

O pré-teste ao EPPLÉ (Exame de Proficiência para Professores de Línguas Estrangeiras): validade e desafios tecnológicos

DOI: <http://dx.doi.org/10.21165/el.v48i2.2202>

Jéssica Nunes Caldeira Cunha¹

Resumo

O Exame de Proficiência para Professores de Línguas Estrangeiras (EPPLÉ) tem sido foco de muitas pesquisas no Brasil. Recentemente, desenvolveu-se o pré-teste ao EPPLÉ com o intuito de fornecer aos candidatos uma informação preliminar sobre sua proficiência, que pudesse prever os resultados do exame. Neste artigo, discutimos alguns aprimoramentos feitos no pré-teste em relação a uma versão anterior, as vantagens e limitações de sua aplicação através de dispositivos tecnológicos móveis, e indicações de sua validade paralela.

Palavras-chave: avaliação; línguas estrangeiras; tecnologia.

¹ Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) São José do Rio Preto, São Paulo, Brasil; jessicancc@gmail.com; <http://orcid.org/0000-0002-6247-0214>

Pretest for the Proficiency Exam for Teachers of Foreign Languages (EPPL): Validity and Technological Challenges

Abstract

The Proficiency Exam for Teachers of Foreign Languages (EPPL) has been the object of many studies in Brazil. Recently, the pretest for this examination was developed, aiming to provide examinees with a preliminary proficiency that could predict their results in the exam. In this article, we discuss how the pretest has been improved in comparison to a previous version, the advantages and drawbacks of its application in mobile devices, and indications of its parallel validity.

Keywords: assessment; foreign languages; technology.

1. Introdução

No Brasil, a área de Linguística Aplicada tem apontado diversas questões sobre a qualidade da formação de professores de línguas estrangeiras. Tratando-se em especial de professores de inglês, sua formação é considerada deficiente em uma parcela significativa dos cursos de licenciatura, tanto em seu aspecto pedagógico quanto em seu aspecto linguístico (ALMEIDA FILHO, 1992; CELANI, 2010, 2008; CONSOLO *et al.*, 1999). A boa formação linguística destes professores é essencial para o sucesso do ensino, uma vez que a língua é o instrumento e o conteúdo das aulas de línguas. Para um bom desempenho de suas atividades, é imperativo que estes profissionais tenham nível de proficiência adequado da língua que ensinam(rão), necessária para que possam produzir insumo linguístico oral e escrito, atuando como modelo para seus alunos e promovendo o uso da língua estrangeira em sala de aula, o que contribui para o desenvolvimento desses alunos (CONSOLO, 2006).

Buscando avaliar essa proficiência linguística do professor de línguas estrangeiras, pesquisas no Brasil resultaram na produção do Exame de Proficiência para Professores de Línguas Estrangeiras (EPPL) (CONSOLO *et al.*, 2009), atualmente disponível em sua versão para professores de língua inglesa (doravante, LI). Neste exame, busca-se avaliar a proficiência do professor considerada tanto em seu aspecto global como em aspectos específicos da profissão de professor de línguas, tais como, dar instruções, esclarecer conteúdos, dentre outros. O conjunto destas proficiências é chamado de proficiência linguístico-comunicativo-pedagógica por Consolo e Silva (2014), e constitui o domínio do exame EPPL.

Pesquisas sobre o EPPL foram desenvolvidas através dos anos, especialmente no âmbito do projeto de pesquisa "Avaliação da proficiência linguístico-comunicativo-pedagógica do professor de línguas: operacionalização de construto no Exame de

Proficiência para Professores de Língua Estrangeira (EPPLÉ)". O trabalho de iniciação científica de Agüena (2016), desenvolvido no escopo do referido projeto, introduziu a proposta de um pré-teste para o EPPLÉ oral, que indicaria um resultado preliminar ao candidato, permitindo que este tomasse uma decisão informada sobre sua participação no exame. Além disso, o desenho do pré-teste foi pensado para permitir acessibilidade e praticidade aos candidatos para obterem essa informação prévia sobre sua proficiência, o que leva à utilização de recursos tecnológicos, atualmente muito visados na área de avaliação e de ensino e aprendizagem de línguas estrangeiras.

Neste artigo, discutimos brevemente um recorte de uma pesquisa de mestrado em andamento, que tem por objetivo desenvolver um novo pré-teste, analisar sua validade de critério (em relação ao EPPLÉ) e disponibilizá-lo em dispositivos tecnológicos móveis (celulares *smartphones*).

2. Justificativa e relevância do tema

Nossa pesquisa se posiciona, conforme mencionado anteriormente, num cenário de estudos sobre avaliação de proficiência linguística, tendo como público específico professores de LI em pré-serviço e em serviço no Brasil. É sabido que as avaliações têm poder de gerar o chamado "efeito retroativo", um conceito complexo, que diz respeito, basicamente, "ao impacto que uma avaliação tem no ensino e aprendizagem feitos em sua preparação" (GREEN, 2013, p. 40, tradução minha). O termo já foi bastante discutido e ampliado, e aqui consideramos efeito retroativo não só as consequências que uma avaliação tem nas práticas de instrução, mas também na administração de uma instituição de ensino, nas atitudes e hábitos de professores e alunos e, mais amplamente, na sociedade em geral.

O EPPLÉ foi elaborado tendo entre suas motivações a geração de um efeito retroativo positivo e redirecionador na formação de professores, oferecendo à classe parâmetros cientificamente embasados do que se espera de um bom professor de LI. Assim, as aplicações do exame se propõem a colaborar com o rompimento de um círculo vicioso de formação precária de professores nas universidades, e posterior ensino deficiente nas escolas, que traria para a universidade estudantes com baixa proficiência, completando-se o círculo (CONSOLO, 2017). Esse exame pode ser considerado de alta relevância no sentido de ter abrangência nacional e ser respeitado entre pesquisadores e professores. No entanto, ainda não foi possível divulgá-lo tão amplamente e levá-lo a todos os estados do país, ou, indo mais além, institucionalizá-lo como certificação da classe de professores de LI. Assim sendo, o efeito retroativo e o impacto do EPPLÉ podem ainda ser aumentados consideravelmente.

O pré-teste acrescenta ao EPPLÉ uma série de vantagens: oferecer um "pacote" que contém o exame principal (EPPLÉ), de alta relevância, e um teste de baixa relevância, fácil

acesso e pequena duração vinculado ao principal; proporcionar uma maior visibilidade ao EPPLE, devido à acessibilidade do pré-teste e facilidade de divulgação, tornando público um indicador do construto avaliado no EPPLE; fornecer dados para estudos mais aprofundados sobre a validade dos dois instrumentos (pré-teste e EPPLE), contribuindo para a melhoria de ambos e de suas faixas, dado que estes estudos pressupõem um detalhamento e definição cada vez mais precisos dos construtos dos testes.

Assim, este trabalho vem somar-se aos estudos do EPPLE, tornando-o mais completo, e, portanto, contribuindo para o desenvolvimento de avaliações que possam vir a gerar mudanças sociais nos cursos de formação de professores de LI no Brasil (modificações nos currículos, nas práticas de formadores, nos hábitos e atitudes dos professores em formação, dentre outras) e na sociedade (por exemplo, aumentar o respeito pela profissão através de certificação e de uma identificação com a excelência).

3. O pré-teste ao EPPLE

Conforme propõe Bachman (1990), o caminho para o desenvolvimento de um teste se inicia pela determinação de seu construto, seguida da operacionalização deste em um instrumento de avaliação, que, por fim, deve ser analisado para a definição de regras e procedimentos de mensuração das informações coletadas através dele.

Para o desenvolvimento do construto do pré-teste, nos voltamos para o construto do EPPLE, uma vez que desejamos alinhá-los. Publicado na tese de doutorado de Anchieta (2015), o construto do EPPLE, que não será aqui transcrito por questões de espaço, traz as seguintes características do exame: tem natureza comunicativa; prevê uso contextualizado da língua; analisa o desempenho global do candidato; expressa o resultado em faixas de proficiência; entende proficiência como um conceito amplo (envolvendo aspectos socioculturais); analisa a proficiência linguística geral e específica da profissão de professor de línguas. Além disso, lembramos que o resultado do EPPLE é expresso em duas faixas de proficiência: a oral e a escrita. O pré-teste inicialmente se propôs a prever a proficiência oral.

A primeira versão do pré-teste (AGUENA, 2016) pretendia oferecer um indicativo do resultado do EPPLE oral. Sua realocação na plataforma *Google Forms* decorreu do fato de que a plataforma utilizada anteriormente, *Lingt Classroom*, apresentava dois problemas. O primeiro, com base no estudo anterior de Consolo e Agüena (2017), foi que a plataforma requeria a instalação em navegadores de internet do Adobe Flash Player, um *software* já obsoleto e incompatível com a maioria dos DTMs atuais. Algumas alternativas foram estudadas e gestores de plataformas *web* de teste foram contatados, buscando suprir nossas necessidades e propósitos. Porém, foram descartadas a princípio por serem dispendiosas, com custos de produção e hospedagem do teste. O segundo motivo foi que a plataforma *Lingt* restringiu os serviços para usuários gratuitos, limitando o

número possível de usuários, as possibilidades de questões, dentre outros, que seriam disponibilizados apenas para usuários *premium*.

A seleção especificamente do *Google Forms* decorre de algumas de suas características, como ser gratuito, oferecer o teste a um número de usuários ilimitado, ser uma ferramenta bastante intuitiva a desenvolvedores e usuários, permitir a elaboração de itens variados, oferecer automaticamente uma tabela com as respostas dos candidatos e ser acessível de qualquer dispositivo conectado à internet. Várias plataformas foram pesquisadas, porém todas ofereciam limitações, sendo que, dentre as gratuitas, o *Google Forms* foi definitivamente o mais adequado aos nossos propósitos.

A maior modificação durante a migração do pré-teste para o *Forms* foi a exclusão mandatória da gravação de áudio por parte do aluno, devido à sua indisponibilidade na plataforma. O pré-teste propõe-se a avaliar a proficiência oral em LI e, por isso, foi pensado inicialmente como um teste em que as tarefas envolveriam compreensão e produção oral. No entanto, foi impossível prosseguir com esse projeto e tivemos que abrir mão da produção oral para garantir que o pré-teste fosse hospedado *on-line* e gratuitamente. Desse modo, o pré-teste manteve as questões em áudio (compreensão oral), mas o formato das respostas passou a ser escrito, no computador ou celular. Portanto, a utilização do *Forms*, com as adaptações decorrentes dessa utilização, acarretou mudanças no construto original e nas faixas de proficiência do pré-teste.

O segundo grupo de modificações foi no conteúdo das questões. Mantivemos um número semelhante de questões, diminuindo apenas duas, passando de dez para oito questões. Tal decisão foi tomada visando uma manutenção na duração da avaliação, que é de trinta minutos. As três primeiras questões do pré-teste avaliam proficiência geral em LI, com perguntas de cunho mais pessoal que permitem uma elaboração mais livre da parte do candidato. Das cinco questões restantes, três são de compreensão de um texto oral mais longo e duas são questões metalinguísticas, uma contextualizada como dúvida de um aluno e outra pedindo para explicar um uso linguístico no texto de compreensão oral. Uma vez que todas as questões foram gravadas em áudio e o candidato não as lê, apenas as escuta, todo o teste já avalia compreensão oral através da produção de respostas condizentes com as instruções ouvidas.

Assim, o pré-teste é um teste de compreensão oral, que visa avaliar a proficiência oral do candidato, sendo que a produção escrita nas respostas é analisada com ênfase nos aspectos linguísticos e de interpretação do texto que foi ouvido pelo candidato. O tratamento das respostas enfatiza aspectos como tópico (se o candidato respondeu à pergunta), escolha lexical e gramatical, e grafia e pontuação apenas quando impedem a compreensão.

4. Aspectos metodológicos

Nessa seção, trazemos alguns aspectos da metodologia utilizada para a coleta de dados realizada neste trabalho. Voltamos a mencionar que os dados são o recorte de uma coleta maior, parte de uma pesquisa de mestrado.

Primeiramente, os sujeitos que participaram das aplicações eram, à época da coleta, estudantes do último ano de Letras – Português/Inglês em uma universidade pública no sudeste do Brasil. No total dessa amostra, 19 estudantes realizaram o pré-teste (em seus celulares), dos quais 11 realizaram também o EPPLÉ num intervalo de dois a dez dias após o pré-teste.

Além dos resultados das próprias avaliações, também realizamos coleta de dados através de relatórios de observação durante as aplicações e de questionários aplicados aos participantes logo após a realização do pré-teste.

5. Desafios tecnológicos

A disponibilização do pré-teste em computadores e dispositivos tecnológicos móveis (DTMs) se alinha com uma tendência de trabalhos sobre usos de tecnologia na Educação, inclusive na área de avaliação (VIBERG, 2015; MARTINS; SILVA; SILVA, 2017). Podemos dizer que o uso de tecnologias afetou o ensino de línguas basicamente com a mesma força com que afetou a avaliação neste contexto (CHAPELLE; VOSS, 2016). Mais especificamente, a acessibilidade do pré-teste via DTMs, e não apenas computadores, vislumbra um cenário em que *smartphones* se tornaram bastante populares no Brasil (MANS, 2016), sendo que o maior número de acessos à internet já se dá por meio deles (GOMES, 2016).

Marques-Schäfer e Rozenfeld (2018) organizaram um livro em torno da utilização de DTMs na educação, com foco especial na aprendizagem de línguas. A obra, *Ensino de línguas e tecnologias móveis: políticas públicas, conceitos, pesquisas e práticas em foco* (2018), é mais um indicativo da importância que os DTMs já têm na vida cotidiana e, crescentemente, na educação. A legislação brasileira ainda tenta barrar a integração de DTMs a práticas educativas, mas as pesquisas no Brasil e no exterior apontam para o caminho contrário, a associação de técnicas e atividades da vida exterior dos estudantes à vida escolar e à aprendizagem.

Conforme Pathan (2012), a tecnologia possibilita várias vantagens para a aplicação de testes, como a superação de obstáculos logísticos associados com transporte de testes e de resultados, a não necessidade de os candidatos estarem todos no mesmo horário e local para realizarem a prova, uma atitude mais positiva em relação ao teste,

dentre outros. Em nosso caso, com a disponibilização do pré-teste em DTMs, poderemos beneficiar professores com tais facilidades antes de se comprometerem com a inscrição para o exame EPPL, que, atualmente, requer a presença física do candidato em horário e local determinados.

O texto digitado na internet, especialmente através do celular, carrega algumas peculiaridades. Primeiramente, no caso do celular, o próprio dispositivo apresenta aspectos que o diferenciam da escrita em papel e no computador, como a velocidade de digitação, que pode variar com maior familiaridade do usuário com o aparelho e também de acordo com o teclado disponível (teclado numérico, QWERTY virtual ou analógico, dentre outros), as autocorreções habilitadas em vários modelos atuais de celulares, a interação com vários aplicativos ao mesmo tempo, sendo que o usuário pode receber notificações enquanto estiver digitando, dentre outros. Além disso, em diversas situações de escrita através do celular e do computador, podemos observar maior uso de abreviações e menor formalidade. Alguns autores chegam a defender que, no *continuum* entre fala e escrita, a comunicação que se faz através da internet se aproxima mais da fala do que da escrita (CRYSTAL, 2004; HALLIDAY, 1996 apud MARCUSCHI, 2010), uma vez que ambas têm se tornado cada vez mais próximas devido às novas possibilidades e aos novos discursos que surgiram com os avanços tecnológicos. Obviamente, em nosso caso, o contexto de avaliação requer um reajuste dos índices de formalidade e do próprio gênero, mas ainda são utilizados os mesmos canais de comunicação. No momento da correção, esses aspectos foram levados em conta ao se considerar o tamanho do texto, sua organização e algumas características consideradas desvios na escrita, especialmente relativas à ortografia.

5.1 Observações de campo

Durante a aplicação do pré-teste, a conexão com a internet fornecida pela universidade era instável, com variações de sinal dentro do ambiente de sala de aula. Alguns participantes optaram por utilizar o próprio programa de dados, porém este também sofria flutuações. A conexão falha resultava, algumas vezes, na atualização da página do pré-teste com perda do progresso do candidato até então, o que levou uma candidata a desistir de usar o próprio celular e esperar pelo de um colega, com menos flutuações de conexão. Além disso, uma candidata não possuía *smartphone*, pegando emprestado o celular da pesquisadora, e outra candidata tinha um modelo um pouco mais antigo de *smartphone*, cujo sistema operacional não processava html5, o código utilizado pelo YouTube (onde estão hospedadas as questões do pré-teste), impedindo o acesso aos áudios. Assim, esta participante foi forçada a esperar que um colega terminasse o teste para pegar seu celular emprestado.

5.2 Impressões dos participantes

As dificuldades relatadas pelos participantes à própria pesquisadora, durante ou logo após a aplicação, foram duas: o problema da autocorreção do celular e a falta de um local para anotações quando escutavam o texto oral. Uma das alunas não teve como utilizar o próprio celular e, ao pegar emprestado o da colega que havia terminado, relatou que o celular desta não possuía dicionário bilíngue, então autocorrigia todas as palavras para uma semelhante em língua portuguesa, o que tornava o exercício de responder as questões bem mais laborioso. Outros alunos relataram o problema do autocorretor também nos questionários. O segundo apontamento foi feito por uma aluna que pediu uma folha para fazer anotações do texto oral. Este texto tem duração de dois minutos e trinta e três segundos, descontando-se o tempo da introdução da questão, e também das questões subsequentes, que estão em áudios separados. Os itens subsequentes dizem respeito a aspectos gerais de interpretação e uso. Ao elaborarmos a prova, acreditamos que não seria necessário fazer anotações, e esta foi a única participante que sentiu necessidade da folha de papel. Ainda assim, seria interessante disponibilizar essa folha opcional aos candidatos que preferirem fazer anotações.

Ao serem questionados sobre as vantagens de realizar o pré-teste em dispositivo móvel, os participantes citaram frequentemente pontos como a liberdade (possibilidade de controlar o áudio – volume, quantas vezes ouvir, possibilidade de fazer as questões na ordem preferida), a facilidade, a rapidez e a interatividade. Já com relação às desvantagens, com exceção de um aluno, que não observou desvantagens, todos os restantes mencionaram dificuldades tecnológicas (flutuações na conexão com a internet, incompatibilidade do dispositivo com algum tipo de arquivo ou função, interrupção da atividade do teste por chegada de mensagens ou notificações de outros aplicativos), o que mostra a importância de se pesquisar como uma avaliação pode ser aplicada com mais eficiência através de DTMs.

Em questão posterior, os participantes opinam sobre sua preferência em realizar o pré-teste por meio do celular ou de um computador. Um participante preferiu o celular pela praticidade de realizar o teste em qualquer local com rede de internet. Dois participantes alegaram indiferença, desde que haja disponibilidade de boa conexão com a internet. Quatro participantes disseram preferir fazer pelo computador, por motivos diversos, como tamanho do teclado e da tela do DTM, presença de autocorretor, digitação mais lenta do que a do computador, maior chance de o computador abrir os arquivos, possibilidade de habilitá-lo para o teste com antecedência, dentre outros. O participante restante disse preferir o computador se pudesse fazer a prova em casa, mas alegou preferir o celular se tivesse que se deslocar para realizá-la.

A parte final do questionário abre espaço para críticas e sugestões dos candidatos, que são, em sua maioria, em relação à tecnologia utilizada. O quadro 1, a seguir, mostra as informações coletadas e um comentário da pesquisadora.

Quadro 1. Sugestões e críticas ao pré-teste no celular

SUGESTÃO/CRÍTICA	COMENTÁRIO
Disponibilização de rede sem fio estável para a realização do teste	Trabalhamos para aplicar o pré-teste em ambientes em que haja essa disponibilidade. No entanto, dependemos da conexão das universidades, que podem vir a sofrer falhas.
Inclusão de questões gramaticais e de múltipla escolha	Já há questões gramaticais no teste, talvez o candidato tenha se referido a questões de formato mais estrutural, como as mencionadas de múltipla escolha. Gostaríamos que o teste fosse mais comunicativo e, para isso, buscamos coletar a produção dos alunos, e criar contextos e situações de uso autênticas. No entanto, não descartamos a possibilidade de inclusão de questões de múltipla escolha devido à sua provável maior adaptabilidade a DTMs.
Inclusão de itens de produção oral	O único motivo pelo qual não incluímos itens de produção oral foi a indisponibilidade de plataformas <i>on-line</i> que oferecessem esse serviço.
Inclusão de itens de leitura	Por ser um teste de proficiência oral, não acreditamos que este seria uma das tarefas mais próximas das situações de uso que gostaríamos de avaliar.
Inclusão de itens com textos menos formais	Os primeiros três itens do teste são de situações informais, pois se tratam de questões pessoais. Os itens posteriores têm caráter mais específico da profissão de professor de LI e, por esse motivo, apresentam linguagem mais profissional e formal.
Resultados expressos em níveis (iniciante, intermediário, avançado), além das notas	Acredito que o participante em questão inferiu que seria avaliado em notas, pois é a maneira mais tradicional, embora isso não tenha sido dito em momento algum, uma vez que os resultados do pré-teste são expressos em faixas de proficiência.

(continua)

Instrução para escrever as respostas em LI	Essa sugestão foi de grande valia, uma vez que deixamos passar essa informação. No decorrer do teste, que é todo em áudio e em LI, não mencionamos que as respostas devem ser escritas também nessa língua. Já realizamos a correção.
Disponibilização do teste <i>off-line</i> (menos distrações e menos problemas com flutuações na conexão)	Seria bastante complicado disponibilizar o teste <i>off-line</i> , pois precisaríamos de um programa, ou, como sugeriu uma aluna, poderíamos enviar os arquivos aos candidatos.
Utilização de áudios apenas, e não vídeos (demora para carregamento)	A plataforma do Google <i>Forms</i> não permite inserção de áudios apenas, e nem o YouTube permite que carreguemos áudios sem imagens que os acompanham.
Inserção de imagens explicativas e/ou frases escritas nos vídeos, como complemento ao áudio	O propósito do teste de avaliar a proficiência oral seria comprometido, uma vez o candidato poderia responder às questões através de sua proficiência escrita, ou da assistência visual das imagens.
Melhor padronização do volume do áudio	Gostaríamos muito de disponibilizar um áudio mais padronizado, como discutimos na seção de metodologia, mas, para isso, precisamos de uma plataforma de voz com boa entonação, ou a disponibilidade de aparelhos (microfones) de boa qualidade, em estúdio, para a gravação.

Fonte: Elaboração própria.

Alguns participantes mencionaram que problemas com a conexão não são necessariamente do teste em si, mas externos a ele. No entanto, ao se aplicar uma avaliação, é preciso levar em conta questões externas a ela, que são essenciais para a sua boa realização. Não importa se o problema de conexão é externo, ainda assim ele deve ser levado em conta no desenvolvimento e aplicação do teste para que os resultados sejam confiáveis e válidos. Um problema mais sério do que esse é a utilização do celular do próprio participante, que pode não ser compatível com o formato da avaliação, ou não processar os vídeos, ou estar com quaisquer outros problemas. Em nosso caso, não temos como ter controle sobre os aparelhos pessoais dos alunos, e isso foi um problema no momento da aplicação dos testes. Uma solução seria a aplicação por meio de *tablets* pré-configurados, o que, no entanto, prejudicaria o propósito de o participante não precisar se locomover para realizar o pré-teste.

Não podemos deixar de mencionar que, nessa mesma seção dos questionários, alguns candidatos teceram elogios, dizendo que gostaram das questões que levam em conta a atuação do professor e que acharam o áudio e as instruções claras.

5.3 Principais dificuldades e atitudes preventivas

A partir da experiência das aplicações e do *feedback* dos participantes, elencamos alguns pontos importantes para futuras aplicações e pesquisas sobre o pré-teste.

Reiteramos, primeiramente, alguns pontos que são essenciais na aplicação de qualquer teste eletrônico. Primeiramente, deve haver, sempre que possível, o teste dos materiais a serem utilizados. No caso dos celulares e fones dos próprios alunos, não tivemos acesso anterior a eles, o que gerou alguns problemas durante a aplicação. Além disso, e como consequência de problemas potenciais, é preciso sempre ter materiais extras: trouxemos muitos fones extras para as aplicações e disponibilizamos nossos aparelhos celulares pessoais para os participantes que tiveram problemas com os seus próprios. Preparando-se com antecedência é muito mais fácil e rápido lidar com os obstáculos que se apresentam no momento da administração do teste.

Uma participante indicou a falta da instrução aos candidatos de desabilitarem outros aplicativos em seus aparelhos celulares. Tal candidata ficou incomodada, e com toda razão, com o fato de que, durante o teste, recebia mensagens instantâneas em sua tela que provinham de outros aplicativos. Porém, instruir os candidatos a desabilitar aplicativos pode ser desafiador, já que eles podem não saber como realizá-lo. Para isso, precisaríamos de examinadores treinados para ajudar os participantes a configurar modelos diferentes de aparelho, muitas vezes com sistemas operacionais distintos, ou mesmo outras versões do mesmo sistema. Não seria uma instrução simples, como desconectar o aparelho da internet, uma vez que isso desabilitaria o candidato de enviar suas respostas do teste, que está hospedado *on-line*, além de não ser garantia de que aplicativos ainda assim enviassem notificações ou lembretes *off-line*, ou mesmo despertadores previamente estabelecidos pelos candidatos.

Tratando agora de quesitos internos ao pré-teste, um ponto que havíamos considerado e que não conseguimos instrumentalizar foi a presença de um cronômetro no teste. Esse aspecto foi levantado também por um participante e nos fez determinar que, no aviso inicial do teste, devemos informar, além da duração e de que vamos avisar quando restarem 5 minutos, que todos os candidatos possuem relógio em seu celular e qual será o horário de término do teste, para que eles mesmos possam acompanhá-lo para além dos avisos dos examinadores. Para que isso seja bem sucedido, é preciso, antes do teste, confirmar se os relógios dos celulares dos alunos estão acertados.

Outro ponto importante foi a limitação de número de palavras. Nas instruções do teste, pedimos que os participantes escrevessem pelo menos 10 palavras. Observando os resultados, percebemos que foi um número que gerou respostas muito próximas de 10 palavras e também muito distantes, chegando a mais de 70 palavras, por exemplo, o que dificultou a correção pela heterogeneidade das respostas disponíveis. Exploramos novas opções no *Forms* e acreditamos que colocar um limite mínimo de caracteres no próprio teste é uma boa alternativa, e esse limite impediria o candidato de submeter uma resposta menor do que a requisitada. Podemos aplicar em cada questão um diferente número mínimo de palavras, sendo, por exemplo, de 200 caracteres em uma, e 400 em outra da qual se espera uma resposta mais longa. Gostaríamos de poder adicionar também um limite máximo, mas o *Forms* só permite incluir uma regra de número de caracteres, e acreditamos que a de número mínimo seria a mais útil, visto que a própria duração do teste impede o participante de escrever respostas muito longas.

6. Validade

Validade é um conceito central na área de avaliação, pois trata-se, *grosso modo*, da medida em que um teste avalia aquilo que se propôs a avaliar. Ademais, Henning (1987 apud ALDERSON; CLAPHAM; WALL, 1995) afirma que um teste é válido para um determinado propósito, o que significa que um teste deve ser julgado, com base em evidências, como mais ou menos válido, de acordo com sua finalidade. Assim, falamos de validade não como característica intrínseca ao teste, mas sim ao seu uso para tomar certas decisões, baseado em que construto decidiu-se que seria adequado medir, e no quanto o teste realmente avalia esse construto.

Portanto, tradicionalmente, validade refere-se “à adequação de um teste ou de algum de seus componentes como uma medida do que esse teste deve medir” (HENNING, 1987, p. 89 apud ALDERSON; CLAPHAM; WALL, 1995, p. 170). Esta visão tradicional passou por reelaborações através dos anos, e o conceito deixou de ser uma característica do próprio teste, para vir a se tornar um argumento relativo à sua interpretação e uso, e como estes podem ser justificados.

O percurso histórico do conceito de validade, de acordo com Kane (2001), parte de uma visão tipológica para uma unificada. Inicialmente, os tipos de validade (critério, conteúdo e construto) eram vistos como uma “*toolbox*” (caixa de ferramentas), na medida em que os pesquisadores lançavam mão dos vários tipos de validade de acordo com seus trabalhos e bom senso. Em tal abordagem, havia potencial para escolhas altamente oportunistas, ou, no mínimo, dúbias, de validades que, ao estudioso, se mostravam mais adequadas, simples ou promissoras para resultados positivos em seus trabalhos de validação.

Na visão atual, unificada, os tipos de validade passaram a funcionar como possíveis evidências para que se valide um teste para certo propósito. O processo de validação

passou a ser mais longo e complexo, e desenvolvimentos recentes no conceito trouxeram discussões sobre a inclusão ou não de consequências como parte da validade de um teste, além de exporem mais minuciosamente possíveis trajetos para o processo de validação (BACHMAN, 2005; XI, 2010).

Em nosso caso, não pretendemos realizar um processo completo de validação do pré-teste, e sim analisar somente sua validade de critério em relação ao EPPLÉ. A validade de critério diz respeito à utilização de uma medida como referência contra a qual se compara o teste que se pretende validar. Dessa maneira, o EPPLÉ é nossa medida de referência e analisamos a validade do pré-teste em relação a ele. Nesse artigo, especificamente, tratamos de um subtipo de validade de critério, a validade paralela, que acontece quando os resultados da medida de referência e do teste a ser validado são coletados simultaneamente (em nosso caso, com intervalo de poucos dias).

6.1 Validade paralela

Para a análise da validade paralela dos dados coletados, comparamos as faixas de proficiência alcançadas pelos participantes nas duas avaliações realizadas, o pré-teste e o EPPLÉ oral. O pré-teste, por ser um teste curto e simples, concede ao participante faixas que vão de A1 a B2, acompanhando as faixas do quadro comum europeu, mas com especificações relativas à proficiência linguística de professores de línguas. Assim, a faixa mais alta que o pré-teste certifica é a B2, uma vez que as tarefas apresentadas aos candidatos não avaliam aspectos distintivos entre candidatos B2 e C1 ou C2. Já para o EPPLÉ oral, haviam sido desenvolvidas faixas que vão de A a E, sendo A a faixa mais alta e E a mais baixa. Outra singularidade do EPPLÉ é que a faixa E abarca todos os resultados abaixo de 60%, considerados insuficientes para o bom desempenho da profissão de professor de línguas. As faixas A, B, C e D contabilizam, portanto, intervalos de apenas 10%.

Organizamos os quadros 2 e 3 com os resultados das avaliações, partindo de diferentes referências. No quadro 2, a seguir, agrupamos os participantes a partir de seus resultados no pré-teste, relacionando-os, então, com seus resultados no EPPLÉ oral. Dentro da coluna do EPPLÉ, a primeira subcoluna representa o número de alunos daquela faixa (dentro o total dos presentes na faixa de referência do pré-teste) e a segunda subcoluna contém a porcentagem que esse número representa dentro também do total de alunos que estavam na faixa de referência. Para as subdivisões de cada faixa do pré-teste, o resultado no EPPLÉ oral com a maioria dos participantes está em negrito e sublinhado. Nenhum participante teve a faixa A1 como resultado, por isso ela não consta no quadro.

Assim, da nossa amostra total de 11 participantes que realizaram o pré-teste e o EPPLÉ, percebemos que 4 ficaram na faixa B2 do pré-teste, 5 na faixa B1 e 2 na faixa A2. Tomando por base, por exemplo, a faixa B2, dos 4 participantes nessa faixa, 2 conseguiram a faixa

C do EPPLE e os outros 2 ficaram nas faixas A e C. Já partindo dos participantes B1 no pré-teste, de um total de 5, 3 ficaram na faixa B e 2 na C. Os dois participantes que ficaram na faixa A2 do pré-teste foram posicionados na faixa C do EPPLE oral.

É possível, a partir dessa correlação, visualizar uma gradação nos resultados. As faixas A2 e B1 estão hierarquizadas com clareza, sendo a faixa B2 a única que apresenta informação parcialmente contraditória, pois 50% dos participantes foram classificados na faixa C (a mesma que foi a totalidade da faixa A2), com o restante dos participantes divididos igualmente entre duas faixas acima (A) e uma abaixo (D). A explicação para isso pode advir do fato de que o pré-teste não avalia a produção oral, mas apenas a compreensão.

Quadro 2. Resultados – Validade paralela – Pré-teste como referência

REFERÊNCIA - PRÉ-TESTE	EPPLE ORAL	
B2	A	
	1	25%
	B	
	0	0%
	C	
	2	50%
	D	
	1	25%
4 PARTICIPANTES	E	
	0	0%

(continua)

B1	A	
	0	0%
	B	
	3	60%
	C	
	2	40%
	D	
	0	0%
5 PARTICIPANTES	E	
	0	0%
A2	A	
	0	0%
	B	
	0	0%
	C	
	2	100%
	D	
	0	0%
2 PARTICIPANTES	E	
	0	0%

Fonte: Elaboração própria.

Diante dessa inesperada relação, estudamos também os mesmos resultados, agora tomando como referência o EPPLE oral, assim como consta no quadro 3, organizado de maneira semelhante ao quadro anterior.

Todos os candidatos que ficaram na faixa B do EPPLE oral estão nas faixas B2 e B1 do pré-teste, sendo, respectivamente, 40% e 60% do EPPLE oral. Portanto, o pré-teste, da maneira que se encontra, garante aos alunos das faixas B1 e B2 grandes chances de conseguirem a faixa B no EPPLE oral.

Nossa amostra de resultados das faixas A do pré-teste ficou bastante diminuída, dado o público de nossa aplicação, que possui, em sua maioria, alta proficiência. Apenas dois candidatos (que realizaram pré-teste e EPPLE) se posicionaram na faixa A2, e ambos tiveram faixa C no EPPLE oral. Tomando como referência a faixa C do EPPLE oral, podemos ver que, na verdade, 60% deles ficaram na faixa B1 do pré-teste, 20% na faixa B2 e 20% na A2. Assim, embora a faixa A2 do pré-teste aparente indicar um resultado C no EPPLE oral, candidatos com notas B2 e, principalmente, B1 também conseguiram a mesma faixa C. Observando as relações feitas nos parágrafos anteriores, concluímos que a faixa B1 do pré-teste está abarcando, no momento, participantes de níveis variados no EPPLE oral, o que pode sugerir, mais uma vez, que é problemático não avaliar também a produção oral no pré-teste.

Quadro 3. Resultados – Validade Paralela – EPPLE₁ oral como referência.

REFERÊNCIA – EPPLE ORAL	PRÉ-TESTE	
B	B2	
	2	40%
	B1	
	3	60%
	A2	
	0	0%
5 PARTICIPANTES	A1	
	0	0%
C	B2	
	2	33%
	B1	
	2	33%
	A2	
	2	33%
6 PARTICIPANTES	A1	
	0	0%

Fonte: Elaboração própria.

Como podemos observar, todos os alunos que fizeram o EPPL oral acabaram por se posicionar entre as faixas B e C. Este não é um resultado inesperado, pelo contrário, dado que lidamos com um público de alta proficiência. Com relação ao quadro, vemos que a maioria dos candidatos, que tirou B ou C no EPPL oral, obteve a faixa B1 ao fazer o pré-teste. No entanto, nenhum candidato que alcançou B no EPPL oral tirou menos do que B1 no pré-teste, enquanto os candidatos que obtiveram C no EPPL oral tiveram faixas variadas no pré-teste (B2, B1 e A2).

7. Considerações finais e encaminhamentos

Com base nos resultados parciais encontrados durante esta pesquisa, podemos propor alguns encaminhamentos e sugestões para pesquisas futuras. Primeiramente, pesquisas que se propusessem a realmente validar o pré-teste, passando por todas as etapas do processo de validação e coletando evidências que as apoiassem, consolidariam o pré-teste como instrumento válido para indicar ao candidato sua aptidão para prestar o EPPL.

Em segundo lugar, embora o pré-teste já tenha sido reformulado para uma segunda versão e esteja em processo de nova adaptação, acreditamos que ainda há arestas a polir e que outras versões do pré-teste seriam úteis, inclusive como versões paralelas para que não haja apenas uma disponível ao público. Seria interessante que essas versões incluíssem a produção oral, que parece estar gerando disparidades nos resultados.

Por fim, pesquisas com tecnologia e avaliação, em especial com a utilização de DTMs ainda são escassas, e percebemos que há vários desafios para os quais não há soluções simples. Essa área de pesquisa, inclusive, se beneficiaria enormemente de interação interdisciplinar com áreas como a programação e a ciência da computação, em geral.

REFERÊNCIAS

ANCHIETA, P. P. *Relações entre o construto e as características do teste escrito do EPPL eletrônico: um processo de validação*. 2015. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2015.

AGUENA, D. M. *Avaliação da proficiência linguístico-comunicativa-pedagógica do professor de línguas: operacionalização de construto no Exame de Proficiência para Professores de Línguas Estrangeiras – ETAPA II*. 2016. Relatório de Iniciação Científica (Licenciatura em Letras) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2016.

ALDERSON, J. C.; CLAPHAM, C.; WALL, D. Validation. In: ALDERSON, J. C.; CLAPHAM, C.; WALL, D. *Language Test Construction and Evaluation*. New York: Cambridge University Press, 1995. p. 170-196.

ALMEIDA FILHO, J. C. P. O professor de Língua Estrangeira sabe a língua que ensina? A questão da instrumentalização lingüística. *Contexturas*, São Paulo: APLIESP, v. 1, n. 1, p. 77-85, 1992.

BACHMAN, L. F. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1990.

BACHMAN, L. F. Building and Supporting a Case for Test Use. *Language Assessment Quarterly*, v. 2, n. 1, p. 1-34, 2005.

CELANI, M. A. A. Perguntas ainda sem resposta na formação de professores de línguas. In: GIMENEZ, T.; MONTEIRO, M. C. G. (org.). *Formação de professores de línguas na América Latina e transformação social*. Campinas: Pontes Editores, 2010.

CELANI, M. A. A. Ensino de línguas estrangeiras: ocupação ou profissão. In: LEFFA, V. J. (org.). *O professor de línguas estrangeiras: construindo a profissão*. Pelotas: EDUCAT, 2008.

CHAPELLE, C. A.; VOSS, E. 20 years of technology and language assessment in Language Learning & Technology. *Language Learning & Technology*, v. 20, n. 2, p. 116-128, 2016. Disponível em: <http://llt.msu.edu/issues/june2016/chapellevoss.pdf>. Acesso em: 17 maio 2017.

CONSOLO, D. A. Assessing EFL teachers' oral proficiency: On the development of teacher education programmes and testing policies in Brazil. In: CONSOLO, D. A.; GATTOLIN, S. R. B.; TEIXEIRA DA SILVA, V. L. (org.). *Perspectivas em Avaliação no Ensino e na Aprendizagem de Línguas: pesquisas e encaminhamentos na formação e na prática docente*. Campinas: Pontes Editores, 2017. p. 59-71.

CONSOLO, D. A. Classroom oral interaction in foreign language lessons and implications for teacher development. *Linguagem & Ensino*, Pelotas, v. 9, n. 2, p. 33-55, 2006.

CONSOLO, D. A.; AGUENA, D. M. Pré-testagem ao Exame de Proficiência para Professores de Língua Estrangeira (EPPL): desenho e desafios tecnológicos para avaliação de proficiência oral em dispositivos tecnológicos móveis. *Revista Polifonia*, Cuiabá, v. 24, n. 35/2, p. 30-49, 2017.

CONSOLO, D. A.; ALVARENGA, M. B.; CONCÁRIO, M.; LANZONI, H. P.; MARTINS, T. H. B.; SILVA, V. L. T. An examination of foreign language proficiency for teachers (EPPL): The initial proposal and implications for the Brazilian context. *In: II Congresso Internacional da ABRAPUI*, 2009, São José do Rio Preto/SP. *The Teaching of English: Towards an Interdisciplinary Approach Between Language and Literature*. São José do Rio Preto/SP: ABRAPUI, 2009. p. 1-15.

CONSOLO, D. A.; BERTOLDO, E. S.; COX, M. I. P.; PETERSON, A. A. A.; VIANA, N.; FÉLIX, A. Desafios e perspectivas na formação de professores de línguas no Brasil. *Estudos Linguísticos*, São Paulo, v. 28, n. 1, p. 112-117, 1999.

CONSOLO, D. A.; SILVA, V. L. T. Em defesa de uma formação linguística de qualidade para professores de línguas estrangeiras: o exame EPPL. *Revista Horizontes de Linguística Aplicada*, v. 13, p. 63-87, 2014.

CRYSTAL, D. *Language and the internet*. Cambridge: Cambridge University Press, 2004.

GOMES, H. S. Smartphone passa PC e vira aparelho nº 1 para acessar internet no Brasil. *Portal G1*. abr. 2016. Disponível em: <https://glo.bo/2G8YGGT>. Acesso em: 17 maio 2017.


GREEN, A. Washback in language assessment. *International Journal of English Studies*, v. 13, n. 2, p. 39-51, 2013.

KANE, M. T. Current Concerns in Validity Theory. *Journal of Educational Measurement*, v. 38, n. 4, p. 319-342, 2001.

MANS, M. Uso de smartphones já alcança 80% dos brasileiros, aponta estudo. *Estado de São Paulo*. out. 2016. Disponível em: <http://bit.ly/2NOb0Tv>. Acesso em: 17 maio 2017.

MARCUSCHI, L. A. Gêneros textuais emergentes no contexto da tecnologia digital. *In: MARCUSCHI, L. A.; XAVIER, A. C. (org.). Hipertexto e gêneros digitais: novas formas de construção de sentido*. 3. ed. São Paulo: Cortez, 2010. p. 15-80.

MARQUES-SCHÄFER, G.; ROZENFELD, C. C. F. (org.). *Ensino de línguas e tecnologias móveis: políticas públicas, conceitos, pesquisas e práticas em foco*. São Paulo: Edições Hipótese, 2018.



MARTINS, T. H. B.; SILVA, V. L. T.; SILVA, W. E. Desafios para se aplicar um exame de proficiência em meio eletrônico. *In: CONSULO, D. A.; GATTOLIN, S. R. B.; TEIXEIRA DA SILVA, V. L. (org.). Perspectivas em Avaliação no Ensino e na Aprendizagem de Línguas: pesquisas e encaminhamentos na formação e na prática docente.* Campinas: Pontes Editores, 2017. p. 191-206.

PATHAN, M. M. Computer Assisted Language Testing (CALT): Advantages, Implications and Limitations. *Research Vistas*, v. 1, n. 4, p. 30-45, 2012.

VIBERG, O. *Design and Use of Mobile Technology in Distance Language Education: Matching Learning Practices with Technologies-in-Practice.* Örebro: Örebro University, 2015.

XI, X. Methods of test validation. *In: SHOHAMY, E.; HORNBERGER, N. H. (ed.). Encyclopedia of Language Education.* 2 ed. v. 7. Language Testing and Assessment. New York: Springer Science+Business Media LLC, 2010. p. 177-196.