

AUTOAVALIAÇÃO E INSTRUMENTOS DE AVALIAÇÃO DE RENDIMENTO NO ENSINO DE LÍNGUA INGLESA

SELF-ASSESSMENT AND PERFORMANCE ASSESSMENT INSTRUMENTS IN ENGLISH LANGUAGE TEACHING

Eliana Kobayashi¹
Sandra Regina Buttros Gattolin²

Data de recebimento do texto: 19/06/2026

Data de aceite: 15/07/2026

Resumo: Este estudo de caso, desenvolvido em uma sala do ensino médio de uma instituição pública na disciplina de língua inglesa, traz os resultados sobre a utilização da autoavaliação voltada ao desempenho (Bachman, 2000; Haughton, Dickinson, 1988; Oscarson, 1989). O objetivo da investigação foi identificar a existência de similaridades entre as notas da autoavaliação e as notas obtidas por meio de outros instrumentos formais de avaliação. Para tanto, os dados foram tratados de modo estatístico por meio da utilização do *T-Test* para amostras pareadas. Os resultados revelam que não há similaridades entre os resultados da autoavaliação e os dois instrumentos avaliativos, além disso, uma análise adicional revelou que participantes com médias mais baixas tendem a se autoavaliar com notas mais elevadas, ao passo que o oposto foi registrado junto àqueles com médias mais elevadas. As conclusões deste estudo corroboram resultados de pesquisas na área que apontam para a necessidade de cautela na utilização desse tipo de autoavaliação.

Palavras-chaves: Autoavaliação. Ensino De Língua Inglesa. Avaliação.

Abstract: This case study, conducted in a public high school English language classroom, presents findings on the use of performance-based self-assessment (Bachman, 2000; Haughton & Dickinson, 1988; Oscarson, 1989). The aim of the study was to identify whether similarities existed between students' self-assessment scores and the scores obtained through other formal assessment instruments. To this end, the data were analyzed statistically using the paired-samples t-test. The results indicate that there were no significant similarities between the self-assessment scores and those obtained from the two assessment instruments. Furthermore, an additional analysis revealed that participants with lower average grades tended to assign themselves higher self-assessment scores, whereas the opposite pattern was observed among those with higher average grades. The findings of this study corroborate previous research in the field, highlighting the need for caution when using this type of performance-based self-assessment.

Keywords: Self-Assessment. English Teaching. Assessment.

¹ Doutora em Linguística Aplicada. Docente do Instituto Federal de São Paulo - IFSP.

² Doutora em Linguística Aplicada. Docente da Universidade Federal de São Carlos - IFSP.

1 Introdução

A busca pela autonomia do aluno tem sido muito discutida no âmbito da aprendizagem (HOLEC, 1981; TUDOR, 1997; OXFORD, 1990, NICOLAIDES, 2003; LEFFA, 2016) e as exigências do mundo globalizado reforçam ainda mais essa necessidade aliada à questão da formação de um cidadão crítico capaz de agir em sociedade, o que pressupõe o domínio da língua (LANKSHEAR, 1997). A Base Nacional Comum Curricular (BNCC) do ensino médio (BRASIL, 2017, p.473) justifica que “por ser um período de vida dos alunos caracterizado por mais autonomia e maior capacidade de abstração e reflexão sobre o mundo, os jovens, gradativamente, ampliam também suas possibilidades de participação na vida pública e na produção cultural.”

Além disso, conforme ressalta Kalantzis (2006), os jovens atualmente mudaram suas posturas e não se contentam com a passividade que antes lhes era usual, preferem ser atores a audiência, montam suas playlists em seus *ipods* ao invés de seguirem uma programação da rádio. A geração dos games está acostumada a dividir a autoria das narrativas e poder influenciar o fim de uma história (GEE, 2005).

Diante disso, torna-se incongruente visar à autonomia da aprendizagem do aluno sem possibilitar a discussão sobre a utilização de instrumentos de avaliação que proporcionem também uma atuação mais autônoma. A autoavaliação, como qualquer instrumento de avaliação, também pode ser utilizada com objetivos diversos podendo, por exemplo, restringir-se ao âmbito de análise do próprio conhecimento e habilidades (LE BLANC, PAINCHAUD, 1985) ou como uma forma de se otimizar tempo e custo para se identificar o nível de proficiência do avaliado (STRONG-KLAUSE, 2000).

Por outro lado, não seria adequado aqui conceber qualquer processo avaliativo como uma ação simples, pelo contrário, pois, apesar de a avaliação em língua inglesa estar presente em diversos setores da sociedade, do educacional ao corporativo, não é possível afirmar que a sua popularidade tenha aumentado a validade ou a confiabilidade dos testes, conceitos estes complexos de serem identificados (MESSICK, 1989; CHAPELLE, ENRICH, JAMIESON, 2004; BACHMAN, 2005). Pelo contrário, estudos apontam diversos questionamentos e problemas nos testes de inglês aplicados em ambientes empresarial e educacional (DIAS, 1995; DUARTE, 2009; KOBAYASHI, 2010), demonstrando a complexidade do ato de avaliar (STOBART, 2003).

Diante disso, reafirma-se a necessidade e relevância de investigações na área de avaliação para a melhor compreensão desse processo. Somados a isso, estão as exigências para a formação de um cidadão capaz de atuar em um mundo cada vez mais sem fronteiras onde o inglês já se solidificou como a língua de uso mundial pela variedade de usuários, de utilizações que apresenta e da predominância nas principais áreas da sociedade (BRASIL, 2017; GRADDOL, 2002).

Este estudo traz uma proposta distinta dos estudos de autoavaliação encontrados no Brasil e em outros países. Observa-se que as pesquisas tendem a focalizar o uso desse instrumento de avaliação em um desses dois espectros: a autoavaliação voltada ao desempenho e a autoavaliação voltada ao desenvolvimento. O primeiro espectro está relacionado ao processo de aprendizagem, nos quais se enquadram também os estudos ligados à autoavaliação e autonomia, o segundo, àquelas investigações que correlacionam a autoavaliação com outros instrumentos formais de avaliação, como testes de proficiência, avaliação dos professores ou de pares. Segundo Haughton e Dickinson (1988), Oskarson (1989) e Bachman (2000), o objetivo da autoavaliação voltada ao desempenho é avaliar, em um determinado momento, o desempenho do candidato, sendo, por isso, voltado a processos seletivos, de nivelamento, de diagnósticos e certificações. Por outro lado, a autoavaliação voltada ao desenvolvimento acompanha os participantes no processo de aprendizagem visando à identificação de padrões de progresso e mudanças.

Assim, na autoavaliação voltada ao desempenho, o aluno ou candidato avalia o seu nível de inglês ou outra língua estrangeira de acordo com critérios estabelecidos, atribuindo notas a si mesmo (LEE, 2011), ou classificando seu desempenho de acordo com escalas ou parâmetros (TAVAKOLI, 2010). A nota, neste estudo, é utilizada por razões similares a de Lee (2011, p.96) em sua investigação sobre a autoavaliação em um programa de inglês e coreano: “notas foram escolhidas especificamente para identificar a relação entre as notas dos professores e as notas da autoavaliação dos alunos. A comparação de notas permite verificar se os alunos forma capazes de atribuir notas similar a dos professores”. Somado a isso, neste trabalho, a nota da autoavaliação visa promover a autonomia do aluno que passa a participar do seu processo avaliativo, antes totalmente centralizado no professor.

Considerando essas discussões, este artigo apresenta um estudo sobre autoavaliação realizado na disciplina de língua inglesa em uma sala de ensino médio de uma instituição pública, cuja pergunta de pesquisa questiona a existência de similaridades entre as notas de autoavaliação e as notas obtidas por meio de instrumentos formais de avaliação. Esta

investigação visa apresentar contribuições sobre a utilização da autoavaliação em sala de aula e promover reflexões sobre o processo de avaliação em língua inglesa.

Este artigo está dividido em cinco partes. Na primeira, os objetivos do estudo, assim como as perguntas de pesquisa em uma breve introdução são apresentados. Em seguida, são discutidos os principais aportes teóricos que embasam a investigação. A natureza metodológica da pesquisa, descrições do contexto e instrumentos de coleta são abordados na terceira parte do artigo, seguidas pela discussão dos resultados. O artigo é encerrado com as considerações finais.

2 Fundamentação teórica

A avaliação em língua inglesa, assim como em outras áreas, envolve a aplicação de instrumentos que possam analisar as habilidades de um candidato segundo critérios pré-estabelecidos. Apesar da variedade de instrumentos, tais como testes, apresentações, trabalhos, portfólios, dinâmicas entre outros, o teste³ ainda predomina nas mais diversas áreas. Um exemplo disso são os exames internacionais de proficiência, cada vez mais utilizados no Brasil, não somente no contexto universitário, mas no ensino regular de escolas particulares (SCARAMUCCI, KOBAYASHI, 2013; KOBAYASHI, RIBEIRO DA SILVA, 2019).

De acordo com Carroll (1986, p. 46), teste é “um procedimento desenhado para elicitare um determinado comportamento do qual é possível fazer inferências sobre certas características de um indivíduo”. Há vários tipos de testes, conforme os objetivos a serem alcançados, como: de proficiência, de rendimento, de nivelamento e de diagnóstico. Quando o propósito consiste em avaliar de acordo com especificações sobre o que o candidato deve saber executar para ser considerado proficiente, independentemente de ter realizado qualquer curso, o teste é classificado como de proficiência. Por outro lado, o teste de rendimento, normalmente elaborado por professores, busca avaliar se os objetivos de aprendizado de um curso estão sendo alcançados. O teste de diagnóstico visa identificar os pontos fracos e fortes do aluno em relação à língua e, assim, possibilitar o direcionamento do aprendizado. Por fim, o teste de nivelamento é utilizado para alocar o aluno no estágio adequado de um

³ Neste artigo, os termos teste, prova e exame são utilizados como sinônimos.

programa de ensino, tendo, assim, o objetivo de identificar o seu nível de proficiência (HUGHES, 1989).

No contexto educacional, o tipo de teste mais regularmente aplicado é o de rendimento, utilizado na composição de nota do aluno como indicativo de seu aproveitamento na disciplina. Apesar de estar relacionado ao passado, como ressaltou Anastasi (1976, p.398), nas discussões iniciais sobre teste de rendimento, ao afirmar que “são elaborados para medir os efeitos de um programa específico de treinamento ou ensino”, o trabalho trazido por Hughes (1988) aproxima testes de proficiência e de rendimento.

Primeiramente, o autor destaca que ter sucesso em testes de rendimento não significa necessariamente a aquisição de uma determinada proficiência. Isso pode ocorrer na medida em que um curso, no qual o teste de rendimento está baseado, pode não ter sido elaborado de forma adequada e, como consequência, a sua avaliação não reflete a proficiência pretendida. Como exemplo, Hughes (1988, p. 36) cita “alguém que obteve pontuação alta em um teste de rendimento de um curso de Francês para Guias que tenha sido mal elaborado pode ainda não ter francês suficiente (ou seja, ser proficiente em francês o suficiente) para atuar com sucesso como um guia.”

Dessa forma, Hughes (1988) aponta que o problema está no conteúdo dos testes de proficiência e de rendimento visto que ambos deveriam ser de acordo com o que é esperado que o candidato ou o aluno seja capaz de realizar na língua. Assim, o autor enfatiza que o teste de rendimento não deve ser baseado no conteúdo ou no programa do curso, mas nos objetivos que, por sua vez, dão origem ao programa implementado no curso, não o inverso. Do mesmo modo, o teste de proficiência é construído de acordo com determinadas expectativas de uso da língua, ou seja, o que se espera que o candidato seja capaz de desempenhar em inglês.

O autor traz a distinção entre dois tipos de testes de rendimento: o de progresso e o final. O primeiro seria aplicado ao longo do curso para acompanhar o progresso do aluno e o final, como o termo indica, no encerramento do curso. Portanto, o autor conclui que, quando programas de cursos são baseados em objetivos de uso da língua os quais delineiam a elaboração do teste de rendimento aplicado no fim do curso, não há distinções entre testes de proficiência e rendimento final.

Desse modo, no contexto da sala de aula de ensino de língua inglesa, o processo avaliativo poderia ser mais efetivo com a utilização, entre os vários instrumentos possíveis, do teste de rendimento, conforme discussão de Hughes (1988,1989), ou seja, voltado aos

objetivos de uso da língua. Além disso, levanta-se nesta investigação a possibilidade de que a autoavaliação possa funcionar como um instrumento avaliativo.

A autoavaliação no contexto educacional tem sido discutida há algum tempo, em seu artigo de 1985, LeBlanc e Paichaud definiram autoavaliação como procedimentos usados para os alunos analisarem seus conhecimentos e habilidades. Para Noonan, Duncan (2005) e Ross (2008), trata-se da avaliação do conhecimento, habilidades e desempenho da própria pessoa.

Há também outros aspectos no escopo da autoavaliação, como o estabelecimento de critérios e a análise da aprendizagem diante desses critérios, sendo elas realizadas pelos próprios alunos (HARRIS, BELL, 1994). Somado a isso, esse instrumento de avaliação possibilita a identificação individual de pontos fracos e fortes na aprendizagem do aluno, o que contribui para a sua melhora (KLENOWSKI, 1995).

Pesquisas revelam que o uso da autoavaliação pode beneficiar alunos e professores. Por exemplo, podem auxiliar o aluno a compreender os propósitos de certas atividades e assumir mais responsabilidade pelo seu aprendizado (ORSMOND, MERRY, REILLING, 1997); colaborar na formação de alunos mais ativos e focados, e mais bem preparados para avaliar seus progressos em termos de comunicação (HARRIS, 1995); elevar o empoderamento do aluno por meio de um envolvimento mais intenso com a avaliação (TAN, 2008).

Há também resultados positivos em termos de aprendizado de habilidades específicas, como no estudo sobre a autoavaliação da produção escrita de Birjandi, Tamjid (2012) e também no desenvolvimento da fluência oral por meio da aplicação da autoavaliação da própria produção gravada, anotações e registros de aprendizado dos alunos, contribuindo para o alcance de metas, identificação de pontos fortes e fracos (MICAN, MEDINA, 2015).

Klenowski (1995) ressalta que os benefícios da autoavaliação são mais prováveis ocorrerem diante de três condições: a negociação entre professor e alunos quanto aos critérios, o diálogo entre professor e alunos focalizar nas evidências para a avaliação, e a autoavaliação contribuir para a nota do aluno, seja pelo estudante somente ou em colaboração com os professores.

Em relação aos propósitos da autoavaliação em cursos de línguas, a autoavaliação voltada ao desempenho, foco deste estudo, apresenta vários aspectos que favorecem a sua utilização, como a eficiência em termos de custo e tempo (STRONG-KLAUSE, 2000) e o

fato de não requerer atenção quanto a problemas de segurança uma vez que o candidato não vai trapacear no teste (LE BLANC e PAINCHAUD, 1985). No entanto, há questionamentos sobre a credibilidade desse tipo de autoavaliação, principalmente quando o seu resultado pode causar consequências sérias na vida do candidato, o que gera hesitação em seu uso por parte dos administradores de testes, conforme afirma Saito (2003). Portanto, o autor recomenda cuidado na interpretação dos resultados, caso o propósito seja nivelamento, certificação, diagnóstico e admissão.

A cautela enfatizada por Saito (2003) encontra-se em consonância com estudos comparativos entre a autoavaliação realizada pelo aluno e a avaliação feita pelo professor ou especialistas diante dos resultados diversos. Por exemplo, Kuncel, Credé e Thomas (2005) identificaram que alunos com desempenho alto declaram notas obtidas na faculdade e no ensino médio precisamente, no entanto, a precisão diminui entre os alunos com baixo desempenho, possivelmente por fatores relacionados à imagem social.

Em um estudo comparativo, Blatchford (1997) confrontou autoavaliações e testes padronizados, identificando uma relação mediada pela idade, na medida em que as autoavaliações correlacionavam-se com os testes em estudantes com 16 anos, mas isso não ocorria com alunos de sete. Além disso, conforme Saint Léger (2009), a confiabilidade da autoavaliação é medida correlacionando a nota de desempenho autoavaliada à nota aferida por um avaliador ou por uma referência externa.

A investigação realizada por David e Henning (1985) indica, baseada em instrumentos de autoclassificação, que os participantes demonstravam tendências de exagerar suas capacidades pessoais. Para Dickison (1987), há outros atores, como professores e especialistas, que poderiam realizar avaliações mais confiáveis, precisas e com maior regularidade do que os próprios alunos. Em um estudo comparativo de autoavaliação em segunda língua em quatro aspectos: leitura, produção oral, compreensão oral e produção escrita, e testes de rendimento que focalizavam o conteúdo do curso, Ross (1998) observou relação mais forte entre a autoavaliação e o teste de leitura e mais baixa entre a autoavaliação e o teste de produção oral.

Em seu estudo sobre um curso de imersão em francês no Canadá, Pierce *et al* (1993) também não identificaram correlação entre a autoavaliação e os instrumentos oficiais de avaliação. A autoavaliação era baseada na comparação realizada pelos alunos com falantes nativos e a reflexão dos próprios estudantes sobre as dificuldades que enfrentavam com a

língua no dia a dia. Esse resultado foi então correlacionado com os resultados obtidos nos testes de proficiência que avaliavam as quatro habilidades.

Blue (1994) aponta para as possíveis consequências da autoavaliação na investigação que analisou a autoavaliação utilizada no programa de ensino de inglês para propósitos acadêmicos. Os resultados da autoavaliação realizada antes e durante o curso não apontaram correlações com as avaliações do professor. Além disso, as conclusões do estudo apontam que os alunos que autoavaliam o seu nível de inglês de uma maneira mais realística têm maiores chances de perseverar no aprendizado do curso do que aqueles cujas autoavaliações foram irrealisticamente baixas ou altas.

Por outro lado, também há estudos que demonstram que a autoavaliação pode ser considerada confiável e válida para avaliar habilidades comunicativas, como a investigação de Bachman e Palmer (1989), na qual as competências gramatical, pragmática e sócio-linguística de 116 falantes não nativos de inglês foram analisadas. Os resultados, segundo os autores, indicam nível de confiabilidade superior ao esperado, com a avaliação da competência gramatical apresentando melhores indicações sobre esse aspecto do que as avaliações das demais competências.

O trabalho de Tavakoli (2010) investigou a relação entre a autoavaliação dos alunos e a avaliação do professor como uma alternativa à prova de língua em uma sala de aula. Trinta e cinco alunos veteranos de literatura inglesa inscritos em um curso de inglês participaram do estudo que envolvia a análise da habilidade de produção oral. Os resultados indicaram correlação moderada entre as avaliações, levando o autor a concluir que a autoavaliação consiste em uma estratégia que pode colaborar para a autonomia, o automonitoramento e o autodirecionamento da aprendizagem.

O estudo de Gholami (2016), por sua vez, investigou o impacto de autoavaliação tanto na autonomia do aluno quanto na proficiência na língua, por meio da utilização de um questionário multidimensional sobre autonomia e grupos experimental e controle. Depois de um período de três meses e 25 sessões de uso do instrumento, o autor verificou que alguns aspectos da autonomia haviam melhorado, mas não houve impacto em termos de proficiência.

No Brasil, observa-se que os estudos sobre a autoavaliação estão voltados para a autoavaliação para o desenvolvimento e, portanto, não utilizam comparações da autoavaliação com outros instrumentos formais de avaliação, como diversos estudos realizados em outros países. Por exemplo, o trabalho de Punhagui e Souza (2012) analisou a

uso da autoavaliação no ensino de língua estrangeira na 8ª série visando à autonomia no aprendizado. Os resultados demonstram que os estudantes passaram a reconhecer as suas dificuldades e pontos fortes, desencadeando mudanças na aprendizagem. Cavalari (2009) discutiu as práticas autoavaliativas de uma aluna em um curso no ambiente *teletandem* (TELLES, 2005) no qual era aprendiz de inglês e interagia via chat com dois alunos norte-americanos que aprendiam português. Os resultados revelam que a estudante estabelecia parâmetros elevados para a sua autoavaliação baseados em critérios afetivos, linguístico-comunicativo e no feedback obtido. Fidalgo (2005) discutiu a validação por parte dos professores de instrumentos de avaliação por meio das representações de suas falas e os resultados de seu estudo demonstram a importância da autoavaliação ser parte intrínseca do processo de ensino e aprendizagem, como uma ferramenta que regula a própria aprendizagem do aluno. Além disso, Cato (2010) demonstrou que as percepções dos alunos de língua inglesa de uma universidade pública sobre autoavaliação não são influenciadas por traços da personalidade, motivação e sua história de aprendizagem. Somado a isso, a visão dos alunos sobre a autoavaliação não é considerada equivocada para a autora.

Assim, existe uma carência de investigações que explorem a relação entre a autoavaliação e os testes de proficiência ou rendimento, o que poderia colaborar para a confiabilidade desse instrumento e, como consequência, auxiliar na reflexão quanto aos modos de utilização em sala.

3 Metodologia

Este trabalho caracteriza-se como um estudo de caso, ou seja, uma investigação de um “exemplo em ação”, que possibilita a análise do funcionamento em contexto do exemplo em questão (ADELMAN *et al*, 1976). Além disso, trata-se de um estudo de caso explicativo uma vez que busca responder questões para explicar as supostas relações existentes em uma intervenção real (YIN, 2003).

Os dados são tratados de modo estatístico por meio da utilização do software *IBM SPSS Statistics* para a execução do *T-Test* para amostras pareadas.

Descrição do cenário e dos participantes

Este estudo foi desenvolvido em uma sala do ensino médio de uma instituição pública de ensino localizada no estado de São Paulo. Participaram da investigação, 37 alunos da disciplina de língua inglesa, obrigatória em todos os anos do curso, que tem carga horária de 100 minutos por semana, equivalendo a duas aulas no contexto investigado.

No início do ano letivo, os alunos fizeram o teste padronizado *Straightforward* da McMillan Publishing que contava com 40 questões de gramática e 10 de vocabulário. De acordo com a grade de correção da editora, de 0 a 35 acertos, o aluno deveria ser classificado como *beginner* e, de 36 a 50, como *elementary*, o que equivaleria aos níveis A1 e A2, respectivamente do Quadro Comum Europeu de Referência (*Common European Framework for References*). Embora esse tipo de teste não reflita as habilidades linguísticas necessárias para o uso da língua, o aspecto da praticidade constitui um dos motivos da escolha, visto que as questões eram projetadas na tela e respondidas pelos alunos no caderno.

Os resultados demonstraram que 70% da sala, ou seja, a maioria estava no nível A2, de acordo com a proposta da McMillan. Apesar do foco do teste não ter sido nas habilidades comunicativas, os resultados refletiram a média obtida pelos alunos nas provas de rendimento que avaliavam a habilidade de leitura.

Descrição dos instrumentos de coleta

O questionário foi o principal instrumento de coleta desta pesquisa diante de sua eficácia em coletar dados primários (COHEN, 2013), elaborados de acordo com os momentos da investigação e os objetivos de aplicação. A sua escolha baseou-se principalmente pela vantagem de possibilitar a coleta das informações em campo (NUNAN, 1992),

O questionário utilizado foi baseado na escala Likert, criada em 1932 pelo cientista social Rensis Likert, com o objetivo de avaliar traços de personalidade e caráter diante da dificuldade em transferir essas qualidades para medidas quantitativas. Assim, o cientista desenvolveu um procedimento para avaliar escalas de atitudes por meio de perguntas com cinco alternativas de respostas: aprovo totalmente (1), aprovo (2), indeciso (3), desaprovo (4) e desaprovo totalmente (5). Atualmente, há variações na Likert, algumas com menos respostas, outras com mais, apesar de tipicamente a escala ter um ponto neutro.

No entanto, a vantagem da utilização da escala Likert neste estudo está no fato de os itens oferecerem uma gradação de respostas, o que evita opções binárias que podem dificultar a autoavaliação do respondente. Especificamente sobre esses questionários, os itens foram divididos em dois componentes principais: o primeiro visou medir aspectos relacionados ao comprometimento do aluno com a disciplina, tais como frequência, participação em sala, organização do material e contribuição com o grupo; e o segundo, analisar como os alunos avaliam as metas de uso da língua estabelecidas para o período.

Assim, como forma de participação em sua própria avaliação e considerando que a avaliação no contexto investigado envolve notas, o participante pôde fornecer uma nota de zero a dez, a qual corresponderia a 20% de sua média final no bimestre.

Descrição dos instrumentos de avaliação de rendimento

A composição da média do bimestre é calculada por meio de dois instrumentos de avaliação elaborados e aplicados pela professora, uma apresentação oral em grupo e um teste escrito individual, além da autoavaliação. O primeiro consistiu em um trabalho desenvolvido em grupos de quatro alunos que focalizou o gênero carta do leitor. Tendo como base a mídia impressa de países cuja língua oficial fosse o inglês, os alunos deveriam encontrar uma carta do leitor e apresentar suas ideias principais. Para tanto, também era preciso que trouxessem os motivos que levaram o leitor a redigir tal carta. Por exemplo, se o objetivo era criticar uma reportagem veiculada, o grupo deveria trazer tal texto e explicá-lo também. Essa tarefa envolveu processos de seleção de fontes, leitura, produção escrita dos slides e produção oral para a apresentação.

Os critérios de avaliação da apresentação oral foram explicitados na mesma aula na qual as orientações para a realização dessa tarefa ocorreram conforme mostra o Quadro 1.

Quadro 1 – Critérios avaliativos da apresentação oral

Critério avaliativo	Descrição
Realização da tarefa	Refere-se à execução da tarefa de acordo com os objetivos comunicativos apresentados.
Gramática e vocabulário	Refere-se à utilização dos recursos linguísticos (estrutura e léxico) para a execução da tarefa.
Fluência	Refere-se à capacidade de manter a adequação do fluxo de fala
Pronúncia	Refere-se à expressão de sons, do ritmo e da entoação.

O teste escrito, por outro lado, foi baseado em questões de interpretação textual, em sua maioria, compreensão das características do gênero textual carta do leitor e das seções que normalmente compõem um jornal, produção de uma manchete e subtítulo de uma reportagem, e aspectos gramaticais do texto de leitura. O objetivo era avaliar a capacidade de leitura de textos que descrevem ações e fatos passados, o entendimento de gêneros específicos, a composição de um jornal e as características de uma notícia e compreensão das características do passado simples. Similarmente à apresentação oral, os objetivos avaliativos foram apresentados previamente, focando nas competências esperadas.

4 Resultados

Esta seção está dividida em três partes. Primeiramente, os resultados obtidos no instrumento de autoavaliação são discutidos e, em seguida, são realizadas as comparações entre as notas fornecidas pelos participantes e aquelas atingidas no teste de rendimento. Finalizando, são apresentados os resultados obtidos na comparação entre as notas mais elevadas e mais baixas de um recorte dos participantes e as notas fornecidas na autoavaliação.

Questionário aplicado aos alunos

A escala tipo Likert apresentou onze questões com cinco alternativas de respostas para cada questão: concordo totalmente, concordo, não concordo nem discordo, discordo e discordo totalmente. Em termos cronológicos de aplicação dos instrumentos avaliativos previstos para o bimestre, o questionário foi aplicado depois da apresentação oral e antes do teste escrito. O objetivo foi verificar se os alunos estavam atingindo as metas estabelecidas para o período do curso e, principalmente, proporcionar uma oportunidade de reflexão dos alunos sobre o seu próprio desempenho.

Tabela 1 – Questionário de autoavaliação - alunos

Questões	Discordo totalmente	Discordo	Não concordo, nem discordo	Concordo	Concordo totalmente
1.Frequência igual ou superior ao bimestre anterior	0%	3%	5%	24%	68%
2. Traz material didático completo em sala	3%	3%	16%	35%	43%

3. Comportamento adequado	0%	5%	19%	27%	49%
4. Fala mais em inglês que bimestre anterior	0%	0%	3%	19%	78%
5. Faz as atividades escritas	0%	0%	3%	16%	81%
6. Apresentação oral satisfatória	0%	14%	22%	38%	27%
7. Fez atividade escrita/oral satisfatoriamente	0%	5%	8%	22%	65%
8. Consegue relatar fatos passados	0%	3%	14%	43%	41%
9. Consegue descrever características de jornais	0%	0%	11%	19%	70%
10. Distingue fatos e opiniões em carta ao editor	0%	0%	0%	16%	84%
11. Entende e sabe produzir manchetes	0%	5%	14%	38%	43%

Quanto à frequência nas aulas de inglês, primeira pergunta do questionário, 68% dos alunos concordam totalmente que a assiduidade no bimestre em que o questionário foi aplicado foi igual ou superior ao bimestre anterior, enquanto 24% concordaram, 5% não concordaram nem discordaram, 3% discordaram e nenhum participante discordou totalmente. Dessa forma, os participantes não identificaram problemas, de um modo geral, em suas frequências nas aulas de inglês.

Em relação à organização do material didático, 43% dos participantes concordam totalmente com a afirmação de que entram em sala com o material completo, não saindo para buscar no armário, pois sabem que isso atrapalha tanto o seu aprendizado quanto o andamento da aula. Além disso, 35% dos alunos concordam com a afirmação, 16% não concordam nem discordam, 5% discordam e nenhum discorda totalmente. Dessa forma, os resultados não apontam que a maioria dos alunos assiste todas as aulas com o material completo.

Sobre o comportamento em sala, 81% dos alunos concordam totalmente que é muito bom, sendo a evidência estabelecida na pergunta o fato de chegarem pontualmente na sala, respeitar a professora e os colegas, e cooperarem com o bom andamento da aula. Um índice bem inferior, 16%, revela que os participantes concordam com a afirmação e 3% não concordam, nem discordam. Na ótica dos alunos, o comportamento, definido na questão, parece ser muito bom para ampla maioria.

Em relação aos aspectos ligados diretamente às habilidades linguísticas, 43% dos alunos concordam totalmente que a sua comunicação oral em inglês na sala, tanto com

colegas quanto com a professora, está superior ao bimestre anterior, ou seja, estão utilizando mais a língua para se comunicar oralmente, seguido por 35% que concordam e 16% que não concordam nem discordam. Por outro lado, 3% discordam da afirmação e também 3% discordam totalmente. Assim, a percepção da maioria, 80% dos participantes, demonstra evolução, pelo menos em termos de frequência de uso da habilidade oral em sala.

Sobre a produção escrita, 78% dos participantes concordam totalmente que estão realizando as atividades escritas propostas em sala melhor do que no bimestre anterior, enquanto que 19% concordam e 3% não concordam nem discordam. Os índices da concordância dos alunos com a questão estão superiores aos apresentados pelos alunos sobre a habilidade de produção oral. A possível justificativa para tanto pode ser atribuída à possibilidade de os participantes terem incluído as lições de casa, os exercícios no livro didático e em websites recomendados.

Especificamente sobre a apresentação oral, realizada como parte da avaliação de rendimento do bimestre, 27% dos participantes concordam totalmente que o desempenho foi satisfatório, de acordo com os critérios apresentados pela professora (Quadro 1), 38% concordam, 22% não concordam nem discordam, e 14% discordam. É possível identificar que os resultados são distintos entre produção oral não avaliativa, caracterizada por atividades rotineiras realizadas em sala, e a avaliativa, que envolve trabalho em grupo e fora da sala também.

Questionados especificamente sobre uma atividade que envolvia habilidades de escrita e oralidade realizada individualmente em casa apresentada oralmente em sala, a maioria dos alunos, 65%, concordam totalmente que conseguiram atingir o objetivo do trabalho, ao passo que 8% não concordaram nem discordaram e 5% discordaram. Apesar da similaridade dessa atividade à apresentação oral avaliativa e a utilização dos mesmos critérios avaliativos como parâmetros para a execução adequada da tarefa, os resultados foram consideravelmente distintos, reforçando as inferências sobre as distinções entre atividade avaliativa e atividade não avaliativa.

Em seguida, 41% dos participantes afirmam que concordam totalmente que conseguem relatar fatos que aconteceram no passado e situações ocorridas momentaneamente também no passado, ao passo que 43% concordam e 14% não concordam nem discordam. Uma minoria, 3%, discorda que consegue realizar esse uso da língua. Essa questão está relacionada às atividades propostas nas perguntas sobre a apresentação oral e a atividade realizada em casa, pois ambas exigiam essa capacidade. No entanto, os resultados

são significativamente distintos. Uma das hipóteses pode estar no fato da pergunta não especificar as habilidades linguísticas, como produção oral e escrita, diferentemente das duas questões anteriores, o que pode ter remetido somente aos exercícios de gramática do Passado Simples e Contínuo, ou seja, a dissociação entre oralidade e gramática ou produção escrita e gramática pode ter ocorrido.

Sobre as características de um artigo jornalístico e dos meios de comunicação jornal e revista, 70% dos participantes concordam totalmente que sabem descrever a categoria e seções dessas mídias, assim como o título, subtítulo e *lead* de um artigo. Por outro lado, 11% afirmam que não concordam nem discordam e 19% concordam. Essa questão abrange o conteúdo trabalhado durante as aulas que exploravam os gêneros textuais jornalísticos. Os resultados demonstram que a maioria dos alunos consegue reconhecer as características de artigo jornalístico, jornal e revista.

Ainda sobre mídia escrita, a maioria dos alunos, 84%, concorda totalmente que sabe diferenciar fatos e opiniões em um texto do gênero carta do leitor, enquanto 16% concorda. Essa questão foi a única em que não houve nem neutralidade na gradação de resposta (não concordo, nem discordo) nem discordância. Com índices de concordância superiores ao da questão anterior sobre o artigo jornalístico, identifica-se a confiança dos participantes quando as questões envolvem reconhecimento dos aspectos dos gêneros textuais trabalhados.

Por outro lado, quando questionados sobre a compreensão das características de uma manchete de notícias e a capacidade de produzir esse gênero, 43% dos alunos concordam totalmente que são capazes, 38% concordam, 14% optaram pela neutralidade e 5% afirmam que não concordam. Esses índices caem significativamente em toda a escala quando são comparadas esta questão com as duas perguntas anteriores, principalmente na resposta “concordo totalmente”.

De uma forma geral, as opiniões e percepções dos alunos revelam desenvolvimento satisfatório tanto das quatro habilidades linguísticas, conforme os objetivos de ensino estabelecidos para o período, quanto às atividades e comportamentos esperados para o acompanhamento das aulas. No entanto, a oralidade, principalmente na atividade avaliativa, surge como a habilidade que os alunos apresentam menos certeza de desempenho satisfatório.

Relação entre a autoavaliação e os instrumentos avaliativos de rendimento

O questionário da autoavaliação teve o objetivo de despertar reflexões do aluno quanto a sua atuação na disciplina de inglês. A nota de 1 a 10 que o participante determinava foi comparada com os outros dois instrumentos de avaliação que integrariam a média final do bimestre, o teste escrito e a apresentação oral em grupo.

A primeira avaliação do bimestre é a apresentação oral, seguida da autoavaliação e finalizando com o teste escrito. No momento em que os alunos se autoavaliam, estão cientes apenas da nota que obtiveram no primeiro instrumento avaliativo, para evitar influências na obtenção da média final.

Assim, primeiramente foram comparados as notas obtidas no teste escrito e a autoavaliação. Para tanto, foi utilizado o software *IBM SPSS Statistics* para a execução do *T-Test* para amostras pareadas, uma vez que são comparadas a mesma população, mas com dados distintos.

Os resultados demonstram que a média do teste escrito ($M=6,58$, $EP=0,3431$) é menor do que a média da autoavaliação ($M=8,662$, $EP=0,1097$), $t(36)=-6,125$, $p<0,05$. Dessa forma, foram identificadas evidências de diferenças significativas entre as notas do teste escrito e as notas da autoavaliação. A Tabela 2 mostra a comparação entre as notas da autoavaliação e do teste escrito.

Tabela 2 – Teste escrito e Autoavaliação

Teste de amostras pareadas

	Média	N	Desvio Padrão	Erro padrão da média
Par 1 Teste escrito	6,581	37	2,0867	,3431
Autoavaliação	8,662	37	,6672	,1097

	Diferenças pareadas							
	Média	Desvio Padrão	Erro padrão na média	95% Intervalo de confiança da diferença Inferior	95% Intervalo de confiança da diferença Superior	t	df	Sig (2extremidades)
Par 1 Testeescrito & autoavaliação	-2,0811	2,0666	,3398	-2,7701	-1,3920	-6,125	36	,000

Na comparação entre as notas da autoavaliação e da apresentação oral, os resultados também apontam para diferenças. A média do teste oral ($M=7,622$, $EP=0,1614$) é menor do que a média da autoavaliação ($M=8,662$, $EP=0,1097$), $t(36)=-5,882$, $p<0,05$. Tal resultado

revela evidências de diferenças significativas entre as notas da apresentação oral e as notas da autoavaliação.

Tabela 3 - Teste oral e Autoavaliação

Teste de amostras pareadas

	Média	N	Desvio Padrão	Erro padrão da média
Par 1 Apresentação oral	7,622	37	,9818	,1614
Autoavaliação	8,662	37	,6672	,1097

	Diferenças pareadas							
	Média	Desvio Padrão	Erro padrão na média	95% Intervalo de confiança da diferença	95% Intervalo de confiança da diferença	t	df	Sig (2extremidades)
Par 1 Apresentação oral & autoavaliação	-1,0405	1,0761	,1769	-1,3993	-,6817	-5,882	36	,000

Esses resultados evidenciaram as diferenças entre os instrumentos de avaliação de rendimentos e a autoavaliação, sendo que as notas que os participantes atribuem a si mesmos tendem a serem superiores. Diante disso, foi realizada uma análise mais detalhada das notas em seus extremos, ou seja, as mais baixas e as mais altas obtidas nesses mesmos instrumentos de avaliação de rendimento.

Comparação entre notas nas extremidades e a autoavaliação

Primeiramente, foi realizado um recorte do público investigado, isolando os cinco alunos que atingiram as notas mais elevadas e os cinco que apresentaram as notas mais baixas na média obtida entre o teste escrito e a apresentação oral. Em seguida, tais notas foram comparadas com as notas de autoavaliação, conforme Tabela 4.

Tabela 4: Notas nas extremidades x notas da autoavaliação

Alunos que atingiram notas mais baixas

Aluno	Prova escrita	Apresentação oral	Autoavaliação
Aluno 5	3,0	6,5	9,0
Aluno 7	5,5	7,5	8,0
Aluno 10	5,5	7,0	8,5
Aluno 11	5,5	6,5	8,5
Aluno 22	2,0	7,0	8,5

Alunos que atingiram notas mais elevadas

Aluno	Prova escrita	Apresentação oral	Autoavaliação
Aluno 4	10,0	10,0	9,0
Aluno 3	9,0	8,5	8,0
Aluno 9	9,5	8,0	9,0
Aluno 16	9,0	8,5	8,5
Aluno 17	9,5	8,0	7,5

Analisando os dois grupos, é possível verificar que os alunos que obtiveram notas mais baixas tanto na prova escrita quanto na apresentação oral tendem a se autoavaliar com notas superiores a esses dois instrumentos, tanto individualmente quanto na média. A diferença mínima entre as notas comparadas de modo individual é 0,5 atingindo até 6,5 pontos de diferença, sempre superior para a autoavaliação.

Por outro lado, os cinco alunos com notas mais altas da sala tanto na prova escrita quanto na apresentação se autoavaliaram com notas abaixo de ambos os instrumentos de avaliação. A diferença mínima de 0,5 e máxima de 1,5, sempre inferior para a autoavaliação.

Assim, essa análise mostra indícios de que alunos que obtêm notas baixas nos instrumentos de avaliação aplicados no bimestre fornecem, na autoavaliação, notas superiores às obtidas nas avaliações da professora. Por outro lado, alunos que atingem as maiores notas nos instrumentos de rendimento se autoavaliam com notas inferiores às obtidas em tais instrumentos. Como a nota da autoavaliação tem peso na constituição da média final, não é possível descartar a hipótese de elevação dessa nota para compensar ou balancear resultados desfavoráveis obtidos em outros instrumentos de avaliação.

5 Discussão dos resultados

Os resultados deste estudo levantam a necessidade de reflexões sobre a utilização da autoavaliação como instrumento que visa avaliar as habilidades linguísticas do aluno, corroborando a preocupação apontada por Saito (2003), David e Henning (1985), Dickison (1987) e Blue (1994) sobre as correlações entre resultados de autoavaliação e testes realizados por especialistas, professores e testes padronizados.

Em relação à pergunta de pesquisa desta investigação sobre a existência de similaridades entre as notas de autoavaliação e as notas obtidas por meio de instrumentos formais de avaliação, o estudo comparou a nota de autoavaliação dos alunos por meio do t-Test com a prova escrita e a apresentação oral. Em ambas as comparações, não houve evidências de similaridades entre as notas. Além disso, a média da autoavaliação foi superior em comparação com os dois outros instrumentos avaliativos.

Em uma análise adicional, o estudo mostrou que alunos que obtêm as notas mais baixas do grupo nos instrumentos avaliativos utilizados pela professora tendem a fornecer notas mais altas na autoavaliação. O oposto ocorre com os alunos que atingem as notas mais altas do grupo, pois se autoavaliam com notas mais baixas do que as obtidas no teste escrito e na apresentação oral.

Entretanto, tal distinção não pode ser considerada um aspecto negativo da autoavaliação na medida em que reflete um recorte de um processo que pode sofrer intervenções para ajustes e esclarecimentos sobre os critérios da avaliação para o alinhamento entre os instrumentos. Como já foi apontado por Ross (2006), critérios e evidências para a autoavaliação, assim como a sua utilização para a nota do aluno, podem provocar benefícios na aprendizagem. Portanto, a recomendação diante desses resultados seria proporcionar esclarecimentos sobre os critérios e objetivos da autoavaliação especialmente junto aos alunos que atingiram médias mais baixas.

Apesar dessas diferenças, é preciso ressaltar também os aspectos positivos da autoavaliação na medida em que consiste em um processo de desenvolvimento dos alunos no contexto investigado, que começaram a assumir mais autonomia em seu aprendizado (ORSMOND, MERRY, REILLING, 1997) por meio deste estudo, tornando-se responsáveis, mesmo que parcialmente, pela composição de sua própria nota, o que pode contribuir para o seu empoderamento (TAN, 2008).

A autoavaliação neste trabalho não foi utilizada como um instrumento único de avaliação ou uma alternativa para algum outro, mas sim em conjunto com outros dois, tendo, inclusive, metade do peso das demais avaliações. Dessa forma, em vista desses resultados, não seria recomendável que a autoavaliação tivesse o papel de instrumento substitutivo de testes elaborados por professores ou padronizados no contexto investigado.

6 Considerações finais

Este artigo investigou o processo de autoavaliação em língua inglesa em uma sala de ensino médio de uma instituição pública fundamentado na utilização de instrumentos e critérios específicos. O objetivo consistiu em traçar uma comparação entre as notas das avaliações de rendimento, elaboradas e aplicadas pela professora no bimestre, e aquelas atribuídas pelos alunos na autoavaliação. Os resultados mostram diferenças nas notas concedidas pelos alunos em todos os instrumentos utilizados, com médias superiores às obtidas nas avaliações de rendimento, assim como indícios de que participantes com médias baixas nesses instrumentos tendem a se autoavaliar com notas mais altas, situação inversa identificada junto aos alunos com médias altas.

Desse modo, recomenda-se cautela na utilização da autoavaliação como um instrumento único voltado à análise do desempenho do aluno, entretanto, há benefícios que podem ser identificados, como oportunidade de reflexão sobre o desenvolvimento na língua e comparação pelo aluno da sua prática com parâmetros específicos.

Uma das limitações deste estudo refere-se a possíveis problemas de conteúdo dos instrumentos de avaliação. Os itens que compõem o questionário da autoavaliação não focalizaram somente questões relacionadas à capacidade de uso da língua, mas também atitudes e comportamentos como “Meu comportamento é bom. Chego pontualmente às aulas” ou “Venho com o material organizado para as aulas”. Embora esse tipo de item corresponda a cerca de 30% do questionário, somado ao fato de que poderiam representar indicativos de comprometimento com a aprendizagem e, conseqüentemente, colaborar para o desenvolvimento das habilidades linguísticas, poderiam, por outro lado, influenciar na elevação da nota da autoavaliação em comparação aos testes de rendimentos que avaliaram somente a capacidade de uso da língua e não aspectos comportamentais. Portanto, um estudo que busca aprofundar a investigação dessa possibilidade está em andamento.

Referências

ADELMAN, C., JENKINS, D., KEMMIS, S. Rethinking case study: notes from the second Cambridge conference. **Cambridge Journal of Education**, v.6, n. 3, p.139-10, 1976.

ASHRAF, H; MAHDINEZHAD, M. The Role of Peer-assessment versus Self-assessment in Promoting Autonomy in Language Use: A Case of EFL Learners. **Iranian Journal of Language Testing**, v. 5, n. 2, p. 110-120, 2015.

BACHMAN, L. F. Learner-directed assessment in ESL. In G. EKBATANI, H. PIERSON (Eds.), **Learner-directed assessment in ESL** (pp. ix-xii). New Jersey: Lawrence Erlbaum Associates, Inc., 2000.

BLATCHFORD, P. Students' self assessment of academic attainment: Accuracy and stability from 7 to 16 years and influence of domain and social comparison group. **Educational Psychology**, 17(3), 345-360, 1997.

BLUE, G.M. Self-assessment: the limit of learner independence. In: BROOKES, A.; GRUNDY, P. (eds.). **Individualisation and autonomy in language learning. ELT documents**, 131. London: Modern English Publications in association with the British Council (Macmillan), 1988.

CATO, M. Percepções de alunos universitários brasileiros aprendizes de inglês como LE sobre a auto-avaliação. Campinas: Unicamp. Dissertação de mestrado, 2016.

CAVALARI, S. M. S. A autoavaliação da aprendizagem de inglês como língua estrangeira em um ambiente tandem à distância. **RBLA, Belo Horizonte**, v. 11, n. 1, p. 247-270, 2011.

COHEN, L., MANION, L., MORRISON, K., EBOOKS Corporation. **Research methods in education** (7th ed.). Abingdon, Oxon; New York: Routledge, 2013.

DAVIDSON F. AND HENNING, G. A self-rating scale of English difficulty: Rasch scalar analysis. **Language Testing** 2: 164–179, 1985.

DENZIN, N. K. **Sociological methods**. Nova York: McGraw-Hill, 1978.

DICKINSON, L. **Self-instruction in language learning**. London: Cambridge University Press, 1987.

FIDALGO, S. S. Auto-avaliação: uma questão de prática? Ou de representações? Self-assessment: Practice or Representation? **the ESpecialist**, v. 26, n. 2, p.137-155, 2005.

GEE, J.P. (2005). **Why video games are good for your soul**: Pleasure and learning. Melbourne, Australia: Common Ground.

GHOLAMI, H. Self Assessment and Learner Autonomy. **Theory and Practice in Language Studies**, v. 6, n. 1, p. 46-51, p.-46-51,2016.

HARRIS, D. AND BELL, C. **Evaluating and Assessing for Learning**. East Brunswick, NJ: Nichols Publishing, 1994.

HARWELL, M. R. (2011). Research design: Qualitative, quantitative, and mixed methods: Pursuing ideas as the keystone of exemplary inquire. In C. Conrad, R. C. Serlin (Eds.), **The Sage handbook for research in education: Pursuing ideas as the keystone of exemplary inquire** (Second Edition ed.). Sage.

HAUGHTON, G., DICKINSON, L. Collaborative assessment by masters' candidates in a tutor based system. *Language Testing*, 5, 233-246, 1988.

HENNING, G. **A guide to language testing**: Development, evaluation, research. Rowley, MA: Newbury House, 1987.

HIATT, J. F. (1986). Spirituality, medicine, and healing. **Southern Medical Journal**, 79, 736–743, 1986.

HUGHES, A. Achievement and Proficiency: The missing link? Testing English for University Study. **ELT Documents**: 127 Modern English Publications, 1988.

JANSSEN-VAN DIETEN, A. The development of a test of Dutch as a foreign language: the validity of self-assessment by inexperienced subjects. *Language Testing*, 6 (1): 30-46, 1989.

JOHNSON, R. B.; ONWUEGBUZIE; A. J. Mixed Methods Research: A Research Paradigm Whose Time Has Come, **Educational Researcher**, vol.33, 2004, pp.14-26, 2004.

KALANTZIS, M. (2006). Changing subjectivities, new learning. **Pedagogies**: An International Journal, 1(1), 7–12.

KLENOWSKI, V. (1995). Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England. **Assessment in Education**, 2(2), 145-163.

KUNCEL, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis. **Review of Educational Research**, 75 (1), 63-82.

LeBLANC, R.; PAINCHAUD, G. Self assessment as a second language placement instrument. **TESOL Quarterly**, v. 19, n. 4, p. 673-687, Dez. 1985.

LEFFAL, V. *Língua estrangeira Ensino e aprendizagem*. Pelotas: EDUCAT, 2016.

LEE, Y. H. Comparing Self-assessment and Teacher's Assessment in Interpreter Training" **T&I Review** 1: 87–11, 2011.

MERRIAM, S. B. Case Study Research in Education: **A Qualitative Approach**. San Francisco: Jossey-Bass, 1988.

MICÁN, A., CUESTA MEDINA, L. Boosting vocabulary learning through self-assessment in an English language teaching context. **Assessment and Evaluation in Higher Education**. v. 42, p. 398-414, 2015.

MINISTÉRIO DA EDUCAÇÃO. Base Nacional Comum Curricular (ensino médio), 2017. NICOLAIDES; I. MOZZILLO; L. PACHALSKI; M. MACHADO; V. FERNANDES. (Org.). O desenvolvimento da autonomia no ambiente de aprendizagem de línguas estrangeiras. Pelotas: UFPEL, 2003.

NUNAN, D. **Research Methods in Language Learning**. Cambridge: Cambridge University Press, 2003.

OSCARSON, M. Self-assessment of language proficiency: Rationale and applications. **Language Testing**, 6, 1-13, 1989.

OXFORD, R.L. **Language learning strategies**: What every teacher should know. New York: Newbury House, 1990.

PEIRCE, B. M., SWAIN, M., HART, D. Self-assessment, French immersion, and locus of control. **Applied Linguistics**, 14, 25-42, 1993.

PUNHAGUI, G. C., SOUZA, N. A. A autoavaliação para a autorregulação da aprendizagem em língua estrangeira: apontamentos. **Cadernos de Educação FaE/PPGE/UFPEl**, v. 42, P. 199-222, 2012.

ROSS, J. The Reliability, Validity, and Utility of Self-Assessment. **Practical Assessment, Research and Evaluation**, v. 11, n. 10, p. 1-13, 2006.

SAITO, Y. The Use of Self-assessment in Second Language Assessment, 2003. Acesso em 02-02-2018. Disponível em: <https://journals.cdrs.columbia.edu/wp-content/uploads/sites/12/2015/05/3.3-Saito-2003.pdf>.

STRONG-KLAUSE, D. Exploring the effectiveness of self-assessment strategies in ESL placement. In G. Ekbatani, H. Pierson (Eds.), **Learner-directed assessment in ESL** (pp.49-73). New Jersey: Lawrence Erlbaum Associates, Inc., 2000.

TAN, K. Qualitatively different ways of experiencing student self assessment. **Higher Education Research and Development**, 27(1), 15-29, 2008.

TAVAKOLI, M. Investigating the relationship between Self-assessment and teacher-assessment in a academic contexts: A Case of Iranian university student. **The Asian EFL Journal Quarterly**, v. 12, n.1, p. 234-260, 2010.

TELLES, J. A. Teletandem Brasil: línguas estrangeiras para todos. Projeto apresentado ao CNPQ nº 061/2005 (Seleção pública de projetos de pesquisa nas áreas de Ciências Humanas, Sociais e Sociais Aplicadas). Faculdade de Ciências e Letras, Departamento de Educação, Universidade Estadual Paulista, Assis, 2005.

VALETTE, R. M. **Modern Language Testing**. New York: Harcourt Brace Jovanovich, 1977.

YIN, R. K. Case study research: Desing and methods. Thousand Oaks: CA: Sage, 2003