensão e produção orais são bastante enfatizadas, condizendo com os efeitos pretendidos. Contudo, carências em relação aos conhecimentos sobre o exame e à natureza da linguagem fazem com que somente alguns aspectos do EPLIS ganhem proeminência em detrimento de outros, levando a um suposto estreitamento do currículo e dos métodos de instrução e a uma preparação mais voltada para o formato do teste.

O efeito retroativo na abordagem de ensino do professor também se altera em função da proficiência dos alunos. As observações de aulas revelaram que, nas turmas menos proficientes, os professores e instrutor tomam uma postura mais controladora e centralizadora, com maior tempo de fala, oferecendo poucas atividades em pares ou em grupo com foco na produção oral. A interação dominante nessas turmas é a de professor-classe, o que reduz a participação dos alunos e faz com que seus turnos de fala sejam mais curtos, afastando-se, assim, dos efeitos pretendidos. Embora, em turmas mais avançadas, também ainda sejam encontradas características estruturalistas na prática pedagógica, os professores e instrutor atuam mais como facilitadores, existem mais oportunidades de interação aluno-aluno e os turnos de fala dos alunos são mais longos, atendendo, assim, ao que se espera em termos de impacto desejado.

Diferenças nos níveis de conhecimentos dos professores e instrutor sobre o teste e a experiência no ensino de inglês aeronáutico também influenciam o tipo de preparação promovida em sala de aula. Nas aulas da professora e coordenadora Pilar, por exemplo, encontramos evidências de efeito retroativo que indicam uma preparação em formato mais

variado, considerada benéfica, por não estreitar os métodos de ensino. É importante ressaltar que Pilar é uma das professoras mais experientes, que liderou a reestruturação do curso de inglês para BCT na EEAR, e vem construindo muitos conhecimentos sobre controle de tráfego aéreo e sobre a natureza da comunicação aeronáutica.

A alta relevância atribuída pelos professores, instrutor e alunos às provas escritas também afetam o efeito retroativo. As observações feitas nas 2ª e 3ª séries corroboraram que as habilidades de compreensão oral e leitura são mais exploradas em sala de aula. A compreensão oral é, na maior parte das vezes, trabalhada isoladamente, indicando um impacto maior da fase I do EPLIS. Raramente, é proposta a integração das habilidades de compreensão e produção orais, encontrada na fase II do exame. Podemos, então, concluir que, embora encontremos diversas tarefas que se assemelham às da fase II do EPLIS, nessas séries, a fase I do exame exerce maior influência nos eventos em sala de aula, pelo fato de as provas escritas focarem nas habilidades de compreensão oral e leitura.

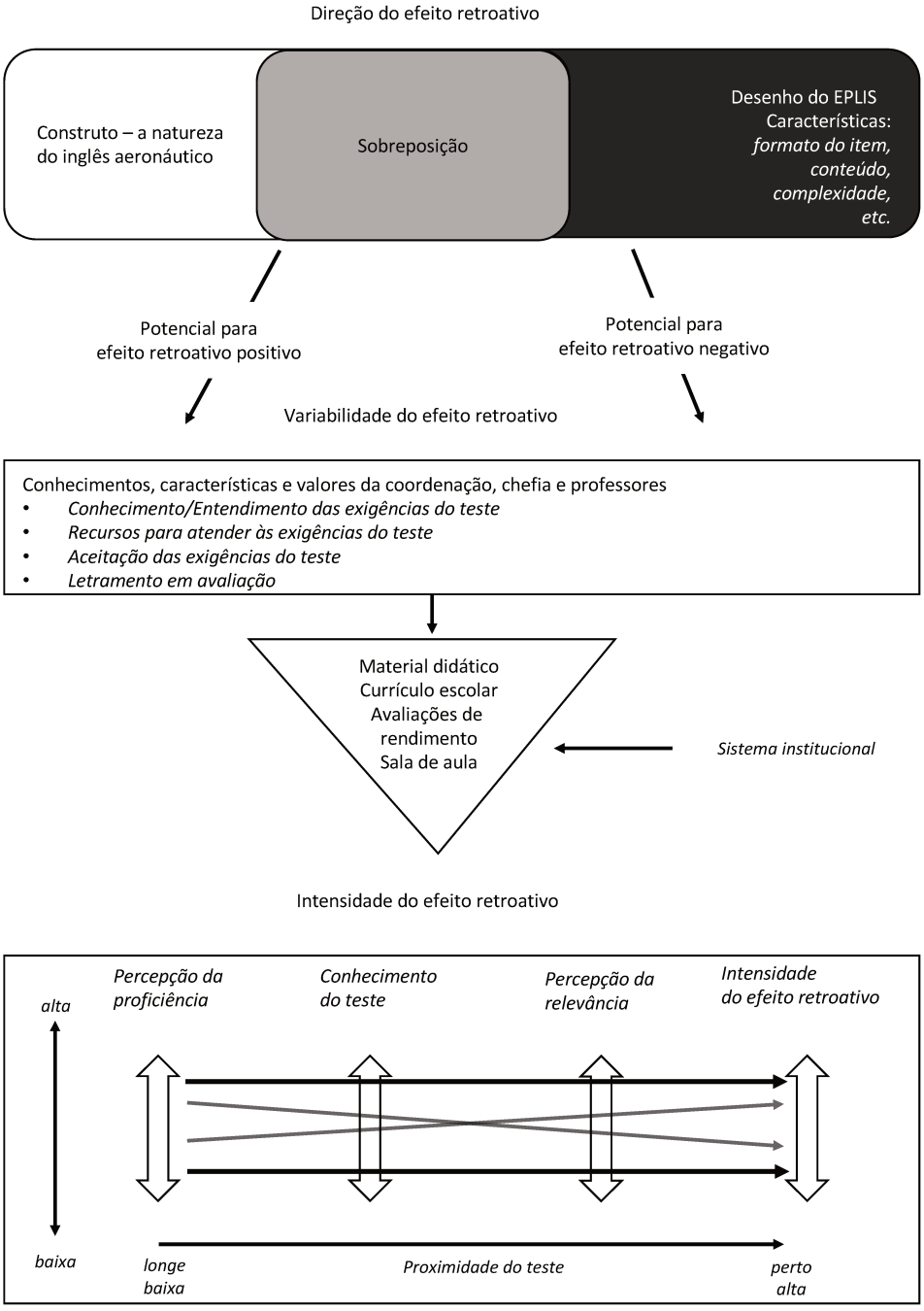
Na 4ª série, os resultados mostraram evidências ainda mais contundentes de um forte impacto do formato da fase I do teste, com tarefas de múltipla escolha sendo redigidas em português, e grande influência da parte 4 da fase II, na qual o candidato deve descrever uma imagem e relatar uma história com base nela, corroborando a alegação de que o efeito retroativo se intensifica com a proximidade do teste. Por outro lado, foi verificada a ausência ou pouca frequência de atividades com foco no gerenciamento do diálogo piloto-controlador e no uso de estratégias de negociação de sentido, bastante importantes para a comunicação aeronáutica, não só na 4ª série, mas também nas 2ª e 3ª séries. Embora essa tarefa esteja contemplada nos livros didáticos e nas provas orais das 2ª e 3ª séries e seja bastante representativa do ambiente de trabalho, os professores e instrutor observados optam por deixá-la em segundo plano, provavelmente, influenciados pelo fato de o EPLIS não a contemplar, provocando um efeito negativo nesse ponto.

6.3 Diagrama do efeito retroativo do EPLIS no cenário escolar

Com base no modelo de efeito retroativo proposto de Green (2007a), proponho o diagrama da Figura 11 que sintetiza o mecanismo de atuação do impacto do EPLIS nas aulas de inglês do curso BCT. O processo de efeito retroativo tem início na concepção do exame, cujas características atuais apresentam potencial para influenciar o ensino tanto positivamente como, por exemplo, na ênfase dada às habilidades orais, quanto negativamente, devido a problemas relacionados ao estreitamento das funções comunicativas avaliadas. Passamos,

então, para o primeiro filtro, posicionado no centro do diagrama, que corresponde às características, conhecimentos e valores da coordenação, chefia, professores e instrutores envolvidos no processo e que irão promover a variação no efeito. No caso do EPLIS, encontramos arestas ligadas à capacitação sobre o exame, à natureza da linguagem avaliada e ao letramento em avaliação, que afetam o modo como esses participantes interpretam as demandas do teste.

Figura 11 – Diagrama do efeito retroativo do EPLIS no curso BCT



Essas interpretações, por sua vez, influenciam como esses profissionais lidam com o material didático, fazendo com que certos tópicos, habilidades, aspectos linguísticos e tarefas ganhem proeminência na elaboração do currículo escolar. O currículo escolar afeta, assim, a construção das avaliações de rendimento que são limitadas pela cultura de avaliar da instituição e pelo nível de letramento em avaliação de seus elaboradores. As avaliações de rendimento, por terem alta-relevância no contexto escolar, direcionam os eventos em sala de aula. Esses filtros estão representados na pirâmide invertida.

O exame chega, então, à sala de aula, como pode ser observado na parte inferior do diagrama. A intensidade de seu efeito torna-se variável em função de sua proximidade, o que faz com que quanto mais próximos à data de realização do teste, mais intensa seja sua influência nos alunos, professores e instrutores. A percepção do nível de proficiência também interfere na intensidade do fenômeno, sendo mais forte para alunos e turmas mais proficientes. Esses alunos buscam mais informações sobre o teste e seus professores e instrutores tendem a suprir essas necessidades. A relevância do exame torna-se maior para esse grupo que, em razão da confiança depositada em sua capacidade de sucesso, engajam-se em práticas preparatórias, aumentando, assim, a intensidade de seu efeito.

As setas destacadas em preto demonstram, portanto, os caminhos mais comuns. Todavia, exceções, representadas pelas setas em cinza, são passíveis de ocorrer, por exemplo, quando um aluno, mesmo ciente de sua baixa proficiência, decide se dedicar ao exame, buscando informações e se preparando de alguma forma. Nesse caso, a relevância do teste para esse aluno aumenta, bem como a intensidade de seu efeito. O caminho inverso também pode ocorrer, por exemplo, quando alunos muito proficientes, ao perceberem a avaliação como fácil demais, não veem a necessidade de se engajar em nenhum tipo de preparação.

6.3 As contribuições do estudo

Os resultados desta pesquisa trazem contribuições teóricas, metodológicas e pedagógicas, bem como implicações para a (re)elaboração e implementação de políticas linguísticas. Discuto, aqui, cada um dessas contribuições e as relaciono com estudos anteriores.

6.3.1 Teóricas

Esta pesquisa, primeiramente, corrobora estudos anteriores que tiveram como foco principal de investigação o professor e cujos resultados demonstraram que: i) o efeito retroativo

varia de professor para professor e, portanto, não é determinista (ALDERSON; WALL, 1993; ALDERSON; HAMP-LYONS, 1996; WATANABE, 1996); ii) a intensidade do efeito também sofre influência do quanto os participantes sabem sobre o exame, dominam o conteúdo avaliado e possuem acesso a recursos para atender às suas demandas (HUGHES, 1993; GREEN, 2007a); iii) quanto mais próximos à sua realização maiores as chances de o teste mobilizar professores e alunos a se engajarem em práticas preparatórias (SMITH, 1991; WATANABE, 1996); iv) os valores, experiência e crenças dos participantes moldam sua interpretação sobre o exame (GIMENEZ, 1999; BURROWS, 2004); e v) exames, por si só, mesmo sendo bem elaborados, não são capazes de conduzir as inovações pretendidas, sendo necessário promover, conjuntamente, ações de desenvolvimento profissional com foco na reflexão crítica e informada (NOBLE; SMITH, 1994; CHENG, 1999; SCARAMUCCI, 1999; QI, 2004; WATANABE, 2004a).

Em relação aos alunos, autores como Alderson e Banerjee (2001), Scaramucci (2004), Stoneman (2006), Cheng et al (2011) apontaram a carência de estudos com foco nesses participantes, mesmo sendo eles um dos principais afetados pelos usos que se faz dos resultados dos testes. Esta pesquisa ajuda, portanto, a sanar essa deficiência e mostra resultados bastante interessantes sobre como o teste os afeta. Uma dessas contribuições refere-se à relação entre à percepção de relevância e o impacto do exame. Encontramos evidências de que a percepção dos alunos sobre a relevância de um teste interfere mais na intensidade do efeito retroativo do que o seu *status*, confirmando as conclusões de Stoneman (2006). Assim, embora as notas obtidas pelos alunos no EPLIS fossem creditadas pelas autoridades competentes no setor de controle de tráfego aéreo, a ausência de consequências imediatas enfraqueceu sua relevância no cenário escolar.

Além disso, os resultados iniciais alcançados por Cheng et al (2011) de que o nível de proficiência declarado pelo aluno afeta a intensidade do efeito retroativo, sendo essa maior para os alunos ditos mais proficientes, foram corroborados por esta pesquisa. Características pessoais dos alunos também se provaram fatores a serem considerados na determinação do efeito retroativo, assim como foi constatado por Bartholomeu (2002), Correia (2003), Lanzoni (2004) e Xie e Andrews (2012). Dessa forma, alunos que se autoavaliaram como mais proficientes foram descritos pelos professores e instrutor como mais confiantes em relação à obtenção de sucesso no exame e, consequentemente, buscavam mais conhecimentos sobre o teste e maneiras de melhor se preparar, confirmando que o efeito retroativo varia também de aluno para aluno (ALDERSON; WALL, 1993).

Outra grande contribuição de cunho teórico envolveu a relação existente entre a abordagem de ensinar e a proficiência dos alunos. Verificamos que, em turmas mais proficientes, os efeitos do exame eram mais intensos na prática pedagógica dos professores e instrutor observados, sendo atribuídos maior tempo de fala ao aluno e mais oportunidades de interação aluno-aluno. Já nas turmas menos proficientes, os professores e instrutor se mostraram mais controladores, dominando o tempo de fala e propondo menos atividades que permitissem a interação entre alunos. Embora Cheng et al (2011) tenham mostrado evidências que sugerem uma relação entre proficiência e efeito retroativo, seus dados não incluíram observações em sala de aula que pudessem investigar como isso ocorre na prática, o que foi feito neste estudo.

6.3.1 Metodológicas

Em termos metodológicos, esta pesquisa se distingue, portanto, da maioria dos estudos anteriores, pelo fato de ter adotado um desenho de métodos mistos, o que permitiu que os avanços teóricos, considerados importantes, se tornassem possíveis. Conforme destacado por Xie e Andrews (2012), as pesquisas exclusivamente qualitativas, embora sejam capazes de identificar inúmeros fatores atuantes no processo de efeito retroativo, mostram-se impedidas de tecer afirmações mais consistentes sobre as relações existentes entre eles. Por outro lado, as pesquisas fundamentadas em dados quantitativos somente, obtidos por meio de questionários, recaem na falta de dados de sala de aula que comprovem o que foi relatado pelos participantes.

Assim, os resultados de Cheng et al (2011) e Xie e Andrews (2012) se basearam em análises exclusivamente quantitativas. Bartholomeu (2002), Correia (2003) e Lanzoni (2004), apesar de terem coletado dados quantitativos e qualitativos, empregaram somente a estatística descritiva para a análise das informações numéricas, o que enfraquece as inferências feitas. No caso desta pesquisa, além da estatística descritiva, foi utilizada a estatística inferencial que possibilitou assegurar que as tendências observadas no grupo de alunos eram genuínas e não fruto do acaso, trazendo, dessa forma, mais embasamento e rigor para os seus resultados. Stoneman (2006), por sua vez, não conduziu observações em sala de aula que corroborassem o que foi relatado pelos participantes nos questionários e entrevistas, diferentemente do que foi feito neste estudo.

Ademais, o fato de esta pesquisa contemplar mais de um grupo de participantes (professores e alunos) e empregar diversos instrumentos de coleta de dados, propiciando não apenas a triangulação entre métodos (quantitativo e qualitativo), mas também entre diversas

fontes (grupo focal, entrevistas individuais e observações de aulas), possibilitou uma visão bem mais completa e abrangente do efeito retroativo, além de mais confiabilidade aos resultados, como pontuaram Quevedo-Camargo e Scaramucci (2016) sobre os benefícios de estudos que ampliam as suas opções metodológicas.

6.3.3 Pedagógicas

Este estudo também contribuiu para o desenvolvimento profissional dos professores e do instrutor observados. É incontestável a influência do professor e sua formação no processo de efeito retroativo, conforme atestam este e demais estudos como os de Bartholomeu (2002) e Correia (2003), respaldando a afirmação de Scaramucci (2006) de que o professor interage com as características do exame, as interpreta com base em suas crenças, conhecimentos e valores e define o que será ensinado e como será ensinado. As entrevistas, nas quais discutimos as ações em sala de aula, provaram ser um bom instrumento na estimulação da reflexão desses participantes. Ao longo da análise, foram transcritos trechos em que pudemos verificar o processo reflexivo dos participantes, em que eles justificavam suas ações, refletiam sobre elas e propunham mudanças futuras. Além disso, por diversas vezes, eles externaram os benefícios advindos dessas discussões e demonstraram o desejo de que a coordenação pedagógica escolar pudesse atuar da mesma forma. Isso demonstra que, por meio da prática reflexiva teoricamente informada, o professor encontra os subsídios necessários para se conscientizar de suas decisões, confrontar suas crenças e propor reconstruções. (LIBERALI, 2004; MAGALHÃES, 2004; OLIVEIRA, 2006; SOUZA, 2008; OLIVEIRA; SOUZA, 2010).

6.3.4 Elaboração e implementação de políticas linguísticas

Na esfera da elaboração e implementação de políticas linguísticas, os resultados deste estudo demonstraram que a política linguística da OACI, adotada pelo DECEA, apresenta dificuldades em promover os efeitos pretendidos somente por meio do EPLIS. Como parte integrante da equipe de elaboração do exame, me comprometi com o DECEA/ICEA a apontar áreas que necessitam de melhorias para o pleno cumprimento dessa política. Portanto, listo aqui intervenções sugeridas. Primeiramente, embora o EPLIS tenha promovido diversas inovações no ensino e aprendizagem de inglês na EEAR, efeitos indesejados foram encontrados devido, principalmente, ao estreitamento nas funções comunicativas avaliadas pelo exame. Assim, faz-

se urgente uma revisão do exame a fim de que a área de sobreposição entre o construto "inglês aeronáutico" e as características do teste, mostrada na Figura 11, seja ampliada.

Em segundo lugar, devem ser promovidas aos coordenadores e professores mais oportunidades de capacitação especializada sobre a natureza da linguagem avaliada, bem como sobre o exame, para que os currículos das disciplinas, elaborados por esses profissionais, sejam embasados nesses conhecimentos e não sejam apenas espelho das unidades do livro didático adotado. Além disso, posto que as avaliações de rendimento escritas carregam maior relevância no cenário escolar, é necessário que elas sejam reconstruídas para melhor representar as necessidades linguísticas do futuro profissional. Precisamos, então, oferecer aos seus elaboradores, professores e instrutores que desempenham duplo papel como facilitadores e avaliadores do processo de ensino e aprendizagem (SCARAMUCCI, 2006; INBAR-LOURIE, 2008b; SCARINO, 2013), as ferramentas necessárias na área de letramento em avaliação. Como vimos no estudo de Fulcher (2012), esse tipo de capacitação deve integrar princípios e conceitos da avaliação em contexto de línguas a conhecimentos procedurais aplicados ao desenvolvimento, manutenção e aplicação de testes. Inbar- Lourie (2008a) também destaca a importância de o professor compreender as informações sobre o desempenho de seus alunos nos testes e utilizá-las em benefício do processo de ensino e aprendizagem, bem como na tomada de decisões mais informadas e apropriadas. Dessa forma, como argumentam Quevedo-Camargo e Scaramucci (2018, p.239), "a proposta de tornar o professor de línguas letrado em avaliação tem como meta principal fazer dele um consumidor e produtor crítico de avaliações".

Em paralelo, as autoridades escolares também precisam ser convencidas de que as avaliações de inglês no curso BCT devem ser diretas, ou seja, necessitam promover o uso da língua em tarefas que se assemelham àquelas encontradas no ambiente profissional, em vez de indiretas, mostrando que existem maneiras seguras de implementá-las mantendo bons níveis de confiabilidade. Assim, promoveríamos a "cultura da avaliação" em oposição à "cultura da testagem", que tem imperado na EEAR por longos anos e que vê a avaliação como mera verificação da aprendizagem – avaliação *da* aprendizagem (SCARAMUCCI, 2006; INBAR-LOURIE, 2008b).

Outro ponto importante seria considerar o desempenho dos alunos, nessas novas avaliações, na escolha de vagas, ou seja, os alunos que apresentarem melhor desempenho seriam direcionados para vagas em órgãos de controle de tráfego internacional, enquanto aqueles com proficiência menor disputariam as vagas nos órgãos que atuam somente na aviação doméstica. Também é necessário discutir e estabelecer uma proficiência mínima nas habilidades de compreensão e produção orais para os ingressantes no Curso de Controle de

Tráfego Aéreo da EEAR. Como foi exposto no Capítulo III e retomado no Capítulo V, o exame de entrada na EEAR avalia somente a compreensão de textos, o que faz com que sejam aprovados candidatos com pouca ou nenhuma proficiência nessas duas habilidades, que são consideradas primordiais na condução do serviço de controle de tráfego aéreo.

6.4 As dificuldades e limitações do estudo

Não obstante as vantagens que a combinação de métodos trouxe para esta pesquisa como, por exemplo, o aprofundamento das relações existentes entre os diversos fatores que interagem com o efeito retroativo, essa escolha implicou dificuldades, dado que a tradição de pesquisas em Linguística Aplicada no Brasil constitui-se de estudos exclusivamente qualitativos, com poucas exceções à regra como, por exemplo, o estudo misto conduzido por minha orientadora (SCARAMUCCI, 1995). Conhecendo melhor os desafios que iríamos enfrentar, ela me aconselhou a procurar por profissionais de estatística para receber uma ajuda mais especializada. Por residir na cidade de São José dos Campos e trabalhar no Comando da Aeronáutica, fui ao Instituto Tecnológico da Aeronáutica (ITA) e comecei a frequentar aulas como ouvinte na disciplina de Estatística e Probabilidade, com o objetivo de construir conhecimentos introdutórios nessa área. Em paralelo, contratei uma empresa júnior de consultoria estatística, formada por alunos de graduação que desenvolvem projetos para a comunidade em geral, supervisionados por docentes da instituição onde estudam, para me auxiliar na construção dos questionários, na aplicação dos testes estatísticos e na análise dos dados. Expus um esboço do estudo e das ferramentas que pretendia empregar à aluna designada para conduzir o meu projeto e ela me confessou que ainda não havia estudado sobre o assunto e que era tão desconhecido para ela quanto para mim.

Para acompanhar os testes que estavam sendo feitos nos dados, comecei a estudar mais a fundo a análise fatorial e os testes de hipóteses, lendo diversos artigos das áreas de avaliação de línguas e educação que haviam empregado essas técnicas estatísticas, além de livros de estatística geral. Essa etapa foi fundamental, pois me possibilitou conhecer mais sobre as vantagens e limitações de cada método e saber interpretar, com segurança, os resultados obtidos. Foi, então, que constatei algumas discrepâncias no relatório apresentado pela consultoria e solicitei uma reunião com a responsável, juntamente com a professora supervisora. Apresentei as inconsistências encontradas e um novo relatório foi gerado. Porém ainda havia lacunas que me deixavam insatisfeita e achei mais prudente refazer as análises. Baixei o *software* SPSS (*Statistical Package for Social Sciences*) no meu computador pessoal

utilizando uma licença de 15 dias grátis e com a ajuda de um canal no *youtube* chamado *Quantitative Specialists*, que continha diversos vídeos explicativos sobre como conduzir análises estatísticas no SPSS, apliquei todos os testes feitos pela consultoria e incluí novos.

Apesar de todo o trabalho, acredito que esse esforço foi recompensado, não apenas pelo fato de eu ter alcançado o refinamento dos dados que tanto buscava, mas também pelo aumento na confiabilidade dos resultados que esse processo propiciou, sem mencionar as habilidades adquiridas ainda que restritas. Ressalto que, mesmo com a contratação de uma consultoria, a construção de conhecimentos estatísticos para um pesquisador em avaliação de línguas é de suma importância, dado que a exploração dos resultados depende da fusão de ambas as áreas.

Como se pode observar, desenvolver um estudo de natureza mista foi uma tarefa bastante árdua diante do cenário pouco convidativo em que eu me encontrava. Fica, portanto, o meu pedido aos programas de pós-graduação em Linguística Aplicada para que busquem alternativas para fomentar o desenvolvimento de pesquisas mistas, oferecendo, por exemplo, minicursos ou disciplinas sobre o assunto.

A fase qualitativa da pesquisa também apresentou dificuldades em torno das observações de aulas. O planejamento inicial contemplava acompanhar o trabalho em sala de aula de um pequeno número de professores, três a quatro, nas 2ª, 3ª e 4ª séries do curso. No entanto, a distribuição de aulas feita pela EEAR não me permitiu isso. A professora Pilar, o professor Eduardo e o instrutor João, voluntários a participar dessa etapa da pesquisa, atuaram nas 2ª e 3ª séries, mas não ministraram aulas para a 4ª série em 2017. A professora Ana, por sua vez, ministrou aulas para a 4ª série naquele ano, porém não foi escalada para as aulas das 2ª e 3ª séries. Embora tenhamos obtido evidências de que a intensidade do efeito do exame aumenta com a proximidade do teste, observar as ações de Pilar, Eduardo e João na 4ª série, bem como as de Ana nas 2ª e 3ª séries, teria proporcionado dados ainda mais ricos.

Ademais, mesmo tendo sido a 4ª série considerada a mais voltada para o exame, somente Ana se voluntariou para as observações. A dificuldade em angariar professores dispostos a abrirem as portas de sua sala de aula a uma pessoa muitas vezes desconhecida não é uma característica do meu estudo, pelo contrário, é enfrentada por muitos pesquisadores. A meu ver, isso se dá não apenas por uma questão de proteção da face, mas também por incerteza do professor de que a pesquisa resultará em contribuição direta para a sua prática pedagógica. Por isso, quis assegurar que as entrevistas conduzidas não fossem apenas instrumento de confiabilidade e validade dos dados, mas uma forma de contribuir para o crescimento profissional dos participantes.

O sistema de rodízio entre professores e instrutores, em que a mesma unidade é ministrada em diferentes turmas, embora tenha contribuído para que eu acompanhasse a atuação dos participantes em grupos de diferentes níveis de proficiência, restringiu o estudo no sentido de que não foi possível observá-los tralhando um repertório mais amplo de tópicos, habilidades e tarefas. Esse problema foi maior em relação ao professor Eduardo, cujas aulas assistidas não contemplaram o desenvolvimento da compreensão oral dos alunos em conversas radiotelefônicas. Somado a isso, o fato de eu ter concentrado as observações em duas semanas para as 2ª e 3ª séries e duas semanas para a 4ª série, a fim de otimizar o meu período de férias do trabalho para a geração de dados, também impôs limitações, fazendo com que não fosse possível assistir as aulas do professor Eduardo e do instrutor João em turmas de níveis bastante distintos de proficiência, como ocorreu com as professoras Pilar e Ana.

6.5 Sugestões para estudos futuros

Esta é a primeira pesquisa que se propôs a investigar sistematicamente o efeito retroativo do EPLIS e, por isso, optei por fazer um estudo panorâmico de sua influência nas percepções, atitudes e ações de professores, instrutores e alunos pesquisados. Embora eu tenha endereçado diversos aspectos do ensino e aprendizagem afetados pelo exame, a decisão de fornecer uma visão mais ampla impossibilitou aprofundá-los, abrindo, assim, a possibilidade para que vários estudos sobre efeito retroativo possam ser conduzidos futuramente.

Uma investigação a respeito do impacto do EPLIS nas avaliações de rendimento na EEAR, por exemplo, seria particularmente interessante. Como este estudo mostrou, as provas escritas receberam um alto nível de relevância, tornando-se direcionadoras dos eventos em sala de aula. Além disso, as questões de compreensão oral das avaliações da 4ª série pareceram ser as mais influenciadas pelo exame e seu formato e, portanto, analisar a natureza desses itens, a visão de linguagem que os embasam e como eles se assemelham aos itens do exame traria grandes contribuições, principalmente, para o aperfeiçoamento das avaliações escolares. Constatamos, também, influência do EPLIS na concepção e a elaboração das provas orais e um estudo sobre o efeito retroativo nessas avaliações e nas percepções de seus elaboradores expandiria nosso entendimento sobre o impacto de exames externos no desenvolvimento de exames de sala de aula e como esses potencializam os efeitos pretendidos.

Outro foco interessante de pesquisa seria o efeito retroativo do EPLIS no material didático da 1ª série. Durante a geração de dados, o currículo dessa série do curso estava passando por uma reestruturação para melhor se adequar às necessidades e interesses dos

alunos, o que compreendia, além da definição de novos objetivos de aprendizagem, a elaboração de uma apostila didática própria e de conjuntos de provas. Analisar esses materiais traria contribuições valiosas sobre o efeito retroativo na elaboração de livros didáticos e nas percepções de seus elaboradores.

Esta pesquisa ofereceu evidências de efeito retroativo do EPLIS e, dado que me coloquei como representante do DECEA e ICEA, mostra a preocupação dessas instituições com as consequências advindas do uso desse instrumento. No entanto, considerando que a validação de um exame é um processo contínuo e requer a coleta e análise de diversos tipos de evidência (SCARAMUCCI, 2011b), são bem vindos e incentivados estudos que se proponham a investigar as interpretações e usos que se fazem dos resultados do EPLIS como, por exemplo, estudos que investiguem o construto de uso da língua no contexto aeronáutico e sua operacionalização no teste, o que traria grandes contribuições para a continuidade no aperfeiçoamento do instrumento.

REFERÊNCIAS

ALDERSON, J. C. The politics of aviation English. **Language Assessment Quarterly**, v. 8, n. 4, p. 386-403, 2011.

_____. Forward. In: CHENG, L., WATANABE, Y.; CURTIS, A. (Ed.). **Washback in language testing**: Research, contexts and methods. New Jersey: Lawrence Erlbaum Associates, 2004. p.ix-xii.

_____. Testing and its administration in ESP. CHAMBERLAIN, D.; BAUMGARDNER, R. J. (Ed.). **ESP in the classroom**: Practice and evaluation. ELT Documents 129, 1988. p. 87-97.

ALDERSON, J. C.; WALL, D. Does washback exist? **Applied Linguistics**, v. 14, n. 2, p. 115-129, 1993.

ALDERSON, J. C.; CLAPHAM, C.; WALL, D. **Language test construction and evaluation**. Cambridge: Cambridge University Press, 1995.

ALDERSON, J. C.; HAMP-LYONS, L. TOEFL preparation courses: A study of washback. **Language Testing**, v. 13, n. 3, p. 281-297, 1996.

ALDERSON, J. C.; BANERJEE, J. Language testing and assessment − Part I. **Language Teaching**, v. 34, n. 4, p. 213-236, 2001.

AL-HOORIE, A. H.; VITTA, J. P. The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. **Language Teaching Research**, 2018. Disponível em: <http://journals.sagepub.com/doi/abs/10.1177/136216881876719>. Acesso em: 30 nov. 2016.

ALLEN, P.; FROHLICH, M.; SPADA, N. The communicative orientation of language teaching: An observation scheme. In: HANDCOMBE, J.; OREN, R.; TAYLOR, B. (Ed.) **On Tesol '83**: The question of control, Washington: TESOL, 1984. p. 231-252.

ALLWRIGHT, D.; BAILEY, K. M. **Focus on the language classroom**: An introduction to classroom research for language teachers. Cambridge: Cambridge, 1991.

ALMEIDA FILHO, J. C. P. **Dimensões comunicativas no ensino de línguas**. 3. ed. Campinas, SP: Pontes, 2002.

ANDRADE, N. R. Evidências de validade em exames de alta relevância: o caso do EPLIS. In: SCARAMUCCI, M. V. R.; TOSQUI-LUCKS, P.; DAMIÃO, S. M. (Org.). **Pesquisas sobre inglês aeronáutico no Brasil**. Campinas: Pontes, 2018, p. 271-293.

ANDREWS, S. Washback and curriculum innovation. In: CHENG, L.; WATANABE, Y.; CURTIS, A. (Ed.). **Washback in language testing**: Research, contexts and methods. New Jersey: Lawrence Erlbaum Associates, 2004. p. 137-150.

ANDREWS, S.; FULLILOVE, J.; WONG, Y. Targeting washback – a case study. **System**, v. 30, n. 2, p. 207-223, 2002.

ARAGÃO, B. F. O uso de critérios autóctones no contexto aeronáutico: Contribuições para uma nova escala de proficiência para controladores de tráfego aéreo. In: SCARAMUCCI, M. V. R.; TOSQUI-LUCKS, P.; DAMIÃO, S. M. (Org.). **Pesquisas sobre inglês aeronáutico no Brasil**. Campinas: Pontes, 2018, p. 243-269.

AZEVEDO, N. C. **Exame de Proficiência de Língua Inglesa (EPLI) para controladores de tráfego aéreo e operadores de estação aeronáutica brasileiros**: impactos potenciais. 2009. 38f. Trabalho de Conclusão de Curso. (Graduação em Licenciatura em Letras) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2009.

BACHMAN, L. F. **Fundamental considerations in language testing**. Oxford: Oxford University Press, 1990.

BACHMAN, L. F.; PALMER, A. S. **Language testing in practice**: designing and developing useful language tests. Oxford: Oxford University Press, 1996.

_____. **Language assessment in practice**. Oxford: Oxford University Press, 2010.

BAILEY, K. Working for washback: A review of the washback concept in language testing. **Language Testing**, v. 13, n. 3, p. 257-279, 1996.

_____. **Washback in language testing**: TOEFL Monograph Series MS – 15. Princeton: Education Testing Service, 1999.

BANDURA, A. Self-efficacy mechanism in human agency. **American Psychologist**, v. 37, n. 2, 1982, p. 122-147.

BARDIN, L. **Análise de conteúdo**. Tradução de Luís Antero Reto e Augusto Pinheiro. Lisboa: Edições 70, 1977.

BARBOUR, R.; SCHOSTAK, J. Interviewing and focus groups. In: SOMEKH, B.; LEWIN, C. (Ed.). **Research Methods in the Social Sciences**. London, Thousand Oaks, New Delhi: Sage Publications, 2005. p. 41-48.

BARTHOLOMEU, M. A. N. **Prova de língua estrangeira (Inglês) dos vestibulares e sua influência nas percepções, atitudes e motivações de alunos do terceiro ano do ensino médio**. 2002. 119f. Dissertação. (Mestrado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2002.

BASTURKMEN, H. ESP teacher education needs. **Language Teaching**, 2017. Disponível em: <http://www.cambridge.org/core/journals/language-teaching/article/esp-teacher-education-needs/94355A07092D3EFDAD6EC5A5D52CCA85>. Acesso em: 30 nov. 2016.

BLOCK, D. Problematizing interview data: Voices in the mind's machine? **Tesol Quarterly**, v. 34, n. 4, p. 757-763, 2000.

BONFIM, L. Grupos focais: conceitos, procedimentos e reflexões baseadas em experiências com o uso da técnica em pesquisas de saúde. **Phisis**, v. 19, n. 3, p. 777-796, 2009. Disponível em: <http://www.cambridge.org/core/journals/language-teaching/article/esp-teacher-education-needs/94355A07092D3EFDAD6EC5A5D52CCA85>. Acesso em: 30 nov. 2016.

BONIFÁCIO, E. M. R. **Critérios de avaliação de livro didático para o ensino −aprendizagem de inglês para controladores de tráfego aéreo brasileiros:** uma proposta de checklist. 2015. 170f. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de Taubaté, São Paulo, 2015.

BORG, S. Teacher cognition in language teaching: A review of research on what teachers think, know, believe, and do. **Language Teaching**, v. 36, p. 81-109, 2003.

BRASIL. Departamento de Controle do Espaço Aéreo. Publicações. **Fraseologia de tráfego aéreo** (MCA 100-16). Rio de Janeiro: DECEA, 2013. Disponível em: <http://publicacoes.decea.gov.br>. Acesso em: 30 nov. 2016.

BRASIL. Departamento de Controle do Espaço Aéreo. Publicações. **Plano de implementação dos requisitos de proficiência em inglês** (PCA 37-9), para o período 2014/2016. Rio de Janeiro: DECEA, 2014. Disponível em: < http://eplis.icea.gov.br/> Acesso em: 05 mar. 2014.

BRASIL. Departamento de Controle do Espaço Aéreo. Publicações. **Plano de implementação dos requisitos de proficiência em inglês** (PCA 37-9), para o período 2011/2013. Rio de Janeiro: DECEA, 2011.

BRASIL. Departamento de Controle do Espaço Aéreo. Publicações. **Plano de implementação dos requisitos de proficiência em inglês** (PCA 37-9), para o período 2008/2010. Rio de Janeiro: DECEA, 2008.

BRASIL. Departamento de Controle do Espaço Aéreo. Publicações. **Programa de elevação de nível na língua inglesa**, Rio de Janeiro: DECEA, 2007.

BURROWS, C. Searching for washback: The impact of the assessment in the Certificate in Spoken and Written English. In: BRINDLEY, G.; BURROWS, C. (Org). **Studies in immigrant English language assessment**. Sydney: National Centre for English Language Teaching and Research Macquarie University, 2001. p. 95-187.

_____. Washback in classroom-based assessment: A study of the washback effect in Australian adult migrant English program. In: CHENG, L.; WATANABE, Y.; CURTIS, A. (Ed.). **Washback in language testing**: Research, contexts and methods. New Jersey: Lawrence Erlbaum Associates, 2004. p. 113-128.

CANALE, M.; SWAIN, M. Theoretical bases of communicative approaches to second language teaching and testing. **Applied Linguistics**, v. 1, p. 1-47, 1980.

CHALHOUB-DEVILLE, M. The intersection of test impact, validation and educational reform policy. **Annual Review of Applied Linguistics**, v. 29, p. 118-131, 2009.

_____. Validity theory: Reform policies, accountability testing and consequences. **Language Testing**, v. 33, n. 4, p. 453-472, 2016.

CHAPELLE, C. A. Validity in language assessment. **Annual Review of Applied Linguistics**, v. 19, p. 254-271, 1999.

CHENG, L. Impact of a public English examination change on students' perceptions and attitudes toward their English learning. **Studies in Educational Evaluation**, v. 24, p. 279-301, 1998.

_____. Changing assessment: washback on teacher's perspectives and actions. **Teaching and Teacher Education**, v. 15, n. 3, p. 253-271, 1999.

_____. The washback effect of a public examination change on teacher's perceptions towards their classroom teaching. In: CHENG, L., WATANABE, Y.; CURTIS, A. (Ed.). **Washback in language testing**: Research, contexts and methods. New Jersey: Lawrence Erlbaum Associates, 2004. p. 147-170.

_____. **Changing language teaching through language testing**: a washback study. Cambridge: Cambridge, 2005.

_____. Washback, impact and consequences. In: SHOHAMY, E.; HORNBERGER, N. (Ed.). **Encyclopedia of language and education**. 2. ed. Language Testing and Assessment. New York: Springer, 2008. p. 349-364. v. 7

CHENG, L.; ANDREWS, S.; YU, Y. Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. **Language Testing**, v. 28, n. 2, p. 221-249, 2011.

CHINI, M. R. R. C. **Ensino-aprendizagem de inglês para o controlador de tráfego aéreo brasileiro**: em busca de novos rumos. 2014. 195 f. Dissertação (Mestrado em Linguística Aplicada) – Faculdade de Ciências Sociais e Letras, Universidade de Taubaté, Taubaté, 2014.

_____. Ensino de língua inglesa para controladores de tráfego aéreo: investigando a melhor rota para o aprendizado efetivo. In: SCARAMUCCI, M. V. R.; TOSQUI-LUCKS, P.; DAMIÃO, S. M (Org.). **Pesquisas sobre inglês aeronáutico no Brasil**. Campinas: Pontes, 2018, p. 133-154.

COE, R. **It's the effect size, stupid**: what ''effect size'' is and why it is important. Paper. Exeter, DVN: Annual Conference of the British Educational Research Association, 2002. Disponível em: <http://www.leeds.ac.uk/educol/ documents/00002182.htm>. Acesso em: 15 maio 2018.

COHEN, A. D. **Strategies in learning and using a second language**. New York: Addison Wesley Longman Inc, 1998.

_____. The coming of age or research on test-taking strategies. **Language Assessment Quarterly**, v. 3, n. 4, p. 307-331, 2006.

COHEN, L.; MANION, L.; MORRISON, K. **Research methods in education**. 6. ed. London/New York: Routledge, 2007.

COHEN, J. **Statistical power analysis for the behavioral sciences**. 2. ed. New York: Lawrence Erlbaum Associates, 1988.

CORREIA, R. M. D. **O efeito retroativo da prova de inglês do vestibular da Unicamp na preparação de alunos em um curso preparatório comunitário**. 2003. 154f. Dissertação (Mestrado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2003.

CRESWELL, J.; CLARK, V. **Designing and conducting Mixed Methods Research**. 2. ed. Thousand Oaks: Sage, 2011.

DAMÁSIO, B. F. Uso da análise fatorial exploratória em psicologia. **Avaliação psicológica**, v. 11, n. 12, p. 213-228, 2012. Disponível em: <http:// pepsic.bvsalud.org/pdf/avp/v11n2/v11n2a07.pdf >. Acesso em: 30 nov. 2016.

DUSHKU, S. Conducting individual and focus group interviews in research in Albania. **Tesol Quarterly**, v. 34, n. 4, p. 763-768, 2000.

EMERY, H. J. Developments in LSP testing 30 years on? The case of aviation English. **Language Assessment Quarterly**, v. 11, n. 2, p. 198-215, 2014.

ESPÍRITO-SANTO, H.; DANIEL, F. Calcular e apresentar tamanhos de efeito em trabalhos científicos: As limitações do $p < 0,05$ na análise de diferenças de médias de dois grupos. **Revista Portuguesa de Investigação Comportamental e Social**, v. 1, n. 1, p. 3-16, 2015.

ESTIVAL, D.; FARRIS, C. Aviation English as a lingua franca. In: ESTIVAL, D.; FARRIS, C.; MOLESWORTH, B. **Aviation English**: A lingua franca for pilots and air traffic controllers. Nova Iorque: Routledge, 2016. p. 1-18.

FREDERIKSEN, J. R.; COLLINS, A. A systems approach to educational testing. **Educational Researcher**, v. 18, n. 9, p. 27-32, 1989.

FROETSCHER, D. Washback on classroom testing: Assessment literacy as a mediating factor. In: LANGUAGE TESTING RESEARCH COLLOQUIUM, 39, 2017, Bogotá. **Resumos**. Bogotá: ILTA, 2017. p. 61-62.

FULCHER, G. Assessment literacy for the language classroom. **Language Assessment Quarterly**, v. 9, n. 2, p. 113-132, 2012.

GEBRIL, A.; EID, M. Test preparation beliefs and practices in a high-stakes context: A teacher's perspective. **Language Assessment Quarterly**, 2017.

GIMENEZ, T. Concepções de linguagem e ensino na preparação de alunos para o vestibular. **Trabalhos em Linguística Aplicada**, v. 34, p. 21-37, 1999.

GREEN, A. **IELTS washback in context**: Preparation for academic writing in higher education. Studies in Language Testing 25. Cambridge: Cambridge University Press and Cambridge ESOL, 2007a.

_____. Washback to learning outcomes: a comparative study of IELTS preparation and university pre-sessional language courses. **Assessment in Education**: Principles, Policy & Practice, v. 14, n. 1, p. 75-97, 2007b.

_____.Washback in language assessment. **International Journal of English Studies**, v. 13, n. 2, p. 39-51, 2013.

GREENE, J. C.; CARACELLI, V. J.; GRAHAM, W. F. Toward a conceptual framework for Mixed-Method evaluation designs. **Educational Evaluation and Policy Analysis**, v. 11, n. 3, p. 255-274, 1989.

HAIR, J. F. et al. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.

HALADYNA, T. M.; NOLEN, S. B.; HAAS, N. S. Raising standardized achievement test scores and the origins of test score pollution. **Educational Researcher**, v. 20, n. 5, p. 2-7, 1991.

HALADYNA, T.; DOWNING, S. Construct irrelevance variance in high stakes testing**.** **Educational Measurement**: Issues and Practice, v. 23, p. 17-27, 2004.

HAMP-LYONS, L. Ethical test preparation: The case of the TOEFL. **Tesol Quarterly**, v. 32, n. 2, p. 329-337, 1998.

HENNING, G. **A guide to language testing**. Cambridge: Cambridge University Press, 1987.

HUGHES, A. **Backwash and TOEFL 2000**. Unpublished manuscript. Reading, UK: University of Reading, 1993.

_____. **Testing for language teachers**. Cambridge: Cambridge University Press, 1989.

HUTCHINSON, T.; WATERS, A. **English for specific purposes**: a learning-centred approach. Cambridge: Cambridge, 1987.

HYMES, D. H. On communicative competence. In: PRIDE, J. B.; HOLMES, J. (Ed.) **Sociolinguistics**: Selected Readings. Harmondsworth: Penguin, p. 269-293.

INBAR-LOURIE, O. Constructing a language assessment knowledge base: A focus on language assessment courses. **Language Testing**, v. 25, n. 3, p. 385-402, 2008a.

_____. Language assessment culture. In: SHOHAMY, E.; HORNBERGER, N. H. (Ed.). **Encyclopedia of language and education**: Language testing and assessment. Springer Science+Business Media LLC, 2008b, p. 285-299. v. 7.

_____. Guest editorial to the special issue on language assessment literacy. **Language Testing**, v. 30, n. 3, p. 301-307, 2013.

JANG, E. E.; WAGNER, M.; PARK, G. Mixed Methods Research in language testing and assessment. **Annual Review of Applied Linguistics**, v. 34, p. 123-153, 2014.

JOHNSON, R. B.; ONWUEGBUZIE, A. J.; TURNER, L. A. Toward a definition of mixed methods research. **Journal of Mixed Methods Research**, v. 1, n. 2, p. 112-133, 2007.

JOHNSON, R. B.; ONWUEGBUZIE, A. J. Mixed Methods Research: a research paradigm whose time has come. **Educational Researcher**, v. 33, n. 7, p. 14-26, 2004.

KANE, M. Validating high-stakes testing programs. **Educational measurement**: Issues and practices, v. 21, n. 1, p. 31-41, 2002.

KOBAYASHI, E. **O efeito retroativo de um exame de proficiência em língua inglesa em um núcleo de línguas do programa inglês sem fronteiras**. 2016. 187f. Tese (Doutorado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2016.

LANKSHEAR, C.; KNOBEL, M. **A handbook for teacher research**. Berkshire: Open University Press, 2004.

LANZONI, H. DE P. **Exame de proficiência em leitura de textos acadêmicos em Inglês: um estudo sobre efeito retroativo**. 2004. 257f. Tese (Doutorado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2004.

LAVILLE, C.; DIONNE, J. **A construção do saber**. Tradução de Heloísa Monteiro e Franciso Settineri. Porto Alegre: Artmed; Belo Horizonte: Editora UFMG, 1999.

LIBERALI, F. C. As linguagens das reflexões. In: MAGALHÃES, M. C. **A formação do professor como um profissional crítico**. Campinas: Mercado de Letras, 2004, p. 87-118.

LONG, M. Focus on form: a design feature in language teaching methodology. In: BOT, K.; GINSBERG, R.; KRAMSCH, C. **Foreign Language Research in Cross-Cultural perspective**. Amsterdam: John Benjamins, 1991, p. 39-52.

MAGALHÃES, M. C. A linguagem na formação de professores como profissionais reflexivos e críticos. In: MAGALHÃES, M. C. **A formação do professor como um profissional crítico**. Campinas: Mercado de Letras, 2004. p. 59-86.

MARCUSCHI, L. A. **Análise da conversação**. 6. ed. São Paulo: Ática, 2008.

MCNAMARA, T. **Measuring second language performance**. London/New York: Addison Wesley London, 1996.

_____. Performance testing. In: CLAPHAM, C.; CORSON, D. (Ed.). **Encyclopedia of language and education**: Language testing and assessment. Netherlands: Kluwer Academic Publishers, 1997. p. 131-139. v. 7.

_____. Language assessment as social practice: Challenge for research. **Language Testing**, v. 18, n. 4, p. 333-349, 2001.

_____. The use of language tests in the service of policy: issues of validity. **Rev. franç. de linguistique appliquée**, v. XV, n. 1, p. 7-23, 2010.

MCNAMARA, T.; ROEVER, T. **Language Testing**: The social dimension. Malden, MA: Blackwell, 2006.

MESSICK, S. Validity. In: LINN, R. (Ed). **Educational measurement**. 3 ed. New York: Macmillan, 1989. p. 13-103.

_____. Validity and washback in language testing. **Language Testing**, v. 13, n. 3, p. 241-255, 1996.

_____. Issues of effectiveness and equity in the coaching controversy: Implications for educational and testing practice. **Educational Psychologist**, n. 17, p. 67-91, 1982.

NOBLE, A.; SMITH, M. **Old and new beliefs about measurement-driven reform**: The more things change, the more they stay the same. CSE Technical Report 373, Tempe, AZ: Arizona State University, CSE, 1994.

OLLAIK, L. G.; ZILLER, H. M. Concepções de validade em pesquisas qualitativas. **Educação e Pesquisa**, v. 38, n. 1, p. 229-241, 2012.

OLIVEIRA, A. L. A. M. **Hermes e as bonecas russas**: um estudo colaborativo para compreender a relação teoria-prática na formação docente. 2006. 247f. Tese de doutorado − Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

OLIVEIRA, A. L. A. M.; SOUZA, P. R. Formação de professores em Linguística Aplicada: poderá Sísifo ser feliz? In: BARROS, S. M.; ASSIS-PETERSON, A. A. **Formação crítica de professores de línguas**: desejos e possibilidades. São Carlos: Pedro & Editores, 2010, p. 71-84.

ORGANIZAÇÃO DE AVIAÇÃO CIVIL INTERNACIONAL. **Manual on the implementation of ICAO language proficiency requirements**: Doc 9835 AN/453.1. ed. Montreal: OACI, 2004.

ORGANIZAÇÃO DA AVIAÇÃO CIVIL INTERNACIONAL/OACI. **Manual on the implementation of ICAO language proficiency requirements**: Doc. 9835 AN/453. 2. ed. Montreal, 2010.

ORGANIZAÇÃO DA AVIAÇÃO CIVIL INTERNACIONAL/OACI. **Guidelines for Aviation English training programmes**: Circular 323-AN/185. Montreal: OACI, 2009.

PEARSON, I. Test as levers for change. In: CHAMBERLAIN, D.; BAUMGARDNER, R. J. (Ed.). **ESP in the classroom**: Practice and evaluation. ELT Documents 129, 1988. p. 98-108.

PLONSKY, L.; OSWALD, F. L. How big is “big”? Interpreting Effect Sizes in L2 Research. **Language Learning**, v. 64, n. 4, p. 878-912, 2014.

POPHAM, W. J. Appropriateness of teachers’ test preparation practices. **Educational Measurement**: Issues and Practice, v. 10, n. 4, p. 12-15, 1991.

PRADO, M. C. A; TOSQUI-LUCKS; P. Are the LPRs focusing on real life communication issues? In: INTERNATIONAL CIVIL AVIATION ENGLISH ASSOCIATION CONFERENCE, 2017, Daytona. **Workshop**. Daytona: ICAEA, 2017.

PRODOMOU, L. The backwash effect: from testing to teaching. **ELT Journal**, v. 49, n. 1, p. 13-25, 1995.

QI, L. Has a High-stakes test produced the intended changes? In: CHENG, L., WATANABE, Y.; CURTIS, A. (Ed.). **Washback in Language Testing: Research, Contexts and Methods.** New Jersey: Lawrence Erlbaum Associates, 2004. p. 171-190.

QUEVEDO-CAMARGO, G.; SCARAMUCCI, M.V.R. Methodology in washback studies. In: GITSAKI, C.; COOMBE, C. (Org.) **Current trends in language evaluation, assessment & testing**: Research perspectives. Cambridge: Cambridge Scholars Publishing, v. 1, 2016, p. 287-304.

_____. O conceito de letramento em avaliação de línguas: Origem de relevância para o contexto brasileiro. **Linguagem: Estudos e Pesquisa**, v. 22, n. 1, p. 225-245, 2018. Disponível em: <https://www.revistas.ufg.br/lep/article/view/54474>. Acesso em: 11 set. 2018.

RETORTA, M. S. **Efeito retroativo do vestibular da Universidade Federal do Paraná no ensino de língua inglesa em nível médio no Paraná**: uma investigação em escolas públicas, particulares e cursos pré-vestibulares. 2007. 483f. Tese (Doutorado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2007.

SÁ, P. P. T. F. **Novos caminhos para o conteúdo programático das disciplinas Inglês 2 e Inglês 3 do Curso de Formação de Controladores de Tráfego Aéreo militares**. 2010. 130f. Dissertação (Mestrado em Linguística Aplicada) – Universidade de Taubaté, Taubaté, 2010.

SCARAMUCCI, M. V. R. **O papel do léxico na compreensão em leitura em língua estrangeira**: foco no produto e no processo. 1995. 345f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

_____. Avaliação de Rendimento no Ensino-Aprendizagem de Português Língua Estrangeira. In: ALMEIDA FILHO, J. C. (Org). **Parâmetros atuais para o ensino de português língua estrangeira**. Campinas: Pontes, p. 75-88, 1997.

_____. Vestibular e ensino de língua estrangeira (Inglês) na escola pública. **Trabalhos de Linguística Aplicada**, v. 34, p. 7-20, 1999.

_____. Proficiência em LE: Considerações terminológicas e conceituais. **Trabalhos em Linguística Aplicada**, v. 36, p. 11-22, 2000.

_____. Efeito retroativo da avaliação no ensino/aprendizagem de línguas: o estado da arte. **Trabalhos de Linguística Aplicada**, v. 43, n. 2, p. 203-226, 2004.

_____. O professor avaliador: sobre a importância da avaliação na formação do professor de língua estrangeira. In: Rottava, L. (Org.). **Ensino-aprendizagem de línguas**: Língua Estrangeira. 1. ed. Ijuí: Editora da UNIJUI, 2006. p. 49-64. v. 1.

_____. Validade e consequências sociais das avaliações em contextos de ensino de línguas. **Linguarum Arena**, v. 2, p. 103-120, 2011a.

_____. O Exame de Proficiência em Língua Inglesa para controladores de voo do SISCEAB − Uma entrevista com Matilde Scaramucci. **Aviation in Focus**, Porto Alegre, v. 2, n. 1, p. 3-12, 2011b.

_____. Letramento em avaliação (em contexto de línguas): Contribuições para a linguística aplicada, educação e sociedade. In: JORDÃO, C. M. (Org). **Linguística aplicada no Brasil**: rumos e passagens. Campinas: Pontes Editora, 2016. p. 141-165.

SCARAMUCCI, M. V. R. et al. **Validação do exame de proficiência em Língua Inglesa para profissionais prestadores de serviços de tráfego aéreo do SISCEAB**. 2010. 243f. Relatório final de pesquisa. Campinas: Unicamp, 2010.

SCARAMUCCI, M. V. R.; TOSQUI-LUCKS, P.; DAMIÃO, S. M. (Org.). **Pesquisas sobre inglês aeronáutico no Brasil**. Campinas: Pontes, 2018.

SCARINO, A. Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. **Language Testing**, v. 30, n. 3, p. 309-327, 2013.

SHOHAMY, E. **The power of tests**: the impact of language tests on teaching and learning. NFLC Occasional Paper. Washington: National Foreign Language Center, 1993.

_____. **Language policy**: Hidden agendas and new approaches**.** New York: Routledge, 2006.

SILVA, A. L. B. C. **Uma análise de necessidades de uso da língua inglesa por oficiais aviadores do Esquadrão de Demonstração Aérea da Força Aérea Brasileira**. 2016. 278f. Dissertação (Mestrado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2016.

SMITH, M. Put to the test: The effects of external testing on teachers. **Educational Researcher**, v. 20, n. 5, p. 8-11, 1991.

SOUZA, P. R. O efeito retroativo do EPLIS nas percepções de professores. **Brazilian English Language Teaching Journal**, v. 8, n. 1, p. 24-39, 2017. Disponível em: < http://revistaseletronicas.pucrs.br/ojs/index.php/belt/issue/view/1150>. Acesso em: 25 jun. 2017.

_____. **As conceptualizações sobre o ensino de gramática de duas professoras em formação inicial, suas práticas de sala de aula e seus processos reflexivos**. 2008. 156f. Dissertação (Mestrado em Linguística Aplicada) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

SPOLSKY, B. Language testing at 25: Maturity and responsibility? **Language Testing**, v. 25, n. 3, p. 297-305, 2008.

SPRATT, M. Washback and the classroom: the implications for teaching and learning of studies of washback from exams. **Language Teaching Journal**, v. 9, n. 1, p. 5-29, 2005.

STENLUND, T.; EKLÖF, H.; LYREN, P. Group differences in test-taking behavior: an

example from a high-stakes testing program. **Assessment in Education**, v. 24, n.1, p. 4-20, 2017.

STONEMAN, B. **The impact of an exit English test on Hong Kong undergraduates**: a study investigation of the test' status on students' test preparation behaviours. 2006. 473p. Unpublished PhD Dissertation. Hong Kong Polytechnic University: Hong Kong, China, 2006.

TAYLOR, L. Developing assessment literacy. **Annual review of applied linguistics**, v. 29, p. 21-36, 2009.

_____. Communication the theory, practice and principles of language testing to test stakeholders: Some reflections. **Language Testing**, v. 30, n. 3, p. 403-412, 2013.

TERREL, S. R. Mixed Methods Research Methodologies. **The Qualitative Report**, v. 17, n. 1, p. 254-280, 2012. Disponível em: http://www.nova.edu/ssss/QR/QR17-1/terrell.pdf. Acesso em: 20 mar. 2016.

TOSQUI-LUCKS, P. T. et al. Ensino e avaliação de Língua Inglesa para controladores de tráfego aéreo como requisito de segurança em voo. **Revista Conexão SIPAER**, v. 7, n. 1, p. 44-54, 2016. Disponível em: http://conexaosipaer.cenipa.gov.br/index.php/sipaer/issue/view/19. Acesso em: 25jun. 2017.

TSAGARI, D. **The complexity of test washback**. Germany: Peter Lang, 2009.

VIEIRA-ABRAÃO, M. H. Algumas reflexões sobre a abordagem comunicativa: O pós método e a prática docente. **EntreLínguas**, v. 1, n. 1, p. 25-41, 2015.

XIE, Q. Does test preparation work: Implications for score validity. **Language Assessment Quarterly**, v. 10, n. 2, p. 196-218, 2013.

XIE, Q.; ANDREWS, S. Do test design and uses influence test preparation? Testing a model of washback with Structural Equation Modeling. **Language Testing**, v. 30, n. 1, p. 49-70, 2012.

WALL, D.; HORAK, T. Using baseline studies in the investigation of test impact. **Assessment in Education**: Principles, Policy & Practice, v. 14, n. 1, p. 99-116, 2007.

WATANABE, Y. Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. **Language Testing**, v. 13, n. 3, p. 318-333, 1996.

_____. Methodology in Washback studies. In: CHENG, L.; WATANABE, Y.; CURTIS, A. (Ed.). **Washback in Language Testing**: Research, Contexts and Methods. New Jersey: Lawrence Erlbaum Associates, 2004a. p. 19-36.

_____. Teacher factor mediating washback. In: CHENG, L., WATANABE, Y.; CURTIS, A. (Ed.). **Washback in Language Testing**: Research, Contexts and Methods. New Jersey: Lawrence Erlbaum Associates, 2004b, p. 129-146.

WINKE, P. Evaluating the validity of a high-stakes ESL Test: Why teachers' perceptions matter. **TESOL Quarterly**, v. 45, n. 4, p. 628-660, 2011.

APÊNDICES

APÊNDICE A – Questionário inicial do professor

Aim: to obtain preliminary data on teacher's perceptions and reactions towards the examination (EPLIS).

Part 1 – Personal Information

1) **Gender:** ( ) female ( ) male

2) **Age:** ( ) 20-30 ( ) 31-40 ( ) 41-50 ( ) above 50

3) **Academic / Professional Qualifications:**

( ) a degree in Teaching English
( ) a postgraduate course in Teaching English
( ) a Master's Degree Area: _____
( ) a Doctor's Degree Area: _____

4) **Number of years you have taught English:**

( ) 1-2 ( ) 3-4 ( ) 5-6 ( ) 7-10 ( ) above 10

5) **Number of years you have taught Aviation English:**

( ) below 1 year ( ) 1-2 ( ) 3-4 ( ) above 5

6) **Before teaching Aviation English, have you had any experience with ESP courses?**

( ) yes ( ) no
What if any?
_____

7) **Which groups do you currently teach?**

( ) CFS grade 1 ( ) CFS grade 2 ( ) CFS grade 3 ( ) CFS grade 4 ( ) CFS ME

8) **The size of your class in terms of student numbers:**

( ) below 20 ( ) 21-30 ( ) 31-40 ( ) above 40

9) **Are the students at your school placed according to their level of English?**

( ) yes ( ) no

10) **The medium of instruction you use in the classroom:**

( ) English only ( ) English with occasional Portuguese ( ) half English half Portuguese ( ) mainly Portuguese

11) **Textbooks you mainly use in Teaching:**

_____
_____
_____

Part 2 – Teacher's perceptions and reactions in relation to EPLIS

12) **How did you get to know about EPLIS?**

( ) pedagogical coordinator
( ) ICEA *website*
( ) workshops and seminars promoted by ICEA
( ) colleagues
( ) Other: _____

13) In a scale from 1 to 5, in which 1 means low and 5 means high, grade the situations:

	1	2	3	4	5
To what extent do you know EPLIS?					
To what extent do you think your students know EPLIS?					
To what extent has EPLIS influenced the school's syllabus?					
To what extent has EPLIS influenced the choice of textbooks?					
To what extent has EPLIS influenced the school's methodology?					
To what extent has EPLIS influenced your teaching?					
To what extent has EPLIS influenced the tasks in the classroom?					
To what extent has EPLIS influenced the school's assessment process?					
To what extent does EPLIS put pressure on you?					
To what extent has EPLIS influenced students' learning strategies?					

14) Do you think you have changed your teaching in some way because of EPLIS?

If yes, how? If now, could you give the main reasons?

_____
_____
_____

Part 3 – Teacher's beliefs on teaching and learning

15) Do you think students are better off having general English or Aviation English classes?

_____
_____
_____

16) Can you specify the percentage of time you spend on each area?

( ) listening ( ) speaking ( ) vocabulary ( ) grammar
( ) reading ( ) writing

17) In which area do students have the most difficulty?

( ) listening ( ) speaking ( ) vocabulary ( ) grammar
( ) reading ( ) writing

APÊNDICE B – Questionário inicial do aluno

Objetivo: coletar dados preliminares sobre as percepções e atitudes de alunos em relação ao EPLIS.

Parte 1 – Informações Gerais

1) **Sexo:** ( ) feminino ( ) masculino

2) **Por quantos anos aproximadamente estuda ou estudou inglês (fora da EEAR)?**

( ) 1-2 ( ) 3-4 ( ) 5-6 ( ) 7-10 ( ) acima de 10

3) **Onde você estuda ou estudou inglês?**

( ) escola pública ( ) escola particular ( ) escolas de idiomas ( ) aulas particulares ( ) outras: _____

4) **Antes de entrar na EEAR, você considera que o seu nível de conhecimento em inglês era:**

( ) iniciante ( ) básico ( ) pré-intermediário ( ) intermediário ( ) intermediário avançado ( ) avançado ( ) proficiente

5) **Você percebeu algum avanço no seu nível de conhecimento em inglês depois que você entrou na EEAR:**

( ) sim ( ) não Por quê? _____

Parte 2 – Percepções e Atitudes em relação ao EPLIS

6) **Como você ficou sabendo sobre o EPLIS?**

( ) não tinha conhecimento sobre o exame até pouco tempo atrás
( ) professores
( ) *site* do ICEA / manual do candidato
( ) palestras na EEAR
( ) colegas de profissão
( ) Outros: _____

7) **Em uma escala de 1 a 5, na qual 1 significa baixo e 5 significa alto, responda as questões abaixo.**

	1	2	3	4	5
Qual o seu grau de conhecimento sobre o EPLIS?					
Com que frequência os seus professores mencionam o EPLIS ?					
As aulas na 4a. série colaboraram para o seu desempenho hoje no EPLIS?					
As tarefas trabalhadas em sala de aula são semelhantes às tarefas no EPLIS?					
O EPLIS representa bem as tarefas de um controlador de tráfego aéreo?					
O EPLIS o motivou a estudar inglês?					
Você se sentiu pressionado pelo EPLIS?					

O EPLIS modificou sua maneira de estudar inglês?					
Você se preparou para o EPLIS fora da sala de aula?					

Parte 3 – Percepções de alunos sobre ensino e aprendizagem de língua

8) Enumere de **1** a **6** as atividades mais trabalhadas pelos seus professores em sala de aula, na qual **1** significa **mais trabalhada** e **6** significa **menos trabalhada**.

( ) exercícios de leitura
( ) exercícios de escrita
( ) exercícios de compreensão oral (áudios e videos)
( ) exercícios de prática oral
( ) exercícios de gramática
( ) exercícios de vocabulário

9) Quais atividades em sala de aula mais colaboraram para o seu desempenho hoje no EPLIS? Você poderá marcar mais de uma opção.

( ) exercícios de leitura
( ) exercícios de escrita
( ) exercícios de compreensão oral
( ) exercícios de prática oral
( ) exercícios de gramática
( ) exercícios de vocabulário

Obrigada por responder ao questionário. Os dados coletados serão de grande valia para a melhoria do nosso processo de capacitação e avaliação. Marque abaixo se você concorda com o uso desses dados para fins de pesquisa.
( ) Concordo.

APÊNDICE C – Questionário principal do professor

Parte 1 – INFORMAÇÕES GERAIS

1. Sexo: ( ) feminino ( ) masculino ( ) não gostaria de opinar

2. Idade: ( ) 20-30 ( ) 31-40 ( ) 41-50 ( ) above 50

3. **Assinale sua maior qualificação na área de ensino de língua inglesa:**

( ) Cursos TEFL, CELTA, CTP011, CTP009, ou similares
( ) Graduação em Letras – Inglês
( ) Pós-Graduação ou Especialização
( ) Mestrado
( ) Doutorado
( ) Não tenho

4. **Você possui formação como controlador de tráfego aéreo?**

( ) Sim.
( ) Não.

5. **Há quantos anos você atua como professor de inglês?**

( ) menos de 1 ano ( ) 1 a 2 ( ) 3 a 5 ( ) 6 a 9 ( ) 10 anos ou mais

6. **Há quantos anos você ensina inglês aeronáutico?**

( ) menos de 1 ano ( ) 1 a 2 ( ) 3 a 5 ( ) 6 a 9 ( ) 10 anos ou mais

7. De forma geral, como você avalia o seu nível de conhecimento sobre ensino para fins específicos?

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

8. De forma geral, como você avalia o seu nível de conhecimento sobre inglês aeronáutico?

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

9. Para quais grupos você leciona inglês atualmente?

( ) CFS 1ª série ( ) CFS 2ª série ( ) CFS 3ª série ( ) CFS 4ª série ( ) Demais especialidades

Part 2 – CONHECIMENTOS SOBRE O EPLIS

10. De forma geral, como você avalia o seu nível de conhecimento sobre o EPLIS?

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

Como você avalia o seu nível de conhecimento sobre os seguintes itens? Assinale de acordo com a legenda abaixo:

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

	1	2	3	4	5
11. O Doc 9835 AN/453 da OACI					
12. O objetivo do EPLIS					
13. O conteúdo avaliado no EPLIS					
14. As tarefas da fase 1 do EPLIS					
15. As tarefas da fase 2 do EPLIS					

16. A tabela de níveis de proficiência da OACI					
17. O desempenho necessário para obter o nível mínimo (nível 4)					

18. Com que frequência são promovidos workshops ou seminários sobre o EPLIS na sua instituição?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

19. Com que frequência você costuma visitar o *site* do EPLIS para obter informações sobre o exame?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

Parte 3 – IMPACTO DO EPLIS NO ENSINO

Considerando o impacto do EPLIS no ensino, analise a alternativa que mais expressa o seu nível de concordância ou discordância em relação às seguintes afirmações, na qual:

1= Discordo totalmente 2= Discordo 3= Nem concordo e nem discordo 4= Concordo 5= Concordo totalmente

	1	2	3	4	5
20. O EPLIS influencia a minha maneira de ensinar.					
21. O EPLIS influencia o conteúdo trabalhado em sala de aula.					
22. O EPLIS influencia a escolha do material didático pela escola.					
23. O EPLIS influencia os tipos de tarefas realizadas em sala de aula.					
24. O EPLIS influencia os projetos desenvolvidos nas disciplinas de língua inglesa.					
25. Os itens de prova, no processo de avaliação dos alunos na EEAR, são influenciados pelas tarefas do exame.					

26. Com que frequência você menciona o EPLIS em sala de aula?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

27. Com que frequência você discute aspectos específicos do exame em sala de aula?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

28. Com que frequência você trabalha questões do EPLIS em sala de aula?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

29. Com que frequência os seus alunos te procuram, fora da sala de aula, para tirar dúvidas sobre o EPLIS?

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

Com que frequência as seguintes atividades realizadas em sala de aula estão relacionadas à AVIAÇÃO E AO CONTROLE DE TRÁFEGO AÉREO? Assinale de acordo com a legenda abaixo:

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

	1	**2**	**3**	**4**	**5**
30. Compreensão de áudios e vídeos					
31. Compreensão de textos escritos					
32. Prática da produção oral (*speaking*)					

33. Prática de vocabulário					
34. Prática de itens gramaticais					

Parte 4 – PREPARAÇÃO DOS ALUNOS PARA O EPLIS

Com que frequência você costuma trabalhar as seguintes atividades com seus alunos em sala de aula? Assinale de acordo com a legenda abaixo:

1 = nunca 2 = raramente 3 = às vezes 4 = frequentemente 5 = sempre

	1	2	3	4	5
35. Responder, oralmente, perguntas sobre aviação e controle de tráfego aéreo					
36. Discutir com os colegas um assunto específico sobre aviação e controle de tráfego aéreo					
37. Fazer uma apresentação oral sobre um assunto relacionado à aviação e ao controle de tráfego aéreo					
38. Descrever uma figura relacionada à aviação e ao controle de tráfego aéreo					
39. Contar uma história com base em uma figura relacionada à aviação e ao controle de tráfego aéreo					
40. Recontar uma história com base em um áudio ou texto relacionado à aviação e ao controle de tráfego aéreo					
41. Explicar, oralmente, uma situação-problema de tráfego aéreo					
42. Dar uma sugestão, oralmente, para um problema de tráfego aéreo					
43. Simular, com um colega, um diálogo entre piloto e controlador					
44. Responder questões sobre um áudio ou vídeo envolvendo uma conversa entre piloto e controlador					
45. Responder questões sobre um texto cujo tema está relacionado à aviação e ao controle de tráfego aéreo					
46. Escrever um diálogo entre piloto e controlador					
47. Ler em voz alta um diálogo entre piloto e controlador					
48. Escrever uma história envolvendo uma situação complicadora de tráfego aéreo					
49. Explicar, oralmente, o significado de uma palavra ou expressão ligada à aviação e ao controle de tráfego aéreo					

50. Como você avalia a preparação recebida pelo aluno na EEAR para a realização do exame?

1 extremamente insuficiente
2 insuficiente
3 razoável
4 suficiente
5 extremamente suficiente

Parte 5 – VISÕES E PERCEPÇÕES SOBRE O EXAME

Analise a alternativa que mais expressa o seu nível de concordância ou discordância em relação às seguintes afirmações, na qual:

1= Discordo totalmente 2= Discordo 3= Nem concordo e nem discordo 4= Concordo 5= Concordo totalmente

	1	2	3	4	5
51. A realização do EPLIS durante o curso de formação na EEAR motiva os alunos a estudarem inglês.					
52. O EPLIS colabora para a elevação do nível de proficiência dos alunos.					
53. O EPLIS representa bem as atividades de um controlador de tráfego aéreo.					
54. O EPLIS avalia precisamente o nível de proficiência dos alunos.					
55. O sucesso no EPLIS é importante para o desempenho do aluno como controlador de tráfego aéreo.					
56. O resultado no EPLIS deve ser considerado para a nota final na disciplina de inglês da 4ª série.					
57. A realização do EPLIS durante o curso de formação na EEAR deve ser opcional.					
58. As aulas de inglês na EEAR ajudam os alunos a se preparam bem para o EPLIS.					
59. O professor sente-se confortável na preparação dos alunos para o EPLIS.					
60. O professor sente-se pressionado a ensinar o conteúdo contemplado pelo exame.					
61. O EPLIS aumenta consideravelmente a carga de trabalho do professor.					
62. O EPLIS aumenta consideravelmente o nível de *stress* / ansiedade dos alunos.					
63. O EPLIS aumenta consideravelmente a carga de estudos dos alunos.					
64. O EPLIS limita substancialmente os tipos de atividades realizadas em sala de aula.					
65. O EPLIS faz com que os alunos foquem em técnicas de realização de provas em detrimento do desenvolvimento de sua competência na língua.					

APÊNDICE D – Questionário principal do aluno

Parte 1 – INFORMAÇÕES GERAIS

1. Sexo: ( ) feminino ( ) masculino ( ) não gostaria de informar

2. Qual série do CFS-BCT você está cursando?

( ) 1ª série ( ) 2ª série ( ) 3ª série ( ) 4ª série

3. A qual turma de inglês você pertence?

1 Turma A 2 Turma B 3 Turma C 4 Turma D 5 Turma E 6 Turma F

4. Por quantos anos, aproximadamente, você estuda ou estudou inglês (não considerar as aulas na EEAR)?

( ) 0 a 1 ( ) 2 a 4 ( ) 5 a 6 ( )7 a 10 ( ) acima de 10

5. Como você avalia o seu nível de inglês de uma maneira geral?

1 iniciante 2 básico 3 pré-intermediário 4 intermediário 5 avançado

Como você avalia seu nível de inglês nas seguintes habilidades? Assinale de acordo com a legenda abaixo:

1 iniciante 2 básico 3 pré-intermediário 4 intermediário 5 avançado

	1	2	3	4	5
6. Ler					
7. Escrever					
8. Ouvir					
9. Falar					

Parte 2 – CONHECIMENTO SOBRE O EPLIS

10. De forma geral, como você avalia o seu nível de conhecimento sobre o EPLIS?

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

Ainda sobre o EPLIS, como você avalia o seu nível de conhecimento sobre os seguintes itens? Assinale de acordo com a legenda abaixo:

1 nenhum 2 pouco 3 razoável 4 bom 5 ótimo

	1	2	3	4	5

11. objetivo do EPLIS					
12. conteúdo avaliado no EPLIS					
13. tarefas da fase 1 do EPLIS					
14. tarefas da fase 2 do EPLIS					
15. a tabela de níveis de proficiência da OACI					
16. o desempenho necessário para obter o nível mínimo (nível 4)					
17. manual do candidato					

18. Com que frequência os professores da série atual mencionam o EPLIS em sala de aula?

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

19. Com que frequência os professores da série atual discutem aspectos específicos do exame em sala de aula?

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

Parte 3 – PREPARAÇÃO PARA O EPLIS

20. Com que frequência os professores da série atual trabalham tarefas semelhantes às do EPLIS em sala de aula?

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

Considerando a série atual, com que frequência as seguintes atividades realizadas em sala de aula estão relacionadas à AVIAÇÃO E AO CONTROLE DE TRÁFEGO AÉREO? Assinale de acordo com a legenda abaixo:

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

	1	**2**	**3**	**4**	**5**
21. Compreensão de áudios e vídeos					
22. Compreensão de textos escritos					
23. Prática da produção oral (*speaking*)					
24. Prática de vocabulário					
25. Prática de itens gramaticais					
Outras? Especifique					

Considerando a série atual, com que frequência você realiza as seguintes atividades em sala de aula? Assinale de acordo com a legenda abaixo:

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

	1	**2**	**3**	**4**	**5**

26. Responder, oralmente, perguntas colocadas pelo(a) professor(a) sobre aviação e controle de tráfego aéreo.					
27. Discutir com os colegas um assunto específico sobre aviação e controle de tráfego aéreo.					
28. Fazer uma apresentação oral sobre um assunto relacionado à aviação e ao controle de tráfego aéreo.					
29. Descrever uma figura relacionada à aviação e ao controle de tráfego aéreo.					
30. Contar uma história com base em uma figura relacionada à aviação e ao controle de tráfego aéreo.					
31. Recontar uma história com base em um áudio ou texto relacionado à aviação e ao controle de tráfego aéreo.					
32. Explicar, oralmente, uma situação-problema de tráfego aéreo.					
33. Dar uma sugestão, oralmente, para um problema de tráfego aéreo.					
34. Simular, com um colega, um diálogo entre piloto e controlador.					
35. Responder questões sobre um áudio ou vídeo envolvendo uma conversa entre piloto e controlador.					
36. Responder questões sobre um texto cujo tema está relacionado à aviação e ao controle de tráfego aéreo.					
37. Escrever um diálogo entre piloto e controlador.					
38. Ler em voz alta um diálogo entre piloto e controlador.					
39. Escrever uma história envolvendo uma situação complicadora de tráfego aéreo.					
40. Explicar, oralmente, o significado de uma palavra ou expressão ligada à aviação e ao controle de tráfego aéreo.					
Outras? Especifique:					

Com que frequência você se envolve nas seguintes atividades EXTRA-CLASSE? Assinale de acordo com a legenda abaixo:

1 nunca 2 raramente 3 às vezes 4 frequentemente 5 sempre

	1	**2**	**3**	**4**	**5**
41. Participo de atividades extras oferecidas pelo Galpão de inglês.					
42. Converso com os colegas em inglês para praticar para o exame.					
43. Tiro dúvidas sobre o exame e seu conteúdo com os professores de inglês.					
44. Participo de grupos de estudos voltados para o exame.					
45. Faço os exercícios de áudio disponíveis no *site* do exame.					

46. Faço aulas com professores particulares voltadas para o exame.					
47. Ouço áudios em inglês de conversas entre controladores e pilotos.					
48. Acompanho blogs sobre aviação e controle de tráfego aéreo.					
49. Assisto aos vídeos da fase 2 do EPLIS disponíveis no *site* do exame.					
50. Faço uso do *sample* da fase 2 do EPLIS disponível no *site* do exame.					
Outras? Especifique:					

51. Como você avalia o tempo que você destina para se preparar para o exame?

1 muito aquém do necessário
2 insuficiente
3 razoável
4 suficiente
5 mais do que suficiente

Parte 4 - VISÕES E PERCEPÇÕES SOBRE O EXAME

Considerando a sua relação com o EPLIS, analise a alternativa que mais expressa o seu nível de concordância ou discordância com as seguintes afirmações, na qual:

1 Discordo totalmente 2 Discordo 3 Nem concordo e nem discordo 4 Concordo 5 Concordo totalmente

	1	**2**	**3**	**4**	**5**
52. O EPLIS me motiva a melhorar meu nível de inglês.					
53. O EPLIS modificou minha maneira de estudar inglês.					
54. O EPLIS me força a estudar inglês fora da sala de aula.					
55. O EPLIS aumenta consideravelmente minha carga de estudos.					
56. O EPLIS aumenta o meu nível de ansiedade/*stress*.					

De modo geral, analise a alternativa que mais expressa o seu nível de concordância ou discordância com as seguintes afirmações, na qual:

1 Discordo totalmente 2 Discordo 3 Nem concordo e nem discordo 4 Concordo 5 Concordo totalmente

	1	**2**	**3**	**4**	**5**
57. O EPLIS representa bem as tarefas de um controlador de tráfego aéreo.					

58. O EPLIS avalia o nível de inglês dos alunos precisamente.					
59. O sucesso no EPLIS é importante para o bom desempenho de um controlador de tráfego aéreo.					
60. As aulas de inglês na EEAR preparam bem os alunos para o EPLIS.					
61. Os projetos desenvolvidos nas disciplinas de língua inglesa na EEAR contribuem para o sucesso no EPLIS.					
62. As avaliações de inglês na EEAR estão em consonância com o EPLIS.					
63. A realização do EPLIS durante o curso de formação deve ser opcional.					
64. O resultado no EPLIS deve ser considerado para a nota final na disciplina de inglês da 4ª série.					

Obrigada por ter respondido ao questionário. Sua contribuição será de grande valia! Paula

APÊNDICE E – Ficha de observação de aula

| FICHA DE OBSERVAÇÃO DE AULA (FOA) | | |

APÊNDICE F – Direcionamento para grupo focal

1) Como vocês avaliam o nível de conhecimentos que possuem sobre o EPLIS?

2) Qual a relevância do EPLIS para vocês, alunos do curso BCT?

3) Qual a influência do EPLIS no ensino e na aprendizagem de inglês na EEAR, em termos de conteúdo, habilidades, tópicos, metodologia, avaliações, entre outros?

4) Qual a ênfase das aulas de inglês na EEAR?

5) Como vocês se preparam para o exame?

6) Como vocês avaliam a preparação para o exame na EEAR?

7) Em quais aspectos ela poderia ser melhorada?

8) Quais os maiores desafios e dificuldades na preparação para o exame?

9) Os professores costumam trabalhar estratégias para o exame? Descreva-as.

10) Vocês recebem incentivo em relação ao EPLIS?

APÊNDICE G – TCLE do aluno

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

(Resolução 466/2012 do CNS)

Os efeitos retroativos do EPLIS no ensino e na aprendizagem de inglês na formação do controlador de tráfego aéreo.

Número do CAAE: 52971216.9.0000.5404

Eu, Paula Ribeiro e Souza, estudante de Doutorado do Programa de Pós-Graduação em Linguística Aplicada, da Universidade Estadual de Campinas − UNICAMP, convido-o(a) a participar como voluntário da pesquisa "Os efeitos retroativos do EPLIS no ensino e na aprendizagem de inglês na formação do controlador de tráfego aéreo", sob a orientação da Profa. Dra. Matilde Virgínia Ricardi Scaramucci.

Este documento, chamado **Termo de Consentimento Livre e Esclarecido**, visa assegurar seus direitos como participante e é elaborado em duas vias, uma que deverá ficar com você e outra com a pesquisadora. Por favor, leia com atenção. Se houver perguntas antes ou mesmo depois de assiná-lo, você poderá esclarecê-las com esta pesquisadora. Se preferir, pode levar para casa e consultar seus familiares ou outras pessoas antes de decidir participar. Se você não quiser participar ou quiser retirar sua autorização, não haverá nenhum tipo de penalização ou prejuízo.

Justificativa e objetivos:

A proposta deste estudo é investigar o impacto do EPLIS em diferentes aspectos do ensino e da aprendizagem de inglês do controlador de tráfego aéreo no Brasil. Apesar da alta-relevância desse exame para os profissionais que se submetem a ele e também para a sociedade, ainda não foram conduzidos estudos sistemáticos sobre os efeitos retroativos desse exame no contexto escolar. Você está sendo convidado a participar desta pesquisa por compor o corpo de alunos do Curso de Controle de Tráfego Aéreo, da Escola de Especialistas de Aeronáutica, instituição onde o estudo será realizado.

Procedimentos:

Você está sendo convidado(a) a participar de três etapas, descritas no quadro abaixo. Assinale as etapas com as quais você concorda em participar. Caso haja alguma tarefa da qual você não gostaria de participar, deixe o campo em branco ou escreva a palavra "não":

( )	Responder a um questionário a respeito das suas percepções sobre o ensino e a aprendizagem de inglês no curso de controle de tráfego aéreo e sobre a influência do EPLIS no contexto escolar e na sua preparação. Para essa tarefa, você deverá disponibilizar 1 (uma) hora, no máximo.
( )	Participar de um grupo focal para o aprofundamento das questões levantadas nos questionários. O grupo focal funciona como uma entrevista em grupo, na qual os participantes expressam suas opiniões interagindo com os demais. O grupo focal será

	realizado em uma das salas de aulas da EEAR, em dia e horário que for mais conveniente a todos os participantes, e com duração mínima prevista de 45 minutos e máxima de 1hora e meia.
( )	Ter algumas aulas de seus professores de inglês observadas pela pesquisadora. Serão observadas 12 horas de aula. A observação de aula não ultrapassará 3 horas por dia e deverá ocorrer em dias não consecutivos.

Para o grupo focal (entrevista em grupo) e as observações de aulas está previsto o uso de um gravador de voz a fim de facilitar a recuperação aos dados, minimizando o risco de perda de informações relevantes. Porém, sabe-se que seu uso pode inibir os participantes a exporem sua opinião. Por essa razão, no final deste termo, será solicitada sua autorização para a gravação em áudio das informações coletadas.

Por questões de tempo e logística, o grupo focal será composto por, no máximo, 12 alunos do Curso de Controle de Tráfego Aéreo que aceitaram participar dessa etapa da pesquisa. Participarão do grupo focal somente aqueles que preencheram o questionário e concordaram com a gravação em áudio das informações coletadas. Aqueles que não preencheram o questionário ou que não consentiram a gravação em áudio das informações coletadas serão excluídos. Se o número de voluntários aptos a participar dessa etapa for maior que 12, os seguintes critérios de seleção serão aplicados: maior disponibilidade e maior desempenho na disciplina de inglês. Os demais participantes serão dispensados dessa etapa. Então, como participante, você não necessariamente terá que participar de todas as etapas da pesquisa.

Como o início das observações está previsto para ocorrer juntamente com o início das aulas, os alunos serão distribuídos em turmas, levando-se em consideração, além do nível de proficiência, o consentimento dado para que esta pesquisadora assista a algumas aulas. As observações de aulas ocorrerão apenas nas turmas em que TODOS os alunos se manifestaram a favor das observações. Caso, na turma observada, exista algum aluno que não autorizou a observação de aulas, ele será realocado em outra turma, de acordo com o nível de proficiência. Para as observações, o uso do gravador poderá ser dispensado. Então, caso haja algum aluno que não tenha consentido a gravação em áudio, esta pesquisadora irá registrar os dados somente por meio de anotações.

Todo o material coletado (questionários, gravações, transcrições e anotações) será destruído assim que esta pesquisa for finalizada.

Desconfortos e riscos:

Esta pesquisa não apresenta riscos previsíveis. Acredita-se, porém, que a participação na pesquisa pode gerar pequeno desconforto e incômodo aos participantes, uma vez que eles poderão sentir-se constrangidos ao expressar sua opinião sobre as questões levantadas nos questionários e no grupo focal. Para minimizar esse desconforto, será assegurado total anonimato na divulgação dos resultados da pesquisa. Além disso, será também garantido o direito dos participantes de se absterem a responder qualquer pergunta que julgarem embaraçosa. A observação de aulas também é um instrumento de coleta de dados que poderá gerar desconforto. A fim de que a presença da pesquisadora seja menos intrusiva para os alunos, as primeiras observações serão mais curtas e esse período de observação será prolongado aos poucos. Dessa forma, tanto os alunos quanto os professores irão, naturalmente, se acostumar com a presença da pesquisadora em sala de aula e irão agir naturalmente como se a pesquisadora não estivesse presente.

Benefícios:

Acredita-se que esta pesquisa poderá gerar benefícios em três diferentes níveis. No nível de tomada de decisões, a pesquisa busca o aperfeiçoamento do EPLIS. Por ser um exame de alta-relevância, o EPLIS funciona como um instrumento político, capaz de direcionar ações e atitudes tomadas no contexto de ensino e aprendizagem. No nível pedagógico, a pesquisa busca contribuir para o desenvolvimento profissional dos professores participantes por meio da reflexão crítica sobre sua prática pedagógica. O trabalho em conjunto com os professores, durante as observações, tem o potencial de afetar suas aulas positivamente, uma vez que suas atitudes e decisões estarão mais bem fundamentadas. Os alunos, por sua vez, irão se beneficiar de aulas mais comunicativas, voltadas para o uso real da língua e para suas necessidades enquanto profissionais, o que pode levar à elevação do nível de proficiência em língua inglesa desses alunos. No nível teórico, espera-se que esta pesquisa contribua para a elaboração de uma teoria satisfatória sobre efeito retroativo, buscando compreender a complexidade do fenômeno, suas causas e os fatores que irão interagir na determinação de seu impacto.

Acompanhamento e assistência:

Durante o período da pesquisa, esta pesquisadora estará à disposição para esclarecer qualquer dúvida sobre o estudo, seus objetivos e os procedimentos para a coleta de dados.

Sigilo e privacidade:

Sua participação é voluntária e a qualquer momento você poderá desistir de participar da pesquisa e retirar seu consentimento. Sua recusa ou desistência não lhe trará nenhum prejuízo profissional, seja em sua relação à pesquisadora, à Instituição onde estuda ou à Universidade Estadual de Campinas.

Todas as informações obtidas durante a pesquisa são confidenciais, sendo assegurado o sigilo sobre sua participação em todas as etapas do estudo. No caso das gravações de áudio, elas serão transcritas por esta pesquisadora, na íntegra, mantendo-as o mais fidedignas possível e preservando a identidade dos participantes. Caso haja menção a nomes, a eles serão atribuídas letras, com garantia de anonimato nos resultados e publicações, impossibilitando sua identificação.

Ressarcimento:

A pesquisa será conduzida na própria escola onde o aluno estuda, respeitando seu horário de descanso. Por esse motivo, não haverá nenhum tipo de ressarcimento de despesas. Porém, caso ocorra qualquer gasto aos participantes, ele será ressarcido por esta pesquisadora.

Contato:

Em caso de dúvidas sobre o estudo, você poderá entrar em contato com esta pesquisadora, na Subdivisão de Língua Estrangeira do Instituto de Controle de Tráfego Aéreo, na Praça Marechal do Ar Eduardo Gomes, n° 50, Jardim, Campus CTA, São José dos Campos – SP, pelo telefone (12) 3945 9240, ou por email paulaprs@icea.gov.br.

Em caso de denúncias ou reclamações sobre sua participação e sobre questões éticas do estudo, você poderá entrar em contato com a secretaria do Comitê de Ética em Pesquisa (CEP) da UNICAMP: Rua: Tessália Vieira de Camargo, 126; CEP 13083-887 Campinas – SP; telefone (19) 3521-8936; fax (19) 3521-7187; e-mail: cep@fcm.unicamp.br

Consentimento livre e esclarecido:

Após ter recebido esclarecimento sobre a natureza da pesquisa, seus objetivos, métodos, benefícios previstos, desconfortos e incômodo que esta possam acarretar,

( )	ACEITO participar do estudo e autorizo o uso do gravador durante TODA a coleta de dados (grupo focal e observações de aulas).
( )	ACEITO participar do estudo, porém autorizo o uso do gravador SOMENTE durante o grupo focal.
( )	ACEITO participar do estudo, porém NÃO autorizo o uso do gravador em NENHUMA etapa da pesquisa.

Nome do(a) participante:

Endereço de e-mail:

Telefone para contato:

_____ Data: ____/_____/_____.
(Assinatura do participante)

Responsabilidade do Pesquisador:

Asseguro ter cumprido as exigências da resolução 466/2012 CNS/MS e complementares na elaboração do protocolo e na obtenção deste Termo de Consentimento Livre e Esclarecido. Asseguro, também, ter explicado e fornecido uma cópia deste documento ao participante. Informo que o estudo foi aprovado pelo CEP perante o qual o projeto foi apresentado. Comprometo-me a utilizar o material e os dados obtidos nesta pesquisa exclusivamente para as finalidades previstas neste documento ou conforme o consentimento dado pelo participante.

_____Data: ____/_____/_____.
(Assinatura do pesquisador)

APÊNDICE H – TCLE do professor

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

(Resolução 466/2012 do CNS)

Os efeitos retroativos do EPLIS no ensino e na aprendizagem de inglês na formação do controlador de tráfego aéreo.

Número do CAAE: 52971216.9.0000.5404

Eu, Paula Ribeiro e Souza, estudante de Doutorado do Programa de Pós-Graduação em Linguística Aplicada, da Universidade Estadual de Campinas − UNICAMP, convido-o(a) a participar como voluntário da pesquisa "Os efeitos retroativos do EPLIS no ensino e na aprendizagem de inglês na formação do controlador de tráfego aéreo", sob a orientação da Profa. Dra. Matilde Virgínia Ricardi Scaramucci.

Este documento, chamado **Termo de Consentimento Livre e Esclarecido**, visa assegurar seus direitos como participante e é elaborado em duas vias, uma que deverá ficar com você e outra com esta pesquisadora. Por favor, leia com atenção. Se houver perguntas antes ou mesmo depois de assiná-lo, você poderá esclarecê-las com esta pesquisadora. Se preferir, pode levar para casa e consultar seus familiares ou outras pessoas antes de decidir participar. Se você não quiser participar ou quiser retirar sua autorização, não haverá nenhum tipo de penalização ou prejuízo.

Justificativa e objetivos:

A proposta deste estudo é investigar o impacto do EPLIS em diferentes aspectos do ensino e da aprendizagem de inglês do controlador de tráfego aéreo no Brasil. Apesar da alta-relevância desse exame para os profissionais que se submetem a ele e também para a sociedade, ainda não foram conduzidos estudos sistemáticos sobre os efeitos retroativos desse exame no contexto escolar. Você está sendo convidado a participar desta pesquisa por compor o corpo de professores de inglês da Escola de Especialistas de Aeronáutica, instituição responsável pela formação inicial do controlador de tráfego aéreo e onde o estudo será realizado.

Procedimentos:

Você está sendo convidado(a) a participar de quatro etapas, descritas no quadro abaixo. Assinale as etapas das quais você tem interesse em participar. Caso haja alguma tarefa da qual você não gostaria de participar, deixe o campo em branco:

( )	Responder a um questionário a respeito de suas percepções sobre o ensino e a aprendizagem de língua inglesa na formação do controlador de tráfego aéreo e sobre a influência do exame no contexto escolar e na sua prática pedagógica. Para essa tarefa, você deverá disponibilizar uma hora, no máximo.
( )	Participar de um grupo focal para o aprofundamento das questões levantadas nos questionários. O grupo focal funciona como uma entrevista em grupo, na qual os participantes expressam suas opiniões interagindo com os demais. O grupo focal será realizado em uma das salas de aulas da EEAR, em dia e horário que for mais conveniente a todos os participantes, e com duração mínima prevista de 45 minutos e máxima de 1hora e meia.

( )	Ter suas aulas observadas pela pesquisadora durante a disciplina de inglês aeronáutico para os alunos do curso básico de Controle de Tráfego Aéreo. Serão observadas 12 horas de aula. A observação de aula não ultrapassará 3 horas por dia e deverá ocorrer em dias não consecutivos.
( )	Participar de entrevistas individuais com esta pesquisadora para discussão dos dados coletados durante as observações de aulas. Essa reunião acontecerá após cada observação de aula e terá a duração máxima de 1 hora. Serão realizados quatro encontros com cada professor participante.

Para o grupo focal (entrevista em grupo), as observações de aulas e as entrevistas está previsto o uso de um gravador de voz para facilitar a recuperação e acesso aos dados, minimizando o risco de perda de informações relevantes. Porém, sabe-se que seu uso pode inibir os participantes a exporem sua opinião. Por essa razão, no final deste termo, será solicitada sua autorização para a gravação em áudio das informações coletadas.

O grupo focal com os professores será composto por, no máximo, 12 participantes que aceitaram participar dessa etapa da pesquisa e que concordaram com a gravação Caso o número de professores interessados em participar do grupo focal seja maior que 12, dois grupos focais serão agendados. Participarão do grupo focal somente aqueles que concordarem com a gravação em áudio.

Por questões de tempo e logística, serão acompanhadas as aulas de apenas quatro professores que mostraram interesse em participar dessa etapa. Esses professores serão selecionados com base nos seguintes critérios: a) estar lecionando para a 4ª série do curso de Controle de Tráfego Aéreo; b) ter maior disponibilidade para entrevistas individuais; e c) ter maior tempo de experiência no ensino de inglês aeronáutico na EEAR. Então, como participante, não necessariamente você terá que participar de todas as etapas da pesquisa.

Os professores que aceitarem ter suas aulas assistidas irão necessariamente participar das entrevistas individuais. Nessas duas últimas etapas, o uso do gravador pode ser dispensado. Então, caso haja algum professor que queira ter suas aulas assistidas, porém não consentiu a gravação em áudio, esta pesquisadora irá registrar os dados somente por meio de anotações. Só serão observadas aulas nas turmas em que TODOS os alunos tenham autorizado a presença da professora em sala de aula.

Todo o material coletado (questionários, gravações, transcrições e anotações) será destruído assim que esta pesquisa for finalizada.

Desconfortos e riscos:

Esta pesquisa não apresenta riscos previsíveis. Acredita-se, porém, que a participação na pesquisa pode gerar pequeno desconforto e incômodo aos participantes, uma vez que eles poderão sentir-se constrangidos ao expressar sua opinião sobre as questões levantadas. Para minimizar esse desconforto, será assegurado total anonimato na divulgação dos resultados da pesquisa. Além disso, será também garantido o direito dos participantes de se absterem a responder qualquer pergunta que julgarem embaraçosa. A observação de aulas também é um instrumento de coleta de dados que poderá gerar desconforto. Pretende-se, portanto, estabelecer com os professores participantes uma relação de confiança mútua. Somente quando o professor se sentir mais à vontade com esta pesquisadora, serão realizadas as observações em sala de aula.

Benefícios:

Acredita-se que esta pesquisa poderá gerar benefícios em três diferentes níveis. No nível de tomada de decisões, a pesquisa busca o aperfeiçoamento do EPLIS. Por ser um exame de alta-relevância, o EPLIS funciona como um instrumento político, capaz de direcionar ações e atitudes tomadas no contexto de ensino e aprendizagem. No nível pedagógico, a pesquisa busca contribuir para o desenvolvimento profissional dos professores participantes por meio da reflexão crítica sobre sua prática pedagógica. O trabalho em conjunto com os professores, durante as observações, tem o potencial de afetar suas aulas positivamente, uma vez que suas atitudes e decisões estarão mais bem fundamentadas. Os alunos, por sua vez, irão se beneficiar de aulas mais comunicativas, voltadas para o uso real da língua e para suas necessidades enquanto profissionais, o que pode levar à elevação do nível de proficiência em língua inglesa desses alunos. No nível teórico, espera-se que esta pesquisa contribua para a elaboração de uma teoria satisfatória sobre efeito retroativo, buscando compreender a complexidade do fenômeno, suas causas e os fatores que irão interagir na determinação de seu impacto.

Acompanhamento e assistência:

Durante o período da pesquisa, esta pesquisadora estará à disposição para esclarecer qualquer dúvida sobre o estudo, seus objetivos e os procedimentos para a coleta de dados. Esta pesquisadora também compromete-se a colaborar com o desenvolvimento profissional dos professores participantes.

Sigilo e privacidade:

Sua participação é voluntária e a qualquer momento você poderá desistir de participar da pesquisa e retirar seu consentimento. Sua recusa ou desistência não lhe trará nenhum prejuízo profissional, seja em sua relação à pesquisadora, à Instituição em que trabalha ou à Universidade Estadual de Campinas.

Todas as informações obtidas durante a pesquisa são confidenciais, sendo assegurado o sigilo sobre sua participação em todas as etapas do estudo. No caso das gravações de áudio, elas serão transcritas por esta pesquisadora, na íntegra, mantendo-as o mais fidedignas possível e preservando a identidade dos participantes. Todo o material transcrito será apresentado aos participantes para validação das informações. Caso haja menção a nomes, a eles serão atribuídas letras, com garantia de anonimato nos resultados e publicações, impossibilitando sua identificação.

Ressarcimento:

A pesquisa será conduzida no próprio ambiente de trabalho do professor respeitando seu horário de expediente. Por esse motivo, não haverá nenhum tipo de ressarcimento de despesas. Porém, caso ocorra qualquer gasto aos participantes, ele será ressarcido por esta pesquisadora.

Contato:

Em caso de dúvidas sobre o estudo, você poderá entrar em contato com esta pesquisadora, na Subdivisão de Língua Estrangeira do Instituto de Controle de Tráfego Aéreo, na Praça Marechal do Ar Eduardo Gomes, n° 50, Jardim, Campus CTA, São José dos Campos – SP, pelo telefone (12) 3945 9240, ou por email paulaprs@icea.gov.br.

Em caso de denúncias ou reclamações sobre sua participação e sobre questões éticas do estudo, você poderá entrar em contato com a secretaria do Comitê de Ética em Pesquisa (CEP) da UNICAMP: Rua: Tessália Vieira de Camargo, 126; CEP 13083-887 Campinas – SP; telefone (19) 3521-8936; fax (19) 3521-7187; e-mail: cep@fcm.unicamp.br

Consentimento livre e esclarecido:

Após ter recebido esclarecimento sobre a natureza da pesquisa, seus objetivos, métodos, benefícios previstos, desconfortos e incômodo que esta possam acarretar, aceito participar do estudo:

( )	ACEITO participar do estudo e autorizo o uso do gravador durante TODA a coleta de dados (grupo focal, observações de aulas e entrevistas).
( )	ACEITO participar do estudo, porém autorizo o uso do gravador SOMENTE durante o grupo focal.
( )	ACEITO participar do estudo, porém NÃO autorizo o uso do gravador em NENHUMA etapa da pesquisa.

Nome do(a) participante:

Endereço de e-mail:

Telefone para contato:

_____ Data: ____/_____/_____.
(Assinatura do participante)

Responsabilidade do Pesquisador:

Asseguro ter cumprido as exigências da resolução 466/2012 CNS/MS e complementares na elaboração do protocolo e na obtenção deste Termo de Consentimento Livre e Esclarecido. Asseguro, também, ter explicado e fornecido uma cópia deste documento ao participante. Informo que o estudo foi aprovado pelo CEP perante o qual o projeto foi apresentado. Comprometo-me a utilizar o material e os dados obtidos nesta pesquisa exclusivamente para as finalidades previstas neste documento ou conforme o consentimento dado pelo participante.

_____ Data: ____/_____/_____.
(Assinatura do pesquisador)

ANEXO A – Parecer do CEP

COMITÊ DE ÉTICA EM PESQUISA DA UNICAMP - CAMPUS CAMPINAS

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA Título da Pesquisa: Os efeitos retroativos do EPLIS no ensino e na aprendizagem de inglês na formação do controlador de tráfego aéreo

Pesquisador: Paula Ribeiro e Souza

Área Temática: Versão: 2 **CAAE:** 52971216.9.0000.5404 **Instituição Proponente:** Instituto de Estudos da Linguagem **Patrocinador Principal:** Financiamento Próprio

DADOS DO PARECER Número do Parecer: 1.436.512

Apresentação do Projeto:

Em 2003, a Organização da Aviação Civil Internacional (OACI) adotou um conjunto de Padrões e Práticas Recomendadas – os Requisitos de Proficiência Linguística (LPR) - com o objetivo de assegurar que todos os profissionais (pilotos, controladores de tráfego aéreo e operadores de estação aeronáutica), envolvidos com a aviação internacional, tivessem proficiência linguística suficiente para lidar com situações inesperadas.

Essas recomendações foram motivadas por estudos que constataram que falhas de comunicações no uso da língua inglesa em situações não-rotineiras - isto é, que fogem às situações normais de voo, como as envolvendo desde problemas com passageiros a pane da aeronave - tinham sido determinantes ou contribuintes na ocorrência de alguns acidentes e incidentes aéreos.

Os países membros da OACI - do qual o Brasil é membro signatário - ficariam, assim, responsáveis por desenvolver e aplicar seus próprios exames, tendo sido estipulada a data de 5 de março de 2008 como prazo final para que os Estados se adequassem às novas recomendações. Em 2007, o Exame de Proficiência de Língua Inglesa do SISCEAB (EPLIS) foi aplicado pela primeira vez a 3.347 profissionais. Desenvolvido pelo Instituto de Controle do Espaço Aéreo (ICEA), o EPLIS vem sendo aplicado anualmente desde então a controladores de tráfego aéreo e operadores de estação aeronáutica.

Este trabalho tem como objetivo investigar o impacto do Exame de Proficiência em Língua Inglesa (EPLIS) aplicado aos controladores de tráfego aéreo no Brasil em diferentes aspectos do ensino e da aprendizagem de inglês na formação desses profissionais. Parte-se do pressuposto que um exame não influencia apenas o conteúdo trabalhado em sala de aula, mas todo o currículo, material didático, avaliações de rendimento, metodologias de ensino, as atitudes dos professores e as estratégias de aprendizagem empregadas pelos alunos. Dessa forma, o

"efeito retroativo" (washback effect) é definido como o impacto ou influência que um exame exerce nos contextos de ensino e aprendizagem e também na sociedade.

De natureza complexa, esse efeito interage com outras forças presentes na sociedade, tais como, as crenças dos professores e alunos, suas experiências, as exigências das instituições em que estão inseridos, além de questões políticas, sociais e econômicas. Acredita-se que exames de alta relevância tendem a provocar maiores mudanças quando comparados a exames com relevância menor.

Porém, apesar dos impactos que os exames de alta-relevância exercem tanto no contexto de ensino e de aprendizagem quanto na sociedade em geral, as pesquisas sobre o assunto ainda são poucas e carecem de mais evidências empíricas, o que torna necessário a coleta de dados sobre os efeitos retroativos de exames no modo como professores e alunos agem em sala de aula.

Estudos sobre efeito retroativo têm apontado para a complexidade desse fenômeno ao alegar que a intensidade, especificidade, duração e intencionalidade do efeito retroativo irão variar de participante para participante e entre eles (professores, alunos, coordenadores, diretores, entre outros), assim como também em função do contexto.

O objetivo principal do estudo é investigar o efeito retroativo de um exame, considerado de alta-relevância, nas percepções e comportamento de alunos e professores em um contexto específico, e buscar compreender i) em que circunstâncias o efeito retroativo ocorre, ii) qual sua intensidade e iii) os diversos fatores que irão interagir com o exame na definição de seu efeito.

De fato, nos estudos sobre efeito retroativo, as percepções e comportamentos de professores têm sido amplamente pesquisados pelo papel chave que eles exercem. Os professores são os atores mais visíveis e mais estudados por serem intermediadores entre o exame e a sala de aula. Porém, carece-se de mais pesquisas que analisem o efeito retroativo dos exames nos alunos e na sua aprendizagem. Pouco ainda se sabe sobre os efeitos retroativos de um teste em diversos aspectos relacionados aos alunos e como eles se preparam para um teste. Esta pesquisa, portanto, busca suprir essa necessidade de maiores investigações sobre as percepções e atitudes de alunos em relação a um exame, ao ter também como foco de estudo o aluno, principal envolvido em um processo avaliativo, uma vez que é o mais afetado pelas implicações que os resultados nesse exame podem trazer.

No contexto de ensino de inglês para fins específicos, particularmente em relação ao ensino de inglês para controle de tráfego aéreo no Brasil, ainda não há investigação sobre o efeito retroativo do EPLIS. Acredita-se que, dada a relevância do EPLIS para a vida profissional dos que se submetem a ele e também para a sociedade, existam de fato efeitos retroativos no sistema educacional em relação à disciplina de inglês.

Espera-se, portanto, que esta pesquisa contribua para a elaboração de uma teoria satisfatória sobre efeito retroativo, buscando compreender a complexidade do fenômeno, suas causas e os fatores que interagem na determinação de seu impacto. Além de esclarecimentos sobre esse fenômeno tão complexo, um estudo a respeito do efeito retroativo do EPLIS pode propiciar orientações para que exames da mesma natureza sejam capazes de maximizar seus efeitos benéficos para o ensino de inglês na aviação, elevando o nível de proficiência dos profissionais de tráfego aéreo, e também para restringir ao máximo seus potenciais efeitos negativos no ensino e na aprendizagem.

Metodologia Proposta:

A pesquisa será conduzida na Escola de Especialistas da Aeronáutica (EEAR), na cidade de Guaratinguetá - SP, onde é oferecido o curso de formação em Controle de Tráfego Aéreo. Serão convidados a participar da pesquisa todos os professores de língua inglesa que atuam nesse curso e seus alunos (cerca de 20 docentes e 300 alunos).

Os dados provenientes das diversas fontes de coleta serão triangulados a fim de se obter um desenho do fenômeno com múltiplas perspectivas. Esse procedimento irá também aumentar a confiabilidade do estudo ao prover ferramentas capazes de distinguir os efeitos produzidos pelo teste daqueles relacionados a outros fatores, sendo somente os primeiros considerados efeitos retroativos do exame.

- Aplicação de Questionários: Uma das vantagens de utilizar questionários é a praticidade que eles oferecem na obtenção de dados numéricos de um número grande de indivíduos. Os questionários, destinados tanto os professores como aos alunos, gerarão dados sobre suas percepções e atitudes em relação ao ensino e à aprendizagem de inglês.

- Grupos Focais: Por meio do grupo focal, é possível coletar dados confiáveis e válidos, dentro de um espaço curto de tempo e sem custos elevados. O que o difere da entrevista é o fato de propiciar uma interação entre os participantes que irão discutir sobre o assunto proposto. Essa interação é uma forma muito rica de obter informações, no sentido de que os pontos de vista diferentes tendem a ser justificados em pormenores. O grupo focal com os professores será composto por, no máximo, 12 participantes. Caso o número de professores interessados em participar do grupo focal seja maior que 12, dois grupos focais serão agendados. Será também agenda um grupo focal composto por alunos (no máximo 12 alunos) da 4a série do curso, selecionados com base, primeiramente, na disponibilidade. Os grupos focais e as entrevistas serão agendados em local e horário mais conveniente aos participantes e de comum acordo entre eles.

- Observações de Aulas: Uma das críticas às primeiras pesquisas sobre efeito retroativo era a falta de dados obtidos originados de observações de aulas, não havendo, portanto, evidências empíricas de que as percepções dos professores e alunos refletiam o que realmente ocorria em sala de aula. As observações se fazem apropriadas para este tipo de estudo, porque, por meio delas, é possível presenciar as mudanças no seu cenário natural. Observações em sala de aula são essenciais para que o pesquisador "não se apoie somente no que professores e alunos dizem, uma vez que essas percepções podem não refletir o que eles realmente fazem". Serão observadas as aulas de, no máximo, 4 professores. Essas observações deverão totalizar 12 horas das aulas de cada um desses quatro professores - não ultrapassando 3 horas por dia e ocorrendo em dias não consecutivos. Essas observações mostrarão o que acontece em sala de aula a partir da perspectiva da pesquisadora.

- Entrevistas: Para este estudo, optou-se por entrevistas semiestruturadas, uma vez que haverá um plano geral dos tópicos que serão tratados durante a entrevista. Esses tópicos de discussão serão originados a partir das reflexões feitas pelo pesquisador nas observações em sala de aula. As entrevistas - produtos da interação entre entrevistador e entrevistado - funcionarão como espaços para que os professores possam expor suas reações, intenções e razões perante os dados coletados de sua prática pedagógica. Elas tendem a aumentar a confiabilidade dos dados ao permitir que a interpretação da pesquisadora sobre um evento seja confrontada com a do professor(a), legitimando-a ou modificando-a. A princípio, as entrevistas serão conduzidas apenas com os professores, individualmente, sendo previstos quatro encontros com cada professor participante, após cada observação de suas aulas. Optou-se por não conduzir entrevistas com os alunos para que não houvesse prejuízo a outras atividades presentes no currículo escolar.

Objetivo da Pesquisa:

O objetivo principal da pesquisa é analisar as consequências, sejam elas intencionais ou não intencionais, do EPLIS a fim de melhorá-lo nos aspectos que podem, potencialmente, apresentar problemas. No nível pedagógico, acredita-se que o aperfeiçoamento desse exame irá, consequentemente, afetar positivamente o ensino e a

aprendizagem de língua inglesa, uma vez que exames de alta-relevância como esse têm o potencial de direcionar mudanças pedagógicas.

Objetivo Primário: Este estudo tem como objetivo investigar o efeito retroativo de um exame de proficiência linguística, como parte das recomendações da Organização da Aviação Civil Internacional, no ensino e aprendizagem de inglês na formação do controlador de tráfego aéreo.

Objetivo Secundário: a) Analisar as percepções e atitudes dos professores em relação ao EPLIS no contexto escolar e nas suas práticas de sala de aula. b) Analisar as percepções e atitudes dos alunos em relação ao efeito do EPLIS nas suas estratégias de aprendizagem e na preparação para o exame.

Avaliação dos Riscos e Benefícios:

Riscos: Não existem riscos previsíveis nesta pesquisa. Acredita-se, porém, que o preenchimento do questionário, a participação no grupo focal e nas entrevistas podem causar certo desconforto nos professores e alunos, uma vez que os participantes podem sentir-se constrangidos ao terem que expressarem sua opinião sobre as questões levantadas. Para minimizar esse desconforto, será assegurado aos participantes total anonimato na divulgação dos resultados da pesquisa. Além disso, será também garantido o direito dos participantes de se absterem a responder qualquer pergunta que julgarem embaraçosa. A observação de aulas também é um instrumento de coleta de dados que pode gerar desconforto. Para minimizar o impacto, pretende-se, primeiramente, estabelecer com os professores uma relação de confiança mútua. Somente quando o professor se sentir mais à vontade com esta pesquisadora, serão realizadas as observações em sala de aula. A fim de que a presença do professor seja menos intrusiva para os alunos, as primeiras observações terão tempo de duração reduzido e esse período será prolongado aos poucos. Existe também a previsão de utilização de gravador de voz durante o grupo focal, a observação de aulas e as entrevistas. O gravador de voz possibilita que a recuperação e acesso aos dados sejam feitos de forma prática e mais confiável, sem o risco de perda de informações relevantes. Porém, ele pode inibir os participantes a exporem sua opinião. Novamente, será reforçado aos participantes que todas as informações obtidas durante a pesquisa são confidenciais, sendo assegurado o sigilo e anonimato sobre sua participação em todas as etapas do estudo. Caso ainda um participante mostre-se relutante ao uso do gravador, esse será dispensado e o registro de dados dar-se-á por meio de tomada de notas.

Benefícios: Em relação aos benefícios gerados, o objetivo principal da pesquisa é analisar as consequências, sejam elas intencionais ou não, de um exame a fim de melhorá-lo nos aspectos que podem, potencialmente, apresentar problemas para o contexto escolar. Acredita-se que o aperfeiçoamento desse exame irá, consequentemente, afetar positivamente o ensino e a aprendizagem de língua inglesa, uma vez que exames de alta-relevância, como o EPIS, apresentam potencial de direcionar mudanças pedagógicas. Para os professores participantes, o benefício será o crescimento profissional. As entrevistas, que ocorrerão após cada aula observada, funcionarão como espaços para que os

professores possam expor suas reações, intenções e razões perante os dados coletados. Essas reuniões visam o desenvolvimento profissional do professor a partir da reflexão crítica sobre sua prática pedagógica. Acredita-se que esse trabalho com o professor afetará positivamente suas aulas, uma vez que suas atitudes e decisões estarão mais bem fundamentadas.

Os alunos, por sua vez, irão se beneficiar de aulas mais comunicativas, voltadas para o uso real da língua e para suas necessidades enquanto profissionais, o que pode levar à elevação do nível de proficiência em língua inglesa desses alunos. Acredita-se que esses benefícios superam de sobremaneira os pequenos desconfortos causados durante o processo de coleta de dados relatados anteriormente.

Comentários e Considerações sobre a Pesquisa:

Título do projeto na folha de rosto - ok Nome do pesquisador responsável na folha de rosto - ok Nome da representante da unidade proponente (nome, função, carimbo) - ok Anuência de instituições externas: - Brigadeiro do Ar Omar Juan Wara - Escola de Especialistas da Aeronáutica (Guaratinguetá/SP) - ok

Proposta de tese de doutoramento.

No campo 'cronograma' do documento gerado pela Plataforma Brasil, as entrevistas e coletas de dados estão previstas para os meses de março a novembro de 2016 - ok

No campo 'orçamento' do documento gerado pela Plataforma Brasil, o pesquisador relata um orçamento de 'R$2500,00'. Esse orçamento é compatível com o orçamento de um projeto de pesquisa financiado pelo próprio pesquisador.

Considerações sobre os Termos de apresentação obrigatória:

São apresentados dois TCLEs distintos: um para os professores e outro para os alunos.

Linguagem acessível ao sujeito da pesquisa - ok Justificativa, objetivos e descrição de procedimentos - ok Desconfortos, riscos e benefícios - ok Garantia de esclarecimentos - ok

Liberdade na recusa ou retirada do consentimento - ok Garantia de Sigilo - ok Menção sobre ressarcimento - ou não - de despesas - ok Anuência sobre a permissão ao armazenamento de material de áudio gravado - ok Menção ao TCLE assinado em duas vias - ok

Menção ao CEP em caso de abusos ou reclamações de cunho ético - ok Nome e contato com o pesquisador da pesquisa - ok Rubrica do pesquisador e do voluntário em TCLEs com mais de uma página - ok

Recomendações: Conclusões ou Pendências e Lista de Inadequações:

Após readequação do TCLE, todos os items previstos pela Resolução 466/2012 foram contemplados.

Considerações Finais a critério do CEP:

- O sujeito de pesquisa deve receber uma via do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (quando aplicável).

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (quando aplicável).

- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado. Se o pesquisador considerar a descontinuação do estudo, esta deve ser justificada e somente ser realizada após análise das razões da descontinuidade pelo CEP que o aprovou. O pesquisador deve aguardar o parecer do CEP quanto à descontinuação, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de uma estratégia diagnóstica ou terapêutica oferecida a um dos grupos da pesquisa, isto é, somente em caso de necessidade de ação imediata com intuito de proteger os participantes.

- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo. É papel do pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.

- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial.

- Relatórios parciais e final devem ser apresentados ao CEP, inicialmente seis meses após a data deste parecer de aprovação e ao término do estudo.

-Lembramos que segundo a Resolução 466/2012, item XI.2 letra e, "cabe ao pesquisador apresentar dados solicitados pelo CEP ou pela CONEP a qualquer momento".

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Situação do Parecer: Aprovado

Necessita Apreciação da CONEP: Não

CAMPINAS, 03 de Março de 2016

Assinado por: Renata Maria dos Santos Celeghini (Coordenador)

Endereço: Rua Tessália Vieira de Camargo, 126 **Bairro:** Barão Geraldo **CEP:** 13.083-887 **UF:** SP **Município:** CAMPINAS **Telefone:** (19)3521-8936 **Fax:** (19)3521-7187 **E-mail:** cep@fcm.unicamp.br

Tipo Documento	Arquivo	Postagem	Situação	Autor
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS _DO_P ROJETO_653265.pdf	01/03/2016 09:25:33	Aceito	
TCLE / Termos de Assentimento / Justificativa de Ausência	tcle_professores_modificado.pdf	01/03/2016 09:24:49	Aceito	Paula Ribeiro e Souza
TCLE / Termos de Assentimento / Justificativa de Ausência	tcle_alunos_modificado.pdf	01/03/2016 09:24:31	Aceito	Paula Ribeiro e Souza
Outros	resposta_pendencias.pdf	01/03/2016 09:23:33	Aceito	Paula Ribeiro e Souza
Projeto Detalhado / Brochura Investigador	projeto_CEP_modificado.pdf	01/03/2016 09:22:45	Aceito	Paula Ribeiro e Souza
Outros	autorizacao_pesquisa.pdf	28/01/2016 14:10:04	Aceito	Paula Ribeiro e Souza
Declaração de Pesquisadores	AtestadoMatricula.pdf	28/01/2016 14:07:48	Aceito	Paula Ribeiro e Souza
Folha de Rosto	folha_de_rosto_coep_assinada.pdf	28/01/2016 13:59:58	Aceito	Paula Ribeiro e Souza

ANEXO B – Escala de níveis de proficiência da OACI

Escala de níveis de proficiência em inglês aeronáutico da OACI

NÍVEL	PRONÚNCIA Pressupõe a utilização de um dialeto e/ou de um sotaque inteligível para a comunidade aeronáutica.	ESTRUTURA A relevância das estruturas gramaticais e dos padrões de oração é determinada pela adequação da linguagem à tarefa a ser executada.	VOCABULÁRIO	FLUÊNCIA	COMPREENSÃO	INTERAÇÕES
Experto **6**	Pronúncia, tonicidade, ritmo e entonação, embora possam ter influência do idioma materno ou variante regional, quase nunca dificultam a compreensão.	As estruturas gramaticais e os padrões de oração, tanto básicos, quanto complexos, são consistentemente bem controlados.	A variedade e a precisão do vocabulário são suficientes para uma comunicação eficaz sobre uma ampla variedade de temas familiares e não familiares. O vocabulário é idiomático, matizado e sensível ao registro.	Capaz de sustentar conversas de forma prolongada com naturalidade e sem esforço. Varia o fluxo da fala para efeitos estilísticos; por exemplo, para enfatizar um argumento. Utiliza espontaneamente os marcadores e conectores discursivos.	A compreensão é constantemente correta em quase todos os contextos, inclusive de sutilezas linguísticas e culturais.	Interage com facilidade em quase todas as situações. É sensível a pistas verbais e não-verbais e responde a elas adequadamente.
Pós-operacional **5**	Pronúncia, tonicidade, ritmo e entonação, embora possam ter influência do idioma materno ou variante regional, raramente dificultam a compreensão.	As estruturas gramaticais e os padrões de oração básicos são consistentemente bem controlados. Há tentativas de uso de estruturas complexas, mas com erros que, às vezes, interferem no sentido.	A variedade e a precisão do vocabulário são suficientes para uma comunicação eficaz sobre temas cotidianos, concretos e relacionados ao trabalho. Capaz de parafrasear sistematicamente e corretamente. O vocabulário é às vezes idiomático.	Capaz de sustentar conversas de forma prolongada, com relativa facilidade, sobre temas familiares, mas pode não variar o fluxo da fala como instrumento estilístico. Capaz de utilizar adequadamente marcadores ou conectores discursivos.	A compreensão é correta para temas cotidianos, concretos e relacionados ao trabalho e na maioria das vezes correta quando o falante é confrontado com uma complicação linguística ou situacional ou uma mudança inesperada dos acontecimentos. Capaz de compreender uma gama de variedades linguísticas (dialeto e/ou sotaques) ou registros.	As respostas são imediatas, adequadas e informativas. Gerencia a relação falante/ouvinte de maneira eficaz.
Operacional **4**	Pronúncia, tonicidade, ritmo e entonação são influenciados pelo idioma materno ou variante regional, mas somente às vezes dificultam a compreensão.	As estruturas gramaticais básicas e os padrões de oração básicos são usados com criatividade e geralmente bem controlados. Erros podem ocorrer, principalmente em circunstâncias incomuns ou inesperadas, mas raramente interferem no sentido.	A variedade e a precisão do vocabulário são geralmente suficientes para uma comunicação eficaz sobre temas cotidianos, concretos e relacionados ao trabalho. Frequentemente capaz de parafrasear com êxito quando falta vocabulário em circunstâncias incomuns ou inesperadas.	Capaz de produzir emunciados em um ritmo adequado. Pode haver perda ocasional de fluência na mudança do discurso ensaiado ou formulaico para uma interação espontânea, mas sem que isso impeça a comunicação efetiva. Utiliza de um modo limitado os conectores ou marcadores do discurso. As pausas preenchidas não são dispersantes.	A compreensão é, na maioria das vezes, correta para temas cotidianos, concretos e relacionados ao trabalho quando o sotaque ou a variedade linguística utilizada é suficientemente inteligível para uma comunidade internacional de usuários. Quando o falante se confronta com uma complicação linguística ou situacional ou uma mudança inesperada dos acontecimentos, a compreensão pode ficar mais lenta ou exigir estratégias de clarificação.	As respostas são geralmente imediatas, adequadas e informativas. Inicia e mantém o diálogo mesmo quando lida com uma mudança inesperada dos acontecimentos. Lida de maneira adequada com aparentes falhas no entendimento verificando, confirmando ou clarificando o que se pretende.
Níveis 1, 2, 3 na página seguinte.						

NÍVEL	PRONÚNCIA Pressupõe a utilização de um dialeto e/ou de um sotaque inteligível para a comunidade aeronáutica.	ESTRUTURA A relevância das estruturas gramaticais e dos padrões de oração é determinada pela adequação da linguagem à tarefa a ser executada.	VOCABULÁRIO	FLUÊNCIA	COMPREENSÃO	INTERAÇÕES
Pré-operacional **3**	Pronúncia, tonicidade, ritmo e entonação são influenciados pelo idioma materno ou variante regional e frequentemente dificultam a compreensão.	As estruturas gramaticais básicas e os padrões de oração básicos associados a situações previsíveis nem sempre são bem dominados. Erros frequentemente interferem no sentido.	A variedade e a precisão do vocabulário são frequentemente suficientes para comunicação sobre temas cotidianos, concretos e relacionados ao trabalho, porém o vocabulário é limitado e a escolha das palavras é, muitas vezes, inadequada. Frequentemente incapaz de parafrasear corretamente quando falta vocabulário.	Capaz de produzir emunciados, mas o fraseamento e as pausas são frequentemente inadequadas. As hesitações ou a lentidão no processamento da linguagem podem impedir uma comunicação efetiva. As pausas preenchidas são, às vezes, um fator de dispersão.	A compreensão é frequentemente correta em temas cotidianos, concretos e relacionados ao trabalho quando o sotaque ou a variedade linguística usada for suficientemente inteligível para uma comunidade internacional de usuários. Pode falhar em compreender uma complicação linguística ou situacional ou uma mudança inesperada dos acontecimentos.	As respostas são às vezes imediatas, adequadas e informativas. Capaz de iniciar e manter diálogos com razoável facilidade sobre temas familiares e em situações previsíveis. Resposta geralmente inadequada perante mudanças imprevistas dos acontecimentos.
Elementar **2**	Pronúncia, tonicidade, ritmo e entonação são fortemente influenciados pelo idioma materno ou variante regional e geralmente dificultam a compreensão.	Demonstra apenas domínio limitado de algumas estruturas gramaticais e padrões de oração simples e memorizados.	Vocabulário limitado, consistindo somente de palavras isoladas e expressões memorizadas.	Capaz de produzir segmentos muito curtos, isolados e memorizados. Para procurar expressões e articular palavras menos familiares, frequentemente utiliza pausas vazias e pausas preenchidas que causam distração no interlocutor.	A compreensão limita-se a frases isoladas e memorizadas, quando elas são cuidadosa e lentamente articuladas.	O tempo de resposta é lento e frequentemente inadequado. A interação limita-se a diálogos simples de rotina.
Pré-Elementar **1**	Desempenho inferior ao nível Elementar.	Desempenho inferior ao nível Elementar.	Desempenho inferior ao nível Elementar.	Desempenho inferior ao nível Elementar.	Desempenho inferior ao nível Elementar.	Desempenho inferior ao nível Elementar.