

**EXPLORANDO OS INDÍCIOS DE DESENVOLVIMENTO DE
LETRAMENTO EM AVALIAÇÃO NA FORMAÇÃO INICIAL DE
PROFESSORES DE LÍNGUAS: CONFIABILIDADE E CONCEPÇÕES DE
AVALIAÇÃO**

*Exploring Evidence of Assessment Literacy Development in Preservice Language
Teacher Education: Reliability and Conceptions of Assessment*

Mariana DAMACENA-DUTRA
Universidade de Brasília
mariana.damacena@unb.br
ORCID: 0000-0002-5909-3745

Vanessa BORGES-ALMEIDA
Universidade de Brasília
borgesalmeida@unb.br
ORCID: 000-0003-0502-5162

RESUMO: Este artigo apresenta e discute os resultados de um estudo de caso interventivo realizado em uma universidade pública brasileira. O objetivo do trabalho é discutir indícios de desenvolvimento em letramento em avaliação de futuros professores durante e após a conclusão de uma disciplina teórico-prática dedicada à temática da avaliação em contexto de línguas. A pesquisa está fundamentada nos estudos sobre letramento em avaliação em contexto de línguas (Inbar-Lourie, 2008; Quevedo-Camargo; Scaramucci, 2018) na formação inicial de professores (Fernandes; Borges-Almeida, 2021; Giraldo, 2018; Giraldo; Murcia, 2018, 2019; Restrepo Bolivar, 2020) e sobre avaliação como mecanismo orientador da aprendizagem (Black, 2015; Black; Wiliam, 2010; 2018; Carless, 2007). Participaram da investigação sete estudantes de Letras Inglês. Os dados foram coletados por meio de observação de aulas, entrevistas e análise documental de tarefas elaboradas pelos participantes, e foram analisados pelo método de análise de conteúdo (Bardin, 2011). Os resultados indicam que há ganhos no letramento dos participantes quanto à concepção de avaliação para apoiar e orientar a aprendizagem, e à compreensão das qualidades de boas práticas avaliativas, especialmente confiabilidade, apoiada por rubricas de correção.

PALAVRAS-CHAVE: letramento em avaliação; professores de línguas; formação inicial; confiabilidade; rubricas

ABSTRACT: This article presents and discusses the results of an intervention case study conducted at a Brazilian public university. The aim of this work is to discuss evidence of preservice teachers' assessment literacy development during and after the completion of a theoretical-practical credit course dedicated to the theme of assessment in language

contexts. This research is based on studies on language assessment literacy (Inbar-Lourie, 2008; Quevedo-Camargo; Scaramucci, 2018) in preservice teacher education (Fernandes; Borges-Almeida, 2021; Giraldo, 2018; Giraldo; Murcia, 2018, 2019; Restrepo Bolivar, 2020) and on assessment as a mechanism to guide learning (Black, 2015; Black; Wiliam, 2010; 2018; Carless, 2007). Seven students of an English Language and Literature degree participated in this investigation. Data were collected through classroom observation, interviews, and document analysis of tasks designed by the participants, and were analyzed using the content analysis method (Bardin, 2011). Results indicate gains in participants' literacy regarding the conception of assessment to support and guide learning, and the understanding of the qualities of good assessment practices, especially reliability, supported by rubrics.

KEYWORDS: language assessment literacy; assessing speaking; preservice teacher education;

INTRODUÇÃO

Tradicionalmente, os cursos de graduação devotam pouca atenção à teoria e prática avaliativa, marginalizando as discussões e reflexões a respeito de avaliação (Quevedo-Camargo, 2020; Scaramucci, 2016; Taylor, 2009). Isso se reflete diretamente na atuação dos professores em sala de aula, onde o ato avaliativo reduz-se à aplicação de provas descontextualizadas para verificar o conhecimento ou desempenho dos alunos em uma etapa já finalizada, sem a mínima observância a princípios de boas práticas avaliativas.

É papel dos cursos de licenciatura em Letras proporcionar que o professor de línguas em formação inicial saiba por que, o que, e como avaliar seus alunos (Inbar-Lourie, 2008). Tal formação refere-se ao conceito de letramento em avaliação, essencial para que o professor se torne “consumidor e produtor crítico de avaliações [...] assim como [saiba] interpretar os resultados obtidos por meio de suas avaliações” (Quevedo-Camargo; Scaramucci, 2018, p. 239).

Giraldo e Murcia (2018) defendem que estudar sobre avaliação na formação inicial de professores contribui para o desenvolvimento profissional como um todo, uma vez que ajuda a construir uma consciência geral sobre o que significa avaliar, quais são as implicações dos processos avaliativos e seu o impacto geral na aprendizagem e ensino de línguas. Embora observemos recentemente uma tendência crescente de estudos que investigam a promoção do letramento em avaliação na formação inicial, pesquisas nesse contexto ainda são poucas quando comparadas a outros contextos. O presente estudo contribui para preencher essa lacuna, ao explorar as contribuições advindas de uma disciplina inserida no currículo de um curso de Licenciatura em Letras Inglês dedicada ao estudo da temática de avaliação em contexto de ensino de línguas. A análise aqui apresentada é recorte de uma investigação para obtenção de titulação de mestrado, e orienta-se pela seguinte pergunta de pesquisa: que indícios há de desenvolvimento de letramento em avaliação dos professores em formação inicial durante e após cursarem a disciplina?.

Este artigo está estruturado em três partes principais. Na fundamentação, discutimos o conceito de letramento em avaliação em contexto de línguas e como ele tem sido promovido e investigado na formação inicial de professores. Em seguida,

descrevemos a metodologia deste estudo. Por fim, analisamos e discutimos os resultados.

REFERENCIAL TEÓRICO

Letramento em avaliação em contexto de línguas

Stiggins (1991) é comumente citado como um dos primeiros proponentes do conceito de letramento em avaliação para referir-se à busca por práticas avaliativas que comuniquem definições claras, específicas e ricas do desempenho dos aprendizes. No contexto de ensino de línguas, a definição de letramento em avaliação foi ampliada para compreender tanto a dimensão avaliativa, quanto a da linguagem (Inbar-Lourie, 2008), dando origem ao conceito de letramento em avaliação em contexto de línguas (LAL).

Para Inbar-Lourie (2013, p. 1), o “letramento em avaliação em contexto de línguas forma a base de conhecimento necessário para conduzir procedimentos de avaliação de línguas, ou seja, para elaborar, administrar, interpretar, utilizar e relatar dados da avaliação de línguas para diferentes propósitos”¹, o que, segundo a autora, demanda não somente conhecer as ferramentas e procedimentos para avaliar, mas também a habilidade de fornecer *feedback* de modo a direcionar as ações dos aprendizes. Além disso, a autora destaca que o construto de letramento em avaliação está diretamente relacionado à “capacidade de fazer e responder a perguntas cruciais sobre o propósito da avaliação, sobre a adequação da ferramenta utilizada, sobre as condições de aplicação de provas e sobre o que vai acontecer com base nos resultados”² (Inbar-Lourie, 2008, p. 389)³.

Davies (2008) propõe três elementos centrais que constituem o LAL: conhecimentos, princípios e habilidades. O autor define que o componente *conhecimentos* abrange tanto os conhecimentos teóricos relacionados à Linguística Aplicada quanto aqueles específicos da avaliação. Os *princípios* referem-se ao “uso adequado de provas

¹ Tradução de “Language assessment literacy forms the knowledge base needed to conduct language assessment procedures, that is, to design, administer, interpret, utilize, and report language assessment data for different purposes” (Inbar-Lourie, 2013, p. 1). Todas as traduções trazidas neste artigo são livres e das próprias autoras.

² Optamos por utilizar a definição proposta por Inbar-Lourie (2013), uma vez que o paradigma socioconstrutivista está alinhado à investigação conduzida neste trabalho, que entende a avaliação como um fenômeno socialmente construído (Inbar-Lourie, 2008).

³ Tradução de: “the capacity to ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results” (Inbar-Lourie, 2008, p. 389)

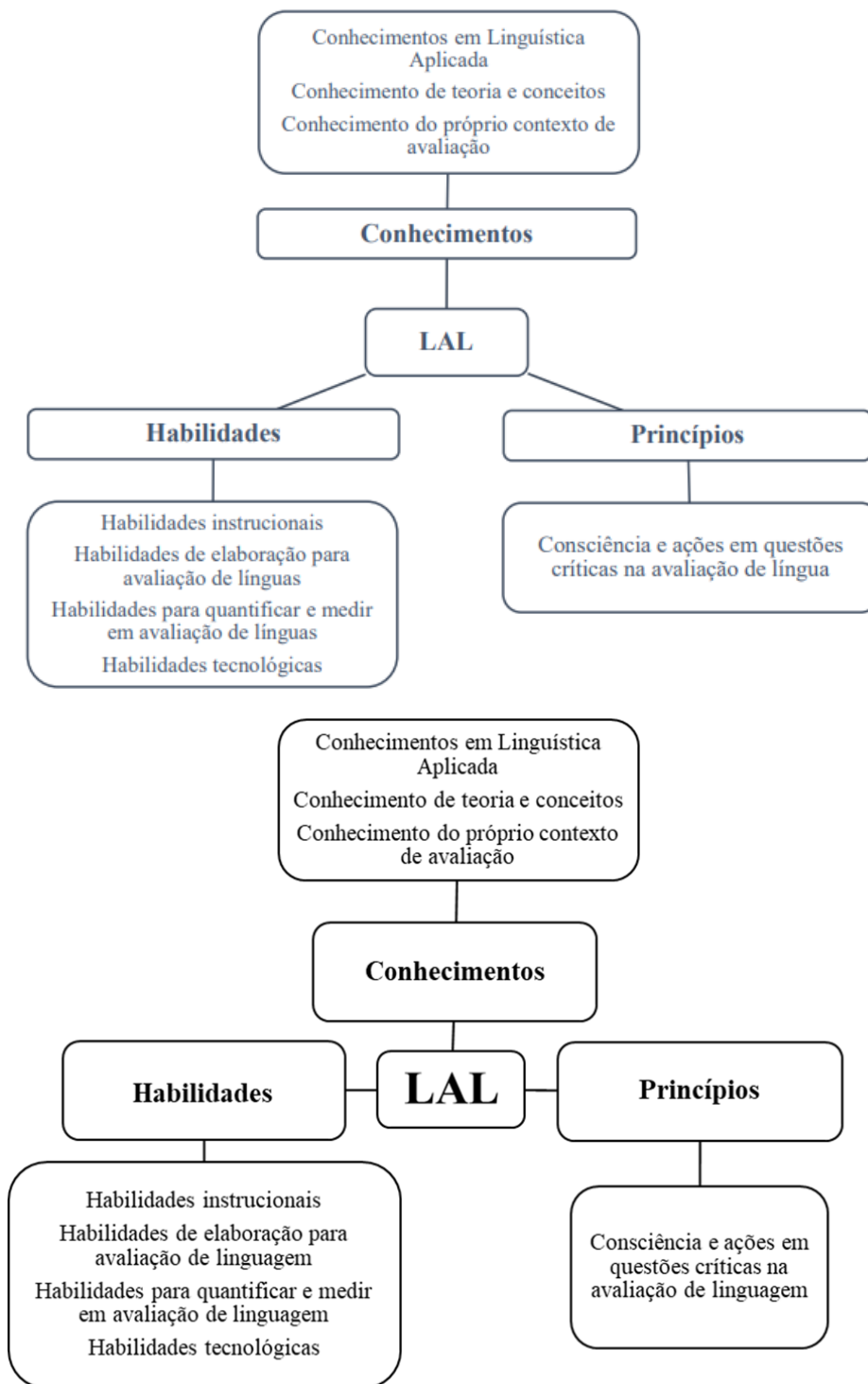
de línguas, imparcialidade e impacto, incluindo questões de ética e profissionalismo” (Davies, 2008, p. 335). Por fim, as *habilidades* dizem respeito à operacionalização dos construtos, incluindo elaboração de itens, análise de resultados e *feedback*. Ademais, o autor salienta a indissociabilidade entre cada um desses componentes, ao destacar que

[...] as habilidades não são sustentáveis sem conhecimento, uma vez que o conhecimento fornece o contexto no qual as habilidades operam: se as habilidades representam ‘como?’, então o conhecimento representa ‘o quê?’ [...] E ainda, à medida que o conhecimento se torna mais amplamente disponível na profissão, a necessidade de explicar, justificar e julgar torna-se importante. [...] Por sua vez, o novo conhecimento baseado em princípios é operacionalizado e incorporado às habilidades. Habilidades, ou seja, técnicas e metodologias, por si só, não são mais suficientes, conhecimento + habilidades são inadequadas sem a adição de princípios. [...] Todos são necessários, mas um sem o(s) outro(s) provavelmente será(ão) mal-entendido(s) e/ou banalizado(s)” (Davies, 2008, p. 335-336).⁴

De lá para cá, diversos autores têm descrito em mais detalhes os componentes do letramento em avaliação (p. ex., Giraldo, 2018; Kremmel; Harding, 2020; Vogt; Tsagari, 2014). Giraldo (2018) parte do modelo de Davies (2008) para propor uma mais detalhada descrição dos três componentes, conhecimentos, habilidades e princípios. Assim, o autor divide cada componente em categorias que, por sua vez, agrupam diferentes descritores do letramento(figura 1).

⁴ Tradução de: “skills are not sustain-able without knowledge since knowledge provides the context in which skills operate: if skills represent ‘how?’, then knowledge rep-resents ‘what?’. [...] And further, as knowledge becomes more widely available in the profession, so the need to explain, to justify and to judge becomes important. [...] In its turn, the new principles-informed knowledge is operationalized and incorporated into skills. Skills, meaning techniques and methodologies, on their own are no longer enough, skills knowledge are inadequate without the addition of principles. [...] All are necessary but one without the other(s) is likely to be misunderstood and/or trivialized” (Davies, 2008, p. 335-336).

Figura 1 – Componentes, categorias e descritores do LAL



Fonte: Giraldo (2018, p. 187, tradução nossa).

Giraldo (2018) argumenta que os descritores podem ser tomados como ponto de partida para uma observação de elementos cruciais do LAL, podendo ser utilizados em ferramentas de autoavaliação ou avaliação de terceiros, possibilitando a identificação de potencialidades e lacunas no LAL. A lista proposta pelo autor também pode ser utilizada como um panorama sobre o assunto, o que pode ser particularmente útil para identificar temas a serem discutidos e aprofundados. No quadro 1, destacamos os descritores que têm relação mais direta com as categorias que discutiremos neste artigo, a saber, o entendimento de avaliação como mecanismo para orientar a aprendizagem (Black, 2015; Black; Wiliam, 2010; 2018; Carless, 2007).

Quadro 1: Descritores do letramento em avaliação

Categoria	Descritores
Conhecimentos	
Teoria e Conceitos da avaliação	Interpreta a validade na avaliação e suas implicações: validade de construto, conteúdo e critério, validade de construto como unificada, validade consequential de Messick (1989); validade como argumento.
	Interpreta as principais qualidades das práticas de avaliação linguística (além de confiabilidade e validade) e suas implicações para a avaliação linguística: autenticidade, praticidade, interatividade, equidade, ética e impacto (incluindo retroalimentação).
	Reconhece o que o feedback implica dentro de um paradigma de avaliação formativa.
Habilidades	
Instrucionais	Fornecer feedback sobre o desempenho dos alunos na avaliação (referenciada pela norma e por critério).
	Proporcionar experiências de avaliação motivadoras, fornecendo feedback encorajador ou criando cenários de autoavaliação.
De elaboração	Projetar avaliações que sejam válidas não apenas em termos de conteúdo do curso, mas também em termos de tarefas do curso.
	Elaborar avaliações que sejam confiáveis, autênticas, justas, éticas, práticas e interativas.
	Elaborar itens abertos (para fala e escrita), juntamente com rubricas de avaliação.
	Elaborar rubricas para avaliações alternativas, como portfólios e avaliação por pares.
Princípios	

Questões críticas da avaliação	Utiliza os resultados das avaliações como feedback para influenciar a aprendizagem da linguagem, não outras fontes irrelevantes (por exemplo, viés pessoal em relação a um aluno).
	Trata todos os alunos, ou usuários da avaliação de linguagem, com respeito.
	Fornecer práticas de avaliação que são justas e não discriminatórias.
	Implementa práticas transparentes de avaliação de linguagem; informa os alunos sobre o quê, como e por que da avaliação.
	Implementa práticas democráticas de avaliação de linguagem, oferecendo aos alunos oportunidades para compartilhar suas opiniões sobre a avaliação.

Fonte: Adaptado de Giraldo (2018, tradução nossa)

A indissociabilidade entre as dimensões pode ser observada, no quadro 1, na recorrência de uma mesma temática em diferentes dimensões. Por exemplo, é preciso compreender o conceito de confiabilidade (descriptor 9) e ser capaz de elaborar rubricas para apoiar a correção baseada em critério (44, 45) para implementar práticas avaliativas transparentes (65). É preciso compreender o papel central do *feedback* (19) para fornecê-lo (27), com base em informações relevantes (58), advindas, por exemplo, das rubricas avaliativas.

Fulcher (2012) organiza hierarquicamente as dimensões do letramento — conhecimentos, habilidades e princípios. Na base, estão os contextos (os aspectos históricos, sociais, políticos e filosóficos da avaliação); apoiada nessa base, há a dimensão dos princípios (os processos, princípios e conceitos que servem como guia para a prática avaliativa), diretamente ligada às questões de boas práticas em avaliação (especialmente validade, confiabilidade e praticidade, que serão detalhadas a seguir); e por fim, na superfície, encontram-se as práticas (o conhecimento e as habilidades necessários para a implementação da avaliação).

Scaramucci (2011) destaca a importância do princípio de validade, um conceito que engloba múltiplas facetas, no qual o construto ocupa posição central (Messick, 1996). Ebel e Frisbie (1991 apud Scaramucci, 2011) explicam que a definição de construto está ligada a conceitualizações teóricas abstratas para a mensuração de aspectos do comportamento humano que não são diretamente observáveis, como a proficiência oral, por exemplo. Desse modo, um construto não pode ser medido ou quantificado diretamente, mas pode ser inferido com base na coleta e análise de evidências.

Quando há validade de construto, os dados obtidos por instrumentos avaliativos são representativos daquilo que o avaliador entende como indícios da habilidade avaliada. Para o escopo desta discussão, o foco recai na validade de conteúdo, uma faceta especialmente importante em contextos de sala de aula, em que o propósito da avaliação é voltado para observar o progresso dos aprendizes em relação ao que foi ministrado anteriormente. Para que haja validade em termos de conteúdo, é preciso que haja relação direta entre os conhecimentos e as habilidades exigidos dos alunos nas situações avaliativas e o que foi de fato trabalhado anteriormente como objetivos de aprendizagem (Brum-Barreto, 2023).

Outro princípio importante a ser considerado é a confiabilidade. Esse conceito diz respeito à consistência e estabilidade de uma avaliação. Como Borges-Almeida (2023, p. 41) destaca,

o conceito de confiabilidade refere-se ao grau de consistência de um instrumento ou método avaliativo. [...] Confiabilidade do método refere-se diretamente aos examinadores ou corretores que atribuem as notas a uma resposta aberta a uma tarefa ou item. Trata-se do quanto uma avaliação é isenta de subjetividade no momento específico da correção ou atribuição de notas.

Segundo Brown e Abeywickrama (2018), um teste confiável é aquele que (1) possui condições consistentes entre duas ou mais aplicações; (2) oferece instruções claras para atribuição de notas e avaliação; (3) possui rubricas uniformes para atribuição de notas e avaliação; (4) permite uma aplicação consistente das rubricas utilizadas pelo avaliador; e (5) contém itens ou tarefas que não são ambíguas para o avaliando. Os autores consideram que diferentes fatores podem incidir sobre a confiabilidade de um instrumento avaliativo. A confiabilidade pode ser afetada, por exemplo, pelo estado físico e emocional do estudante (confiabilidade relacionada ao estudante), o que pode fazer com que o resultado daquele aluno não seja fidedigno ao seu potencial desempenho em outro momento. Já a confiabilidade do instrumento pode ser prejudicada pela redação de itens e/ou comandos ambíguos e pelo tempo insuficiente para a finalização das respostas. Em avaliações subjetivas, respostas abertas, que dependem do julgamento de valor/qualidade do professor para serem consideradas adequadas ou inadequadas, podem impactar na confiabilidade do processo avaliativo, caso não haja uma predefinição clara dos critérios a serem observados.

Por fim, o princípio da praticidade é mais comumente diretamente relacionado às questões de logística na aplicação e correção de uma avaliação (Brown; Abeywickrama, 2018). Porém, refere-se não apenas à simplicidade ou facilidade de uso dos instrumentos, mas ao quanto um processo avaliativo e tudo o que ele envolve é praticável e exequível. Para que seja considerado prático, é preciso que esteja dentro dos limites de custo e tempo, além de ser executável em sua elaboração, administração, correção e modo de reportar os resultados.

Esses conhecimentos, juntamente com as habilidades para conduzir processos avaliativos apoiados em princípios são essenciais ao fazer docente, já que avaliar é inerente e indissociável do processo de ensino-aprendizagem. O papel do professor-avaliador é significativo durante o processo de avaliação, pois está diretamente ligado ao desempenho do aluno. Nesse sentido, é preciso que os conhecimentos, habilidades e princípios sejam integrados para que esse professor-avaliador possa fornecer feedback de qualidade aos seus aprendizes. Tal feedback é um aliado para desencadear mudanças relevantes no processo de ensino-aprendizagem, uma vez que “deve ser projetado para deixar espaço suficiente para que os alunos ganhem confiança para tomar suas próprias decisões de maneira informada e ponderada” (Black, 2015, p. 171).

Essa perspectiva se relaciona com o conceito de feedback como andaime (*feedforward*) proposto por Carless (2007) como um dos pilares da aprendizagem orientada pela avaliação. Para ser considerado *feedforward*, é preciso que haja uma orientação para ações futuras, ou seja, o professor-avaliador deve propor *feedback* simultaneamente retrospectivo (detalhando os pontos fortes do desempenho do aprendiz na tarefa avaliativa e apontando para eventuais inconsistências para serem sanadas) e, principalmente, prospectivo (evidenciando estratégias para mudar seu desempenho a fim de melhorá-lo em uma nova tarefa semelhante).

Trabalhando com o letramento de professores de línguas em formação inicial

Conforme afirmamos anteriormente, recentemente temos visto crescer o volume de estudos sobre letramento em avaliação na formação inicial de professores, especialmente na América do Sul. Discutimos, a seguir, os estudos que foram desenvolvidos em contextos semelhantes ao nosso, em disciplinas formais de cursos de

licenciatura.

Na Colômbia, Giraldo e Murcia (2019) investigaram o desenvolvimento do LAL de professores em formação inicial em uma disciplina ministrada para os alunos do oitavo semestre do curso equivalente à graduação em Letras numa universidade colombiana. Tanto o contexto de pesquisa, quanto os instrumentos para a coleta de dados (questionários, entrevistas semiestruturadas e observação de aulas), guardam semelhança com a proposta do presente trabalho. Durante o curso, os alunos foram apresentados a uma visão geral das questões teóricas fundamentais em avaliação (por exemplo, validade e confiabilidade), estudando essa teoria por meio da análise de amostras de avaliações. Alguns dos instrumentos de avaliação discutidos possuíam problemas de design para permitir que eles usassem a teoria para tecer críticas contundentes. Posteriormente, os alunos aprenderam a formular itens e tarefas avaliativas e receberam *feedback* sobre seu desempenho.

Os resultados do estudo de Giraldo e Murcia (2019) mostram que houve uma mudança significativa nas concepções de avaliação de línguas e um desenvolvimento dos repertórios teóricos e práticos dos participantes. Os autores relatam que “antes do curso, os participantes pensavam que avaliação de línguas era sobre notas ou testes. No entanto, durante o curso, [...] os professores em formação desenvolveram uma perspectiva mais ampla sobre avaliação de línguas” (Giraldo; Murcia, 2019, p. 249).

Também foi apontado no estudo que os momentos de prática, voltados para a análise e elaboração de atividades avaliativas, estimularam o uso da teoria e o crescimento da conscientização sobre a necessidade de contextualizar os instrumentos de avaliação. Segundo os autores, o ato de elaborar as atividades impactou seu desenvolvimento profissional em aspectos teóricos, técnicos, práticos, contextuais e, também, nas dimensões críticas da avaliação. Isso ocorreu, pois esse é “um exercício que envolve dimensões teóricas, técnicas e operacionais” (Giraldo; Murcia, 2019, p. 251). Por fim, os autores concluem que cursos de avaliação de línguas para professores em formação devem oportunizar momentos para a elaboração de tarefas, porque esses momentos desencadeiam decisões conscientes baseadas em referenciais teóricos adequados.

Restrepo Bolivar (2020), também na Colômbia, investigou o processo de desenvolvimento do letramento de 18 professores de língua estrangeira durante uma disciplina semestral sobre avaliação em um curso de graduação por meio de diários

reflexivos. Os resultados encontrados por meio da análise de conteúdo revelaram que os participantes apresentaram ganhos em seu letramento em cinco dimensões: (a) concepção de avaliação e seu propósito; (b) papéis de professor e aluno; (c) clareza e distinção entre objetos e métodos de avaliação; (d) reconhecimento de questões éticas de modo mais fundamentado; e (e) consciência sobre as implicações de más práticas em avaliação.

Interessa-nos, particularmente para a discussão que trazemos neste artigo, observar que o componente curricular formal sobre avaliação contribuiu para que os professores em formação modificassem sua concepção prévia de avaliação como verificação e de avaliação dissociada entre somativa e formativa para uma compreensão de avaliação como processo de (re)direcionamento pedagógico. Além disso, houve indícios de que também mudou a visão inicial de que ao professor cabe um papel ativo, e ao aluno resta o passivo (receber feedback).

Embora o trabalho contribua imensamente para compreendermos os processos reflexivos que permitem o desenvolvimento do letramento, é preciso considerar as limitações metodológicas do estudo, que consistem principalmente no uso dos diários como único instrumento de coleta de dados. Diários exigem tempo, organização, responsabilidade, motivação e alto engajamento em atividades extra classe por parte dos alunos (conforme refletem Jaramillo Delgado e Gil Bedoya, 2019, estudantes que participaram do curso), e limitações em qualquer uma desses requisitos pode impactar a confiabilidade dos resultados.

Ademais, apesar de capturarem aspectos reflexivos que outros instrumentos não são capazes de acessar, diários apresentam os produtos dos discursos dos alunos, que podem muitas vezes, ecoar os dizeres sobre avaliação ouvidos de agentes culturalmente considerados as vozes do saber, como os professores formadores. Assim, são instrumentos que permitem-nos inferir mais precisamente os componentes do letramento de conhecimento e princípios, e menos o de habilidades. No Brasil, o efeito de componentes curriculares sobre o LAL de professores de línguas estrangeiras em formação inicial também tem sido recentemente debatido em trabalhos muito relevantes para a nossa discussão, uma vez que compartilham a mesma cultura educacional local.

O primeiro trabalho que conhecemos que investigou os efeitos de uma disciplina sobre avaliação em contexto de línguas sobre o letramento de alunos de Letras é o de Fernandes (2019), relatado também em Fernandes e Borges-Almeida (2021). O estudo

revelou que, antes do curso, os professores de espanhol em formação inicial possuíam pouco conhecimento formal sobre avaliação, porém, traziam consigo uma bagagem empírica, proveniente de suas experiências anteriores com diferentes modelos avaliativos. Ao longo das aulas da disciplina, os estudantes demonstraram avanços ao passo que apropriaram dos termos específicos, ou seja, passaram a empregar termos técnicos para se referirem a conceitos da avaliação, e foram capazes de explicá-los nas atividades práticas (Fernandes, 2019; Fernandes; Borges-Almeida, 2021). Os resultados apontaram ainda para uma ampliação do entendimento sobre a função da avaliação, alinhada à perspectiva qualitativa e formativa, corroborando os estudos de Giraldo e Murcia (2019) e Restrepo Bolivar (2020).

Uma limitação do estudo é que o objetivo do componente curricular proposto era voltado para o estímulo à reflexão crítica e aumento do conhecimento dos professores em formação inicial sobre a temática, e, pela limitação da carga horária, não foram propostas atividades de análise e/ou elaboração de instrumentos avaliativos. Pelo fato de o LAL ser composto não apenas de conhecimentos e princípios, mas também de habilidades práticas, entendemos que as ações para sua promoção não devem apenas ser teóricas; precisam aliar reflexões e prática articulada com a realidade educacional (Giraldo; Murcia, 2018; 2019).

Já no contexto de estágio obrigatório, Brum-Barreto (2023) conduziu um estudo cujo objetivo era promover e analisar uma intervenção pedagógica voltada para o desenvolvimento do letramento em avaliação em contexto de línguas de professores de espanhol como língua estrangeira em formação inicial. Nessa pesquisa-ação, a autora buscou proporcionar insumos teóricos e, conjuntamente, oportunidades práticas, para que os licenciandos participantes pudessem criar seus instrumentos avaliativos e, em seguida, aplicá-los, sob supervisão da pesquisadora, aos alunos dos contextos em que lecionaram como estagiários. A autora afirma que “foi possível observar indícios de uma ampliação da visão de avaliação, mais alinhada com sua dimensão formativa e embasada em princípios éticos” (p. 167).

No entanto, também identificou algumas lacunas nesse letramento inicial, especialmente no que diz respeito à criação de rubricas, critérios de avaliação e descritores que estejam em sintonia com uma concepção de língua centrada no uso, bem como uma visão de avaliação holística e formativa. Isso sugere que a cultura de avaliação das

participantes, influenciada por uma visão estruturalista da língua e pela tendência a testar (entender avaliação como sinônimo de aplicação de provas), atua como um quadro interpretativo que impacta o desenvolvimento da competência em avaliação no contexto linguístico das participantes.

Outro aspecto relevante desse estudo, discutido posteriormente por Brum-Barreto e Borges-Almeida (2023, p. 362), é que, apesar de terem sido observadas contribuições para o desenvolvimento do letramento em avaliação, não se pode “afirmar que houve uma reconfiguração da visão de língua, que permanece caracterizada como centrada na precisão das formas gramaticais”. As autoras recomendam a realização de estudos complementares para buscar compreender a relação entre o desenvolvimento do letramento em avaliação e a reconstrução da cultura de avaliar de professores, o que precisa ser considerado ao elaborar novas propostas didáticas para formação de professores.

METODOLOGIA

Esta pesquisa qualitativa é um estudo de caso interventivo (Faltis, 1997) que buscou investigar o impacto de uma disciplina teórico-prática intitulada “Inglês: Laboratório de Ensino 3 (avaliação)” sobre o letramento em avaliação de alunos de Letras na Universidade de Brasília.⁵ A disciplina, cuja ementa explicita como objetivo geral fomentar o “letramento em avaliação no contexto de línguas (inglês), a partir da conceituação de construtos elementares, considerando necessidades, objetivos e funções da avaliação na prática escolar” (Universidade de Brasília, 2019), possui carga horária de 90h e foi implementada em 2019, a partir de uma ampla reforma curricular. Não há, até então, trabalhos publicados sobre os resultados dessa implementação. Dos treze alunos matriculados, sete são participantes deste estudo (que aqui serão referidos como E1, E2,..., E7), estão na faixa etária de 19 a 25 anos e têm pouca (até um ano) ou nenhuma experiência docente.⁶ A professora formadora possui sólida trajetória acadêmica em

⁵ Este artigo é recorte da dissertação de mestrado da primeira autora, orientada pela segunda. A metodologia aqui apresentada descreve o estudo geral para fins de esclarecimento do leitor. Os instrumentos e dados aqui analisados referem-se especificamente ao recorte. O texto do artigo em si não é mera reprodução do trabalho original, mas foi produzido por ambas as autoras.

⁶ Aqueles que têm alguma experiência, são professores auxiliares, como estagiários, que participam da dinâmica escolar, porém não são responsáveis por elaborar e/ou implementar propostas pedagógicas e

avaliação de línguas estrangeiras, tendo realizado suas pesquisas de mestrado e doutorado na área. Além disso, é credenciada na *International Language Testing Association* (ILTA) e na *Latin American Association for Language Testing and Assessment* (LAALTA).

Entre novembro e dezembro de 2022, foram observadas nove aulas, das quais quatro abordaram princípios da avaliação (validade, confiabilidade, praticidade, autenticidade, efeito retroativo) e cinco inseriram-se num módulo sobre avaliação de produção oral. Tipicamente, os alunos envolviam-se em tarefas de discussão e análise de materiais e aplicações de provas, com base em leituras previamente orientadas. No módulo sobre produção oral, o enfoque foi dado a avaliações formais, do tipo entrevista oral, a partir de exemplos de aplicações de exames de larga escala, como os de Cambridge ESOL.

Para o recorte deste artigo, analisamos os dados coletados por meio de duas tarefas práticas (uma diagnóstica e uma final), seguidas de duas entrevistas estruturadas, aplicadas antes e depois da formação, para fins de comparação. Além disso, notas de campo foram tomadas ao longo das observações de aula para que pudéssemos acompanhar e compreender melhor o processo de ensino. As tarefas eram parte integrante da avaliação na disciplina, e foram construídas colaborativamente entre as autoras deste texto e a docente formadora. Tiveram a finalidade de levantar informações que permitissem inferir as concepções, o conhecimento e as habilidades em avaliação por meio de ações práticas de elaboração de um instrumento para avaliação da produção oral em uma situação hipotética (figura 2).

Foi disponibilizado também o acesso à unidade completa indicada do livro didático (*American Inside Out - Elementary*). Os instrumentos elaborados não foram aplicados. A tarefa final foi uma versão paralela da tarefa inicial, alterando-se a situação hipotética quanto à unidade do livro didático e o nível de proficiência (A2 e B1). Para o recorte deste artigo, analisamos apenas as propostas dos alunos para os critérios de correção. As entrevistas, por sua vez, contribuíram para que os participantes explicassem suas decisões e apresentassem seus pontos de vista sobre avaliação. Mais detalhes sobre essas decisões podem ser obtidos lendo a dissertação.

avaliativas.

Figura 2 – Tarefa inicial

TAREFA I

Você é professor regente de uma escola de idiomas e precisa elaborar um instrumento para avaliar o *speaking* seus alunos. As opções de instrumentos permitidos pela escola são: entrevista ou roleplay. Veja abaixo as informações sobre sua turma e o material didático utilizado e elabore sua avaliação.

Nível	Idade dos alunos	Quantidade de alunos na turma	Abordagem/ Metodologia
Básico (QCREL: A2)	15 a 17 anos	20 alunos	Comunicativa

Conteúdos			
Units & topics	Speaking & Writing	Reading & Listening	Grammar, Vocabulary, & Pronunciation
1 Travel Introductions Nationalities Hotels	S Meeting on a plane S Hotel facilities WB W Using capital letters and periods Completing a form with personal information	L Going through customs L Airline telephone numbers R The Beach House Hotel Language for Life: Asking for clarification Life Skills Social Skills: Giving and Receiving Feedback	G be: simple present; a/an; there is/there are V Countries/Nationalities, Numbers 0-90, Objects, The alphabet. P Sounds of the alphabet

Fonte: Livro de nível A2 “American Inside Out – Elementary” (disponível para consulta)

Durante o processo de confecção de sua avaliação, responda (em português ou inglês) as questões abaixo:

1. Quais serão os objetivos desta avaliação?
2. Quais serão os critérios avaliativos?
3. Qual o papel do examinador e do aluno durante a realização do processo avaliativo?
4. Será utilizado algum material de apoio?
5. Como você daria o feedback da atividade avaliativa para seus alunos? Por quê?
6. Como será a aplicação da atividade?

Em seguida, formule o instrumento avaliativo a ser utilizado e envie juntamente com as respostas das perguntas acima.

Fonte: Damacena-Dutra (2023), p. 84.

Todos os dados foram analisados segundo o método de análise de conteúdo proposto por Bardin (2011), buscando-se estabelecer comparação entre os dados da etapa diagnóstica e da final. Neste artigo, discutimos os resultados relacionados a categorias e subcategorias que nasceram dos dados e se relacionam com indícios de letramento alinhados a uma perspectiva de avaliação como mediação da aprendizagem.

Finalmente, quanto aos procedimentos éticos, o estudo foi aprovado pelo Comitê de Ética em Pesquisa em Ciências Humanas e Sociais (CEP/CHS) da Universidade de Brasília, sob o número de protocolo 5.480.945 e Certificado de Apresentação de Apreciação Ética número 57696622.0.0000.5540.

RESULTADOS

Nesta seção, apresentamos e discutimos indícios de desenvolvimento do LAL contrastando os dados coletados antes e durante/após os módulos da disciplina. Para organizar nossa discussão, dividimos a seção em duas partes. Na primeira, discutimos os indícios de aprendizagem dos conceitos de validade e praticidade, destacando a relevância da experiência empírica dos participantes e sua oscilação entre dois papéis sociais. Na segunda, abordamos os efeitos da disciplina sobre o entendimento do conceito de confiabilidade e as concepções de avaliação.

Da experiência como aluno para o letramento como professor

Os modelos avaliativos aos quais somos submetidos ao longo de nossas vidas escolares e como membros de uma dada sociedade moldam nossas concepções e expectativas sobre funções, práticas e procedimentos apropriados em situações de avaliação (Rolim, 1998).

Os dados coletados no início desta investigação revelaram que os licenciandos traziam para a disciplina um tipo de conhecimento prévio, informal e empiricamente construído, adquirido nessas experiências anteriores como discentes (excertos 1, 2 e 3), e resignificavam seus olhares a partir delas, como veremos mais adiante.

Excerto 1:

[E5]: Geralmente, a maioria das vezes era só a nota e um "very good" ou alguma coisa assim. E aí, juntamente o professor corrigia a prova toda com a turma: pegava questão por questão. Mas não tinha muito *feedback* individual. [...] Nunca teve, por exemplo, *feedback* de speaking, tipo em teste oral que a gente fez... nunca teve. Assim, era só a nota mesmo. Você pega e vai feliz pra sua casa - ou triste." (6:25 p 5 em Entrevista Inicial E5)

Excerto 2:

[E4]: [...] quando aprendi francês, um professor fazia o seguinte: na entrevista, o professor fazia as anotações dele e, em uma aula seguinte, quando ele terminava de fazer todas as entrevistas com todos os alunos, ele passava por cima de tudo, tipo: escrevia no quadro e falava "Gente, vamos reforçar isso?", "Isso não se fala dessa forma... não se usa desse jeito" (sem especificar o aluno que cometeu tal erro), "Vocês se lembram, que eu falei isso?" e tal. É muito interessante porque você pode revisar conteúdo sem deixar as pessoas constrangidas. [...] Isso foi totalmente uma influência para mim, porque acho que, como eu disse, eu não recebia muito *feedback* quando eu fazia inglês... não que eu me lembre, assim, especificamente. E esse professor começou a fazer isso e me marcou bastante nessa questão de corrigir erros de uma forma

diferenciada. (7:35 p 5 em Entrevista Inicial E4)

Excerto 3:

[E4]: Essa atividade [tarefa inicial] foi totalmente empírica, baseada na minha experiência como aluno, então, muito do que eu fiz foi o que achava era certo.” (11:30 pp 1 – 2 em Entrevista Final E4)

Considerando as qualidades de boas práticas avaliativas, inicialmente vemos que questões relacionadas à validade de conteúdo (Scaramucci, 2011) foram a categoria mais recorrente nas respostas dos participantes à pergunta “para você, uma boa avaliação é aquela que...”. Sobre essa temática, os licenciandos destacaram a necessidade de haver correspondência entre o que foi estudado e o que é “cobrado” na avaliação (excerto 4).

Excerto 4:

[E5] “Testa o que realmente o aluno foi ensinado.” (6:6 p 1 em Entrevista Inicial E5)

[E2] “Uma boa avaliação é uma avaliação que corresponde com aquilo que foi passado, porque, às vezes, acho que muitas coisas acontecem, e as questões ou que tá sendo avaliado na prova oral não são o que a gente aprendeu durante o semestre ou durante as aulas. Eu acho que corresponde ao que foi ensinado e às expectativas que a gente tem como aluno, né? É. Eu acho isso.” (4:6 p 2 em Entrevista Inicial E2)

O segundo aspecto mais relevante, na perspectiva deles, é a preocupação de que o processo avaliativo seja respeitoso para com os alunos, sem que se torne ferramenta punitiva ou sem que o professor exponha os aprendizes a situações embaraçosas (excerto 5).

Excerto 5:

[E4] “Uma boa avaliação é aquela que é feita de forma a deixar os alunos um pouco mais à vontade e que de fato tenha... por exemplo, foi ensinado tal coisa, então vai ser cobrado essas coisas para que haja correspondência de conteúdo ensinado e a matéria cobrada.” (7:13 p 2 em Entrevista Inicial E4)

[E1] “Acho que é aquela que não assusta. Acho que grande de parte dos problemas com a avaliação e com notas ruins (e com tudo que a gente vê por aí, de reclamações), é porque ela assusta.” (5:10 p 4 em Entrevista Inicial E1)

Tais resultados são semelhantes àqueles encontrados por Restrepo Bolivar (2020), que identificou que os licenciandos inicialmente “associam uma avaliação justa à transparência e compartilhamento de critérios avaliativos e à inclusão de conteúdos e tarefas familiares aos estudantes.” Embora realmente essas sejam questões importantes,

no início da disciplina, os participantes ainda olham para a questão pela ótica do aluno, e não do professor, embora o componente curricular em que nossa investigação se insere seja cursado ao final do segundo ano de formação no curso, e tenha como pré-requisito a disciplina de Metodologia de Ensino de Línguas (90h). Dessa forma, os participantes, que não têm experiência docente, mencionam a importância de pensar aspectos voltados para os interesses, emoções, contextos e necessidades individuais dos alunos. Porém, questões que dizem respeito às ações e decisões do professor, como o planejamento do processo avaliativo (como desenho de tarefas, elaboração de provas ou definição de critérios) e a aplicação da avaliação (por exemplo, recursos disponíveis, o tempo e os critérios de avaliação) não são mencionadas ou parecem-lhes de pouca relevância, como no excerto 6, em que E4 não atenta para a pouca praticidade de conduzir entrevistas individuais no contexto hipotético proposto, seja por questões de tempo, seja por espaço físico.

Excerto 6:

[Comando]: Como será a aplicação da atividade?

[E4]: “Será uma entrevista individual, em um espaço à parte da sala: para que o aluno tenha privacidade e se sinta mais à vontade. (20:11 p 1 em *Tarefa Inicial*- E4)

Em contraste, na tarefa e na entrevista finais essas questões já aparecem, como discutido por E4 (excerto 7), e inclusive são citadas como das mais relevantes na disciplina por E5 (excerto 8).

Excerto 7:

[E4] “Ah, eu acho que [...] [a avaliação] deve considerar também os recursos disponíveis para o professor no momento da avaliação. A gente falou muito disso na aula, então, por isso, é uma coisa que vem sempre na minha mente: “Se eu estivesse aplicando uma avaliação de *speaking*, o que os outros alunos fariam?” Então precisa pensar em toda a infraestrutura disponível para o professor.” (11:19 p 3 em Entrevista Final - E4)

Excerto 8:

[E5]: “[A disciplina] ajudou a planejar e ver a avaliação de uma forma muito mais ampla e muito mais completa... pensando em todos os aspectos daquele momento da avaliação. Não só a avaliação material em si, mas todas as outras variáveis que poderiam estar acontecendo ali naquele meio. (2:20 p 1 em Entrevista Final – E5)

Há estudos mostrando que as questões de praticidade, como as mencionadas no excerto 7, figura entre as maiores preocupações dos professores em serviço quando se

trata de avaliação (Armstrong; Marín; Patarroyo, 2004). Como professores em formação inicial e sem experiência docente, que nunca tiveram de envolver-se com os desafios que cercam a aplicação de provas orais, os participantes adquiriram outra perspectiva, a de professores. Ademais, o baixo nível de letramento dos discentes, ao início do processo de formação, é fruto de práticas escolares que não os envolvem na avaliação, a qual fica concentrada nas mãos do professor, “é abstrata” e “fica escondida [...] na sombra”, como bem reconhece E5 (excerto 9).

Excerto 9:

[E5] “Acho que talvez os critérios avaliativos [...] é, talvez, a parte mais abstrata, que é uma parte que geralmente fica escondida da gente. Pra a gente, enquanto aluno, é a parte que fica mais na sombra. A gente não ouviu muito falar.” (6:22 p 5 em Entrevista Final E5)

Restrepo Bolivar (2020) também observou que no início da formação específica na disciplina sobre avaliação, os professores em formação inicial compreendiam que um papel ativo durante os processos avaliativos é desempenhado pelo professor, enquanto aos alunos cabe o papel passivo de receber feedback. Esse é o resultado de uma cultura que concentra nas mãos do professor o processo educacional. Inbar-Lourie (2008) advoga que é preciso envolver os alunos em processos de autoavaliação e avaliação por pares para trazer os alunos a um papel ativo.

Evidentemente, as concepções sobre os processos avaliativos, adquiridas empiricamente, não podem ser ignoradas, porque possuem relação direta com as práticas de ensino e avaliação a serem adotadas futuramente pelos professores em formação. Entretanto, é preciso que haja uma tomada de consciência do discente quanto a seu repertório e à necessidade de sua des- e reconstrução, de modo que sejam estabelecidas relações explícitas entre suas experiências como aluno e sua prática futura como professor (Berry et al., 2019 apud Restrepo Bolivar, 2020). Os próprios participantes percebem, ao final da formação, a diferença entre a ótica do aluno e a do professor (excerto 10).

Excerto 10:

[E1]: Com certeza foram muito úteis as aulas, os materiais, as atividades [...] foram muito úteis pra mudar bastante né? A visão sobre avaliação que a gente tem é visão de avaliação muito rasa, não tem transparência na maioria das vezes. Então, acho que me mostrou a minha profissão como algo profissional realmente, como algo que exige técnicas, que exige muito estudo, muita dedicação [...] 12:17 p 1 em Entrevista Final – E1)

[E5]: Eu acho que foi uma das disciplinas que mais tocou assim no assunto do “ser professor”, da prática do professor no dia a dia, porque eu acho que as outras disciplinas eram muito teóricas e, nessa, realmente, a gente veio pra prática: “é isso que você vai fazer”, “é esse o trabalho que você vai fazer no dia a dia”. (12:29 p 5 em Entrevista Final - E5)

Os resultados dão suporte à discussão de que a construção do letramento em avaliação na formação inicial se dá por meio do conflito e da complementação de saberes, proporcionados pelas reflexões que se tornam possíveis por meio das atividades práticas, em consonância com Giraldo e Murcia (2019), e corroboram a afirmação de que o futuro professor é, na verdade, um “*aprendente e ensinante*” (Jordão; Buhner, 2013, p. 670). Isso significa que suas identidades e culturas (de *aprender, ensinar e avaliar*) prévias coexistem e se ressignificam através das reflexões sobre docência propostas na licenciatura. Assim, a aprendizagem e acomodação dos conhecimentos, habilidades e princípios da avaliação ocorrem em um contexto de transição, por isso, podem ser percebidas intersecções e contradições entre os dois papéis assumidos, uma vez que o “aluna/o-professor/a [...] ocupa simultaneamente as duas posições” (Damacena; Mastrella-de-Andrade, 2018).

Confiabilidade e concepções de avaliação

Quanto às concepções de avaliação antes da formação na disciplina, o cruzamento dos dados evidenciou que os participantes trazem para a sala de aula a concepção de avaliação como verificação (excerto 11), corroborando os resultados amplamente reportados na literatura (Giraldo; Murcia, 2019; Restrepo Bolivar, 2020; Brum-Barreto; Borges-Almeida, 2023; Fernandes, 2019). Mesmo a metade dos alunos que ecoou falas de avaliação formativa em resposta à primeira entrevista, não foi capaz de apresentar, na tarefa inicial, evidências que dessem qualquer suporte ao uso da avaliação como “mola propulsora da aprendizagem” (Scaramucci, 2016), como a definição de critérios minimamente claros ou a preocupação com a autenticidade e potencial motivacional da tarefa. Após a instrução formal, há o início de um movimento em direção à valorização da clareza de critérios, porém as limitações do estudo não nos permitem identificar se e como os participantes de fato conduziram avaliações formativas, uma vez que não há

contexto de aplicação.

Excerto 11:

[E6] “Eu penso que a avaliação serve pra saber se os alunos realmente adquiriram aquela língua, né? No caso, nosso conteúdo. [...] ‘será que eles realmente aprenderam como usar a gramática?’ *Eu penso que a avaliação é assim: uma forma de você ver mesmo se os alunos entenderam.*” (2:3 p 1 em Entrevista Inicial E6)

[E1] “Eu acho a avaliação muito interessante no ensino de língua estrangeira. Acho que o sentido geral dela é checar [...] se o aluno está aprendendo ou se precisa trabalhar mais.” (5:5 p 2 em Entrevista Inicial E1)

Quanto à confiabilidade, na tarefa e na entrevista iniciais⁷, os participantes nomearam os critérios em tópicos, e nenhum deles apresentou ferramentas como grades de correção para demonstrar a operacionalização desses critérios (quadros 2 e 4). Já na tarefa e entrevista finais, vemos ferramentas de correção como escalas (quadro 3) e rubricas (quadro 5). Embora concordemos que careçam de melhorias para que possam ser adequados às tarefas (por exemplo, o quadro 23 é a escala geral de proficiência do nível B2 de Cambridge) e orientar a correção confiável e o feedback ao aluno, não deixam de indicar que os participantes avançaram na compreensão de que é preciso operacionalizar os critérios para que haja confiabilidade e transparência, qualidades essenciais de boas práticas avaliativas.

Quadro 2 – Tarefa inicial de E5

2. Quais serão os critérios avaliativos?

- *Conhecimento da diferença entre Country e Nationality (Russia, Russian)*
- *Domínio da estrutura do Simple Present*
- *Domínio do vocabulário de apresentação (Introduction)*
- *Domínio do vocabulário de Travel*
- *Domínio do sons do Alfabeto em inglês*

Fonte: Damacena-Dutra (2023), p. 133.

⁷ Para fins de ilustração, trazemos, no Anexo, a tarefa inicial elaborada por um dos participantes. Pelo escopo proposto e limitações de espaço, os dados que discutimos neste artigo referem-se apenas a como os critérios propostos são trazidos.

Quadro 3 – Tarefa final de E5

2. Quais serão os critérios avaliativos?

		Needs improvement	Good	Excellent
1	Uso do vocabulário sobre <i>Friends</i> e <i>Friendships expressions</i>			
2	Uso gramaticalmente correto da estrutura dos <i>Question forms</i>			
3	Uso gramaticalmente correto da estrutura dos <i>Adverbs of frequency</i>			
4	Inteligibilidade da pronúncia			
5	Organização do discurso			

Fonte: Damacena-Dutra (2023), p. 133.

Quadro 4 – Tarefa inicial de E2

2. Quais serão os critérios avaliativos?

“2.0. Boa comunicação com o professor e com os colegas
2.1. Conhecimento do tema
2.2. Uso do tempo verbal estudado
2.3. Concisão
2.4. Pronúncia”

Fonte: Damacena-Dutra (2023), p. 133

Quadro 5 – Tarefa final de E2

2. Quais serão os critérios avaliativos?

“Os critérios serão os mesmos definidos pelo QCER: gramática e vocabulário, pronúncia, interactive communication e discourse management.
Os alunos terão acesso à rubrica antes da prova para tirar dúvidas e entender melhor o processo de avaliação.”

	1	2	3	4	5
GRAMMAR AND VOCABULARY	Shows sufficient control of simple grammatical forms. Uses a limited range of appropriate vocabulary to talk about familiar topics.	Between 1 and 3.	Shows a good degree of control of simple grammatical forms. Uses a range of appropriate vocabulary when talking about familiar topics.	Between 3 and 5.	Shows a good degree of control of simple grammatical forms, and attempts some complex grammatical forms. Uses a range of appropriate vocabulary to give and exchange on familiar topics.
DISCOURSE MANAGEMENT	Produces responses which are characterized by short phrases and frequent hesitation. Repeats information or digresses from the top.	Between 1 and 3.	Produces responses which are extended beyond short phrases, despite hesitation. Contributions are mostly relevant, but there may be some repetition. Uses basic cohesive devices.	Between 3 and 5.	Produces extended stretches of language despite some hesitation. Contributions are relevant despite some repetition. Uses a range of cohesive devices.
PRONUNCIATION	Is mostly intelligible, despite limited control of phonological features.	Between 1 and 3.	Is mostly intelligible, and has some control of phonological features at both utterance and word levels.	Between 3 and 5.	Is intelligible. Intonation is generally appropriate. Sentence and word stress is generally accurately placed. Individual sounds are generally articulated clearly.
INTERACTIVE COMMUNICATION	Maintains simple exchanges, despite some difficulty. Requires prompting and support.	Between 1 and 3.	Initiates and responds appropriately. Keeps the interaction going with very little prompting and support.	Between 3 and 5.	Initiates and responds appropriately. Maintains and develops the interactions and negotiates towards an outcome with very little support.

Fonte: Damacena-Dutra (2023), p. 134

Atribuímos esses resultados ao trabalho desenvolvido na disciplina, a partir de atividades reflexivas e práticas. Por exemplo, ao abordar o tema das questões a serem consideradas ao planejar uma avaliação de produção oral, a professora formadora propôs uma discussão a partir de afirmativas para as quais os alunos deviam decidir se eram “verdadeiras” ou “falsas” (quadro 6). Nesse momento, os participantes tiveram, pela primeira vez no curso, contato com o termo “grade de avaliação”. A partir desse tópico, iniciou-se um debate sobre a relevância de serem estabelecidos critérios avaliativos claros e condizentes com o objetivo da avaliação para que, com base em características observáveis, fosse possível avaliar a produção oral de maneira mais válida, confiável e prática. Depois disso, ao longo das aulas, a professora também encorajou os alunos a elaborarem suas próprias grades de correção, apresentando diversos modelos. E5, por exemplo, optou por utilizar uma escala de valores (quadro 3), enquanto E2 selecionou uma escala mais complexa, com descritores para cada categoria (quadro 5).

Quadro 6 – Exemplo de atividade de reflexão usada em Laboratório 3

<i>Activity 3</i>		
<i>When we want to assess L2 learners' speaking ability, we need to consider a number of issues. Read some statements below and decide whether they are 'true' or 'false'. A third option is also included (i.e. 'I can't decide') for you to choose when you think that whether a statement is true or false depends on the context. The answer key and some comments are provided in the Suggested Answers file.</i>		
Question 6: <i>The examiner/teacher can correct the learner's mistakes/errors while they are in interaction.</i>		
<i>a. True</i>	<i>b. False</i>	<i>c. I can't decide</i>
Question 8: <i>The scoring of spoken performance is more reliable when an analytic rating scale is used.</i>		

<i>a. True</i>	<i>b. False</i>	<i>c. I can't decide</i>
----------------	-----------------	--------------------------

Fonte: Adaptada da apostila do curso elaborada pela professora formadora. Módulo 2: *Assessing Speaking*.

Para além dos ganhos no letramento quanto à confiabilidade da correção, o trabalho desenvolvido na disciplina permitiu que os participantes compreendessem a importância da transparência dos critérios e da orientação aos alunos, ou seja, o conhecimento sobre como implementar práticas avaliativas formativas, que iniciam muito antes da entrega dos resultados de uma prova (excertos 12 e 13).

Excerto 12:

[E3] O feedback será disponibilizado no próprio ambiente virtual em que o arquivo foi enviado. O professor irá utilizar os *rubrics* anteriormente disponibilizados para a turma. (p. 2 em E3- Tarefa Final)

Excerto 13:

[E2] Os alunos terão acesso à rubrica antes da prova para tirar dúvidas e entender melhor o processo de avaliação (p. 2 em E2- Tarefa Final)

A importância do *feedback* baseado no uso de rubricas, com o objetivo de orientar o aluno sobre sua aprendizagem, também foi tópico abordado nas discussões durante a disciplina pela professora formadora. No momento da discussão, E6 narrou que, em sua experiência enquanto aluna, raramente o *feedback* é útil: “às vezes, só recebo o comentário ‘*good job*’ e pronto” (Notas de campo, página 4, 08/11/2022). Outros licenciandos relataram experiências parecidas e a professora formadora aproveitou a discussão para enfatizar a importância do retorno formativo, que comunica ao aluno sobre seu processo de ensino-aprendizagem e sobre o desenvolvimento de suas habilidades linguísticas.

Os resultados revelam que os ganhos em desenvolvimento em letramento quanto ao componente confiabilidade, por meio de definição de critérios operacionalizados em instrumentos de correção, foi um dos aspectos mais proeminentes nesta investigação, como observa E2 (excerto 14):

Excerto 14:

[Pesquisadora]: E, bom, teve algum conhecimento específico, algum tema, que te marcou mais durante a disciplina?

[E2]: “Eu acho que foi a apresentação da rubrica mesmo, de colocar bem explícito pro aluno e pra mim também, para ficar mais fácil para o instrutor

avaliar e colocar as notas. [...] Na minha experiência, eu nunca fiquei sabendo como que era avaliada ou como o professor chegava a uma nota e o que estava sendo avaliado também nunca me foi dito. E depois do primeiro *Hands-on* [Tarefa Final] que a gente fez, eu comecei a usar e estou melhorando, porque é um pouco difícil de fazer também [...] ajuda muito na hora de corrigir.” (9:54 p 1- 2 em Entrevista Final - E2)

Apesar disso, entendemos que os resultados apontam para a preocupação dos participantes quanto à transparência, consistência e atribuição de notas (como dividir a pontuação entre os critérios), mais do que quanto a como julgar as qualidades de uma produção oral e orientar os alunos na compreensão de sua aprendizagem. Ou seja, há maior atenção à planificação da rubrica e informação ao aluno sobre como o professor avaliará/avaliou, do que no construto linguístico (excertos 14 e 15). Isso demonstra ainda uma visão reducionista, embora amplamente difundida em nossa sociedade, de que a nota é a parte central da avaliação, afastando os processos avaliativos dos objetivos essencialmente formativos.

Excerto 15:

[E5] “Acho que talvez os critérios avaliativos seriam um dos mais difíceis assim. [...] Não sei porque, mas, geralmente, não é algo que é muito falado [...] geralmente, a maioria dos meus professores, eu quase nunca vejo dando critérios. [...] Só avalia e dá a nota. [...] Mas acho que é isso: os critérios foram onde eu tive mais dificuldade, assim, porque, realmente: onde que você vai dar a nota, né? Tem professor que faz dando pra cada critério um uma nota específica, né? Mas é difícil saber o que pesa mais. O que que é mais relevante, né? E é, talvez, a parte mais abstrata, que é uma parte que geralmente fica escondida da gente. Pra a gente, enquanto aluno, é a parte que fica mais na sombra. A gente não ouve muito falar.” (6:22 p 5 em Entrevista E5)

Acreditamos que um fator que possa ter contribuído para reforçar essa concepção tenha sido a decisão de trabalhar com provas orais empregadas em exames de larga escala, cuja principal finalidade não é promover a aprendizagem, mas informar para decisões de terceiros agentes. Essa é uma lacuna que consideramos relevante na formação inicial de professores e que deve ser mais aprofundadamente explorada nos cursos de formação e em pesquisas futuras. Compreendendo que os dados coletados pela avaliação são fundamentais para orientar e promover a aprendizagem, ratificamos a imprescindibilidade de mais estudos sobre como podemos incorporar a elaboração de rubricas em cursos de formação para desafiar visões de língua como sistema e contribuir ainda mais para o letramento em avaliação de professores.

CONSIDERAÇÕES FINAIS

Este trabalho buscou observar o processo de construção do letramento em avaliação de professores de inglês em formação inicial. Estamos convencidas de que há resultados positivos para o desenvolvimento do letramento em avaliação em disciplinas formais durante a formação inicial, calcando as bases iniciais para que esse letramento possa continuar se desenvolvendo ao longo da formação e da vida profissional. Ademais, entendemos que os professores em formação inicial constroem seu letramento em avaliação a partir de suas experiências como discentes, e que o contato com as temáticas que envolvem os conhecimentos, as habilidades e os princípios que compõem o letramento têm forte potencial para levar o licenciando a começar a adquirir a necessária perspectiva de professor.

Com base nos dados coletados, identificamos indícios de desenvolvimento de letramento em avaliação, sobretudo no que diz respeito ao amadurecimento no entendimento dos futuros professores em relação à avaliação, passando de uma visão tradicional centrada em testes de verificação para uma compreensão mais ampla e crítica do processo avaliativo. Essa mudança importante se deu a partir das discussões teóricas e práticas na disciplina sobre avaliação, sendo evidenciada pela elaboração de instrumentos avaliativos mais contextualizados e significativos, bem como pelo aumento da reflexão sobre o impacto da avaliação sobre a aprendizagem dos alunos.

Há duas limitações principais nesta investigação. A primeira deve-se aos necessários recortes metodológicos, o que nos leva a apoiar que estudos futuros investiguem o desenvolvimento de outros componentes e categorias do letramento em cursos semelhantes. A segunda limitação deste estudo deve-se ao contexto em que foi realizado, que não inclui práticas avaliativas por parte dos professores em formação inicial, não lhes permitindo experienciar as aplicações de avaliações e as tomadas de decisão, tanto durante a correção quanto para orientar a aprendizagem. Estudos em contextos típicos de estágios supervisionados também são necessários e recomendados.

Investigar, por meio de estudos como este, o impacto de iniciativas para a promoção do letramento em avaliação desde a formação inicial é essencial para que a importância desses espaços seja difundida para além do universo acadêmico e científico. Dessa forma, este estudo reforça a urgência da inclusão da temática nos currículos de

cursos de licenciatura em língua estrangeira, bem como de estudos futuros no tema.

REFERÊNCIAS

ARMSTRONG, M. C. F.; MARÍN, M. I. R.; PATARROYO, L. M. Analysis of assessment instruments used in foreign language teaching. **Íkala Revista de Lenguaje y Cultura**, Medellín, v. 9, n. 15, p. 171-201, 2004. Disponível em: <https://www.redalyc.org/pdf/2550/255025901007.pdf>. Acesso em: 28 out. 2023.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2011.

BLACK, P. Formative assessment: an optimistic but incomplete vision. **Assessment in Education: Principles, Policy & Practice**, v. 22, n. 1, p. 161-177, 2015.

BLACK, P.; WILIAM, D. Classroom assessment and pedagogy. **Assessment in Education: Principles, Policy & Practice**, v. 25, n. 6, p. 1-25, 2018. DOI:10.1080/0969594X.2018.1441807.

BLACK, P.; WILIAM, D. Inside the black box: raising standards through classroom assessment. **Phi Delta Kappan**, v. 92, n. 1, p. 4-92, 2010. DOI: 10.1177/003172171009200119.

BROOKHART, S. M. Appropriate Criteria: Key to Effective Rubrics. *Frontiers in*

BROWN, D.; ABEYWICKRAMA, P. **Language assessment: principles and classroom practices**. 3 ed. [s.l.]: Pearson, 2018.

BRUM-BARRETO, A. **Letramento em avaliação na formação do professor de espanhol: visão de língua e cultura de avaliar no estágio obrigatório**. 2023. Dissertação (Mestrado em Linguística Aplicada) — Universidade de Brasília, Brasília, 2023.

BRUM-BARRETO, A.; BORGES-ALMEIDA, V. Cultura de avaliar e letramento em avaliação de professores de espanhol no estágio obrigatório. **Entretextos**, Londrina, v. 23, n. 4, p. 346–366, 2023. DOI: 10.5433/1519-5392.2023v23n4p346-366. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/entretextos/article/view/49329>. Acesso em: 25 fev. 2024.

CARLESS, D. Learning-oriented assessment: conceptual bases and practical implications. **Innovations in Education and Teaching International**, [s.l.], v. 44, n. 1, p. 57-66, 2007. DOI: <https://doi.org/10.1080/14703290601081332>. Acesso em: 04 nov. 2023.

DAMACENA, M.; MASTRELLA-DE-ANDRADE, M. Práticas de letramento na escola pública como espaço de (re) construção identitária para professores/as de inglês em formação. **Gláuks-Revista de Letras e Artes**, [s. l.], v. 18, n. 1, p. 221-241, jan./jun. 2018.

DAMACENA-DUTRA, M. A construção do letramento em avaliação na formação inicial de professores de inglês: avaliação de produção oral em foco. 2024. 180 f., il. Dissertação (Mestrado em Linguística Aplicada) — Universidade de Brasília, Brasília, 2023. Disponível em: <http://repositorio.unb.br/jspui/handle/10482/47361>. Acesso em: 6 mai. 2024.

DAVIES, A. Textbook trends in teaching language testing. **Language Testing**, [s. l.], v. 25, n. 3, p. 327-347, 2008. DOI: <https://doi.org/10.1177/0265532208090156>. Acesso em: 04 jul. 2023.

Education, [s. l.], v. 3, n. 22, p. 1-12, 2018. DOI: 10.3389/feduc.2018.00022. Disponível em:

FALTIS, C. Case study methods in researching language and education. In: HORNBERGER, N.; CORSON, D. (ed.). In: HORNBERGER, N.H.; CORSON, D. (ed.). **Research Methods in Language and Education. Encyclopedia of Language and Education**. Dordrecht: Springer, 1997. p. 145-152. Volume 8.

FERNANDES, M. **Letramento em avaliação de professores em formação inicial em um curso de Letras Espanhol: uma pesquisa-ação**, 2019. Dissertação (Mestrado em Linguística Aplicada) — Universidade de Brasília, Brasília, 2019.

FERNANDES, M.; BORGES-ALMEIDA, V. **Letramento em avaliação na formação inicial de professores**. 1. ed. Curitiba: CRV, 2021.

FULCHER, G. Assessment literacy for the language classroom. **Language Assessment Quarterly**, [s.l.] v. 9, p. 113-132, mai. 2012. DOI: <https://doi.org/10.1080/15434303.2011.642041>. Acesso em: 17 nov. 2023.

GIRALDO, F. Language assessment literacy: implications for language teachers. **Profile Issues in Teachers Professional Development**, Bogotá, v. 20, n. 1, p. 179-195, jan./jun. 2018. DOI: <https://doi.org/10.15446/profile.v20n1.62089>. Disponível em: http://www.scielo.org.co/scielo.php?pid=S1657-07902018000100179&script=sci_arttext&tlng=en. Acesso em: 17 no. 2023.

GIRALDO, F.; MURCIA, D. Language assessment literacy and the professional development of pre-service language teachers. **Colombian Applied Linguistics Journal**, [s. l.], v. 21, n. 2, p. 243-259, 2019. DOI: <https://doi.org/10.14483/22487085.14514>. Disponível em: <http://www.scielo.org.co/pdf/calj/v21n2/0123-4641-calj-21-02-243.pdf>. Acesso em: 17 nov. 2023.

GIRALDO, F.; MURCIA, D. Language assessment literacy for pre-service teachers: course expectations from different stakeholders. **GiST Education and Learning Research Journal**, [s. l.], [s.v.], n. 16, p. 56-77, jan./jun. 2018. Disponível em: <https://files.eric.ed.gov/fulltext/EJ1184916.pdf>. Acesso em 17 nov. 2023.

<https://www.frontiersin.org/articles/10.3389/feduc.2018.00022/full>. Acesso em: 5 mar. 2024.

<https://www.tandfonline.com/doi/epdf/10.1080/15434303.2019.1674855?needAccess=t>

role=button. Acesso em: 04 jul. 2023.

INBAR-LOURIE, O. Constructing a language assessment knowledge base: a focus on language assessment courses. **Language Testing**, [s. l.], v. 25, n. 3, p. 385-402, 2008. DOI: <https://doi.org/10.1177/0265532208090158>. Acesso em: 17 nov. 2023.

INBAR-LOURIE, O. Guest editorial to the special issue on language assessment literacy. **Language Testing**, [s. l.], v. 30, n. 3, p. 301-307, 2013. DOI: <https://doi.org/10.1177/0265532213480126>. Acesso em: 04 jul. 2023.

JARAMILLO DELGADO, J.; GIL BEDOYA, A. M. Pre-service English language teachers' use of reflective journals in an assessment and testing course. **JSR Funlam Journal of Students' Research**, n. 4, p. 210-218, 2019. DOI: [10.21501/25007858.3010](https://doi.org/10.21501/25007858.3010). Disponível em: <https://revistas.ucatolicaluisamigo.edu.co/index.php/JSR/article/view/3010/2655>. Acesso em 07 mai. 2024.

JORDÃO, C.; BÜHRER, E. A condição de aluno-professor de língua inglesa em discussão: estágio, identidade e agência. **Educação & Realidade**, Porto Alegre, v. 38, n. 2, p. 669-682, abri./jun. 2013. Disponível em: <https://www.scielo.br/j/edreal/a/HyfVGJp3yZBS7fD4XrLQw7M/?format=pdf&lang=pt>. Acesso em: 17 nov. 2023.

KREMMEL, B.; HARDING, L. Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: developing the language assessment literacy survey. **Language Assessment Quarterly**, [s. l.], v. 17, n. 1, p. 100-120, 2020. DOI: <https://doi.org/10.1080/15434303.2019.1674855>. Disponível em:

MESSICK, S. Validity and washback in language testing. **Language Testing**, [s.l.], v. 13, n. 3, p. 241-256, 1996. DOI: <https://doi.org/10.1177/026553229601300302>. Acesso em 6 mai. 2024.

PANADERO, E.; JONSSON, A. The use of scoring rubrics for formative assessment purposes revisited: a review. **Educational Research Review**, [s. l.], v. 9, p. 129-144, 2013. DOI: <https://doi.org/10.1016/j.edurev.2013.01.002>. Acesso em: 17 nov. 2023.

QUEVEDO-CAMARGO, G. Formação de professores de línguas adicionais e letramento em avaliação: breve panorama e desafios para os cursos de licenciatura em LEM no Brasil. **Calidoscópio**, v. 18, n. 2, 2020.

QUEVEDO-CAMARGO, G.; SCARAMUCCI, M. V. R. O conceito de letramento em avaliação de línguas: origem de relevância para o contexto brasileiro. **LING. – Est. e Pesq.**, Catalão, GO, v. 22, n. 1, p. 225-245, jan./jun. 2018. DOI: <https://doi.org/10.5216/lep.v22i1.54474>. Acesso em: 17 nov. 2023.

RESTREPO BOLÍVAR, E. M. Monitoring preservice teachers' language assessment literacy development through journal writing. **Malaysian Journal of ELT Research**, v. 17, n. 1, p. 38-52, 2020. Disponível em:

https://melta.org.my/journals/MAJER/downloads/majer17_01_03.pdf. Acesso em: 07 mai. 2024.

ROLIM, A. C. **A cultura de avaliar de professoras de LE (inglês) no contexto da escola pública**, 1998. Dissertação (Mestrado em Linguística Aplicada) - Instituto de Estudos da Linguagem, UNICAMP, 1998.

SCARAMUCCI, M. V. R. Letramento em avaliação (em contexto de línguas): contribuições para a linguística aplicada, educação e sociedade. *In*: JORDÃO, C. M. (org.). **A Linguística Aplicada no Brasil: rumos e passagens**. Campinas: Pontes, 2016. p. 141-166.

SCARAMUCCI, M. V. R. Validade e consequências sociais das avaliações em contextos de ensino de línguas. *Linguarum Arena*, v. 2, p. 103-120, 2011.

STIGGINS, R. J. Assessment literacy. **Phi Delta Kappan**, [s. l.], v. 72, n. 7, p. 354-539, mar. 1991.

TAYLOR, L. Developing assessment literacy. **Annual Review of Applied Linguistics**, [s. l.], v. 29, p. 21-36, 2009. DOI: <https://doi.org/10.1017/S0267190509090035>. Acesso em: 17 nov. 2023.

UNIVERSIDADE DE BRASÍLIA. *Projeto Pedagógico do Curso de Licenciatura em Língua Inglesa e Respectiva Literatura da Universidade de Brasília*. Brasília, 2019.

VOGT, K.; TSAGARI, D. Assessment literacy of foreign language teachers: findings of a European study. **Language Assessment Quarterly**, [s. l.], v. 11, n. 4, p. 374-402, 2014. DOI: <https://doi.org/10.1080/15434303.2014.960046>. Acesso em: 04 jul. 2023.

Recebido em: Fev. 2024.

Aceito em: Abr. 2024.

ANEXO
Tarefa inicial de E4***Como será feita a aplicação?***

“Primeiramente separarei a turma em 5 grupos de 4 alunos cada, aonde cada um irá participar de uma encenação, roleplay. Para cada grupo, escolherei uma situação aonde os alunos deverão ter que interpretar outras pessoas nessas situações. Os grupos receberão roteiros apenas com atos que devem acontecer na encenação sem a parte dos diálogos, aonde eles deverão tentar se comunicar para que esses atos aconteçam. Terminado o primeiro grupo de se apresentar o próximo será chamado e assim sucessivamente até que todos os grupos tenham apresentado.”

Fonte: Damacena-Dutra (2023).