

MACE: Medida de Avaliação da Composição Escrita no Ensino Básico Português

SOFIA OLIVEIRA^I

JANETE SILVA MOREIRA^{II}

ANA LÚCIA AGOSTINHO^{III}

ANA MARGARIDA VEIGA-SIMÃO^{IV}

<http://dx.doi.org/10.22347/2175-2753v11i33.2217>

Resumo

Enquanto sistema de comunicação e aprendizagem, a escrita assume um papel central em diversos contextos, ocorrendo o seu ensino durante vários anos de escolaridade. Porém, a avaliação da escrita em língua portuguesa carece de instrumentos validados. Assim, procedeu-se ao estudo de precisão e validação da Medida de Avaliação da Composição Escrita (MACE), uma medida de observação construída para avaliar a composição escrita de textos narrativos no ensino básico. Participaram na pesquisa 280 alunos do 3º e 4º ano de escolaridade de escolas públicas de Lisboa. A recolha dos 560 textos foi realizada em dois momentos, cumprindo o mesmo procedimento. Os resultados evidenciaram uma estabilidade da medida ao longo do tempo e uma precisão na avaliação por diferentes juízes. A validade de conteúdo da MACE e a sua sensibilidade ao nível de competência de escrita foram também demonstradas. Discutem-se implicações da utilização do instrumento para as práticas de avaliação da composição escrita.

Palavras-chave: Escrita. Avaliação. Instrumento. Ensino Básico.

Submetido em: 06/03/2019

Aprovado em: 28/10/2019

^I Universidade de Lisboa, Faculdade de Psicologia, Lisboa, Portugal; <https://orcid.org/0000-0003-3941-1387>; E-mail: a-oliveira@campus.ul.pt.

^{II} Universidade de Lisboa, Faculdade de Psicologia, Lisboa, Portugal; <https://orcid.org/0000-0001-7504-0291>; E-mail: janetepsilva@gmail.com.

^{III} Universidade de Lisboa, Faculdade de Psicologia, Lisboa, Portugal; <https://orcid.org/0000-0003-4705-4041>; E-mail: ana_agostinho@hotmail.com.

^{IV} Universidade de Lisboa, Faculdade de Psicologia, Lisboa, Portugal; <https://orcid.org/0000-0003-3652-5573>; E-mail: amsimao@psicologia.ulisboa.pt.

MACE: Written Composition Assessment Measure in Portuguese Elementary Education

Abstract

As a communication and learning system, writing plays a central role in various contexts, and its teaching occurs during several years of schooling. However, writing assessment in Portuguese language lacks validated instruments. Thus, a study of accuracy and validation of the Written Composition Assessment Measure (MACE), a measure of observation built to evaluate the written composition of narrative texts in elementary education, was carried out. A total of 280 students from the 3rd and 4th grade of public schools of Lisbon participated in the research. The data collection of 560 narrative texts was performed in two moments, following the same procedure. The results showed a stability of the measurement over time and accuracy in its assessment by different judges. The content validity of MACE and its sensitivity to the level of writing competence were also demonstrated. Implications of the use of the instrument for the written composition evaluation practices were discussed.

Keywords: Writing. Assessment. Instrument. Elementary Education.

MACE: Medida de Evaluación de la Composición Escrita en la Educación Primaria Portuguesa

Resumen

Como sistema de comunicación y aprendizaje, la escritura asume un papel central en diferentes contextos, por lo cual su enseñanza se imparte durante varios años de escolaridad. Sin embargo, la evaluación de la escritura en la lengua portuguesa carece de instrumentos validados. Por lo tanto, se procedió al estudio de precisión y validación de la Medida de Evaluación de la Composición Escrita (MACE), una medida de observación construida para evaluar la composición escrita de textos narrativos en la educación primaria. En el estudio participaron 280 alumnos de 3^o y 4^o grado de escuelas públicas de Lisboa. Los 560 textos fueron recopilados en dos momentos, siguiendo el mismo procedimiento. Los resultados mostraron una estabilidad de la medición en el tiempo y una precisión en la evaluación por diferentes jueces. También se demostró la validez de contenido de MACE y su sensibilidad al nivel de competencia en escritura. Se discuten las implicaciones del uso del instrumento para las prácticas de evaluación de la composición escrita.

Palabras clave: Escrita. Evaluación. Instrumento. Educación Primaria.

Introdução

A escrita é uma importante forma de comunicação, que ocupa um lugar de destaque não só em contexto académico, enquanto meio privilegiado de acesso, partilha e expressão do saber, mas também como forma de participação na vida da sociedade atual (CARVALHO; BARBEIRO, 2013). Consiste, portanto, num processo complexo que envolve a mobilização de um conjunto de competências variadas e que, conseqüentemente, exige uma instrução explícita ao longo de vários anos de escolaridade (STRUTHERS; LAPADAT; MACMILLAN, 2013). Em Portugal, o Perfil do Aluno prevê que até ao final do 4º ano de escolaridade, os alunos sejam capazes de

redigir textos com utilização correta das formas de representação escrita (grafia, pontuação e translineação, configuração gráfica e sinais auxiliares da escrita); escrever textos, organizados em parágrafos, coesos, coerentes e adequados às convenções de representação gráfica (APRENDIZAGENS, 2018a, p. 11).

Nos anos subsequentes, o currículo visa apoiar os alunos no desenvolvimento do domínio de processos, estratégias, capacidades e conhecimentos para a escrita de diversos géneros de textos, com organização discursiva adequada, diversidade e propriedade vocabular, correção linguística e ortográfica (APRENDIZAGENS, 2018b).

Abordagens psicoeducativas para avaliar a composição escrita

Ao longo da escolaridade obrigatória, a avaliação ocupa, no processo de ensino explícito da escrita, um lugar de extrema importância ao permitir que os professores e outros profissionais da educação compreendam as necessidades de apoio específicas de cada aprendiz de forma a andaimar as suas aprendizagens (STRUTHERS; LAPADAT; MACMILLAN, 2013), respeitando o processo idiossincrático da escrita de cada aluno (CAPRISTANO, 2010). Neste sentido, uma avaliação adequada da escrita permite a adaptação do ensino ao nível de desenvolvimento individual dos alunos, bem como aos seus objetivos e metas de aprendizagem (DUNSMUIR et al., 2015; MARQUES; SILVESTRE; GONÇALVES, 2012). Tendo em vista a sua importância enquanto meio de comunicação, é fundamental que os profissionais de educação sejam capazes de identificar precocemente as dificuldades de escrita dos seus alunos, de forma a prevenir o insucesso escolar desde os primeiros anos de escolaridade (ALLEN; POCH; LEMBKE, 2018).

Dado o relevo atribuído à escrita, tem vindo a ser reconhecida a importância da adoção de processos de avaliação compreensiva desde o 1º ciclo. Este tipo de abordagem significa que, no processo de avaliação, são tidos em conta os aspetos mecânicos e os de conteúdo do produto escrito. Quer isto dizer que, através de uma abordagem analítica (ver item abaixo sobre Medidas analíticas de avaliação da composição escrita), o avaliador pode atender aos aspetos do produto final e também aos aspetos do processo de elaboração do texto (MALPIQUE; VEIGA SIMÃO, 2019). Ainda assim, apesar de se reconhecer o papel da escrita e a importância da sua avaliação compreensiva, a construção de medidas válidas e fiáveis para a avaliação da produção escrita, em particular do texto narrativo, tem-se evidenciado uma lacuna da investigação atual, quer no contexto português, quer ao nível internacional (COLE et al., 1997; DUNSMUIR et al., 2015; LEITE et al., 2018). Especificamente no contexto brasileiro, realizou-se um levantamento dos instrumentos disponíveis para a revisão de textos em língua portuguesa (LEÓN et al., 2016), tendo sido identificados 33 instrumentos dos quais: 21 avaliam a ortografia, oito, a produção textual, dois avaliam a ortografia e a produção de texto, um, a grafia, e um, a grafia e produção textual. Contudo, estes instrumentos não incluem aspetos relevantes da competência escrita considerados nos critérios das Aprendizagens Essenciais do Ensino Básico para a avaliação da escrita, tais como o conteúdo e a audiência (APRENDIZAGENS, 2018a).

Nos últimos anos, a literatura tem, por isso, procurado alertar para a necessidade de investir na construção e adaptação de instrumentos de avaliação adequados e, nesse sentido, têm sido discutidos os benefícios da utilização de medidas holísticas e analíticas. O primeiro tipo consiste numa avaliação geral da qualidade do produto escrito, baseada num conjunto de critérios predefinidos (por exemplo, erros ortográficos), que permite comparar o desempenho dos alunos (ESPIN; WEISSENBURGER; BENSON, 2004; KIM et al., 2014). Contudo a validade e a precisão de procedimentos holísticos têm sido questionadas (STRUTHERS; LAPADAT; MACMILLAN, 2013), evidenciando-se diversas vantagens na utilização de medidas analíticas, em contraste com as medidas holísticas (DUNSMUIR et al., 2015; OHTA; PLAKANS; GEBRIL, 2018). Além de ser um tipo de medida mais fiável, pela definição clara e objetiva do sistema de cotação, as medidas analíticas possibilitam a adequação das intervenções psicopedagógicas decorrente da possibilidade de detalhe na leitura do produto escrito. Esta abordagem contribui para traçar perfis de aprendizagem da escrita que monitoriza os pontos fortes e fracos da qualidade da escrita de cada

aluno ao longo do tempo (STRUTHERS; LAPADAT; MACMILLAN, 2013). A explicitação de uma pontuação analítica apresenta ainda potencialidades do ponto de vista do aplicador, na medida em que permite fornecer aos escritores um *feedback* específico associado à melhoria da escrita (BORGES; SILVA, 2017; MARQUES; SILVESTRE; GONÇALVES, 2012; VEIGA SIMÃO, 2008).

Medidas analíticas de avaliação da composição escrita

Em Portugal, os critérios de avaliação propostos pelo Instituto de Avaliação Educativa (IAVE) para os vários ciclos de aprendizagem têm adotado uma visão mais compreensiva do processo de escrita, considerando os seguintes parâmetros: extensão, formato textual, tema e pertinência da informação, organização e coesão textuais, sintaxe e morfologia, pontuação, repertório vocabular, e ortografia (IAVE, 2019). A opção por critérios menos dependentes da produção mecânica (por exemplo, número de palavras escritas) vem permitir a valorização da qualidade do produto escrito para além da extensão do texto, que pouca informação oferece na compreensão do processo de aprendizagem (ALLEN; POCH; LEMBKE, 2018). Pelas suas características, a adoção destas abordagens (menos centradas nos aspetos mecânicos) permite promover o equilíbrio nos critérios de avaliação da escrita (ou seja, mecânicos e de conteúdo) aquando da construção de instrumentos para este fim, estando em linha com o que é proposto por literatura recente nesta área de estudo (MCGRANE; HUMPHRY; HELDSINGER, 2018).

Apesar de se reconhecerem vantagens na utilização de medidas analíticas, verifica-se, à data, uma carência de medidas validadas e específicas para a avaliação da qualidade da composição escrita (LEITE et al., 2018). Especificamente no âmbito do ensino do Português, a escrita tem sido frequentemente avaliada através de instrumentos não validados, construídos com base no currículo e nos critérios de classificação dos exames nacionais e provas de aferição. Consequentemente, dado o panorama atual do sistema educativo em Portugal, verifica-se a necessidade de desenvolver instrumentos deste tipo, adequados aos vários níveis de ensino e que permitam uma avaliação compreensiva do processo de aprendizagem (MARQUES; SILVESTRE; GONÇALVES, 2012).

No contexto português, conhece-se apenas um instrumento de avaliação analítica da qualidade da composição escrita, ou seja, Avaliação Informal da Composição Escrita (GONÇALVES, 2012), adaptado a partir do *ESL Composition Profile* (JACOBS et al., 1981) e que inclui seis áreas de análise do produto escrito:

conteúdo do texto, estrutura global, vocabulário, estrutura das frases, aspetos mecânicos (ortografia, acentuação, pontuação), e aspetos específicos (operações de revisão, originalidade e perspetiva pessoal). Porém, não só este instrumento foi desenvolvido com vista à avaliação do desempenho de escritores experientes (MARQUES; SILVESTRE; GONÇALVES, 2012) e com bom domínio da escrita como carece ainda de estudos de validação (GONÇALVES, 2012), permanecendo a necessidade de construir instrumentos para crianças ainda em fase de aprendizagem da escrita.

No contexto internacional, salientam-se dois instrumentos que permitem não só avaliar as competências das crianças sobre os aspetos mais formais da escrita, como as suas competências para formular e desenvolver ideias num estilo discursivo coerente. Na Austrália, foi desenvolvida uma medida de avaliação da escrita com base no currículo, *Online Literacy and Numeracy Assessment – OLNA* (SCHOOL CURRICULUM AND STANDARDS AUTHORITY, 2015), que considera sete critérios de qualidade: audiência, estrutura e organização, vocabulário, coesão, estrutura frásica, pontuação e ortografia. Este instrumento standardizado, ao avaliar a qualidade da escrita numa escala de quatro pontos para cada um dos critérios mencionados, permite traçar um perfil de desempenho da escrita para cada aluno, bem como monitorizar a sua evolução. De um modo semelhante, no Reino Unido, foi construído um instrumento de avaliação da escrita com base no currículo para crianças entre os 7 e os 11 anos, *Writing Assessment Measure – WAM* (DUNSMUIR et al., 2015), que considera igualmente quatro níveis de avaliação para cada um dos seguintes critérios: caligrafia, ortografia, pontuação, estrutura frásica e gramática, vocabulário, organização e estrutura geral e ideias.

Apesar da existência destas medidas no âmbito internacional, no contexto português elas poderão apenas ser tidas como base de orientação para a construção ou adaptação de um instrumento de avaliação da composição escrita em língua portuguesa, uma vez que, em resultado das especificidades da língua e do currículo escolar dos países anglo-saxónicos, a sua mera tradução e adaptação não seria adequada.

Construção da Medida de Avaliação da Composição Escrita (MACE)

Considerando o número reduzido de instrumentos analíticos para a avaliação compreensiva da composição escrita de alunos do ensino básico em língua portuguesa, considerou-se pertinente a construção de um instrumento para este fim. Tendo em conta

o currículo escolar português, foi desenvolvido um instrumento de avaliação da qualidade da escrita de crianças do 1º ciclo, Medida de Avaliação da Composição Escrita (MACE).

Para a seleção dos critérios de avaliação analítica, foi realizada uma revisão de literatura sobre instrumentos semelhantes (por exemplo, OLNA, WAM), bem como uma análise dos parâmetros considerados pelo Ministério da Educação – MEC (APRENDIZAGENS, 2018a; APRENDIZAGENS, 2018b). Chegou-se, assim, à identificação de seis critérios de avaliação, que incluem quer aspetos mecânicos quer de conteúdo: estrutura e organização, conteúdo, vocabulário, audiência, convenções gramaticais e originalidade.

Atendendo à pertinência da utilização de medidas válidas e fiáveis na avaliação da escrita, este instrumento tem como objetivos:

1. construir uma medida, válida e precisa, que permita a identificação de necessidades específicas na aprendizagem da escrita da língua portuguesa;
2. permitir ao profissional de educação a identificação de necessidades específicas nos seus alunos, no domínio da escrita;
3. contribuir para o planeamento estratégico de intervenções que visam a promoção de competências de escrita adequadas a cada aluno;
4. possibilitar *feedback* que auxilie a aprendizagem da escrita.

Especificamente, a investigação foi orientada pela seguinte questão de investigação: o instrumento MACE, que avalia estrutura/organização, conteúdo, vocabulário, audiência, convenções gramaticais e originalidade, é uma medida que identifica, de forma válida e precisa, as diferenças na evolução da composição escrita dos alunos do 3º e 4º ano?

Método

Participantes

Nesta pesquisa, participaram 280 alunos [Média (*M*) = 8,58 anos; Desvio-padrão (*DP*) = 0,79; 55% meninos] do 3º (*n* = 170; *M* = 8,10 anos; *DP* = 0,47; 61% meninos) e 4º (*n* = 110; *M* = 9,31 anos; *DP* = 0,59; 56% meninas) anos do 1º Ciclo do Ensino Básico. A amostra de conveniência foi recolhida em cinco escolas públicas da região metropolitana de Lisboa, em Portugal, das quais três estavam inseridas no programa Territórios Educativos de Intervenção Prioritária (ou seja, escolas localizadas em territórios económica e

socialmente desfavorecidos). Todos os participantes falavam Português fluentemente e não tinham diagnóstico de Necessidade Educativas Especiais.

Instrumentos

MACE. Esta medida de observação permite a avaliação da qualidade da escrita de textos narrativos, tendo sido desenvolvida para o contexto português em coerência com as Metas Curriculares para o Português (BUESCU et al., 2015) e tomando como referência as escalas *Wechsler Individual Achievement Test* (DUNSMUIR et al., 2015; WECHSLER, 2005) e OLNA (SCHOOL CURRICULUM AND STANDARDS AUTHORITY, 2015).

A MACE é uma escala de avaliação analítica das competências de escrita, composta por seis critérios: Estrutura/Organização (refere-se à organização dos componentes do texto numa estrutura efetiva), Conteúdo (refere-se à apresentação da informação de forma coerente e coesa), Vocabulário (refere-se à amplitude e à precisão de escolhas linguísticas apropriadas ao contexto), Audiência (refere-se à capacidade do autor de orientar e envolver o leitor), Convenções gramaticais (refere-se à capacidade do autor em utilizar as convenções gramaticais do Português escrito) e Originalidade (refere-se à vertente imaginativa na escrita, à criatividade e à criação de recursos inovadores por parte do autor). Cada critério é avaliado numa escala de quatro pontos, cada um operacionalizado com uma descrição específica e clara para evitar enviesamentos na cotação (ver Anexo 1). Esta escala permite assim a avaliação do texto narrativo, podendo esta ser interpretada com base no valor individual de cada critério ou numa cotação global da escala.

Procedimento

Para a realização da presente investigação, foi pedido o consentimento informado aos diretores de escola, professores, encarregados de educação e alunos. A recolha de dados decorreu em sala de aula, com a presença de um investigador e do professor titular de turma. Foram considerados dois momentos de recolha de dados (ou seja, Texto 1 e Texto 2), distados em aproximadamente 45 dias, para avaliar a estabilidade da medida ao longo do tempo (precisão teste-reteste). Em ambos os momentos, foi pedido aos participantes que, num tempo máximo de 30 minutos (justificado pela literatura e pelos relatos da prática dos professores), redigissem um

texto narrativo com o tema "Uma aventura". Foi produzido um total de 560 textos narrativos.

Posteriormente, todos os textos foram avaliados através da MACE, por juízes independentes (ou seja, quatro psicólogos educacionais especialistas na avaliação de competências de composição escrita). De modo a assegurar a avaliação cega dos textos narrativos, um código alfanumérico foi aleatoriamente atribuído a cada participante. Para garantir a consistência da avaliação, numa primeira fase, todos os juízes participaram em reuniões dedicadas à seleção e operacionalização dos critérios e respetivos descritores considerados na MACE. Posteriormente, para garantir a fiabilidade do processo de análise, o painel de especialistas avaliou 10% dos textos produzidos (AMADO, 2014). Finalmente, os restantes textos foram divididos pelos quatro juízes, que procederam a uma avaliação independente. De forma a atestar o cálculo da precisão teste-reteste, cada juiz avaliou os dois textos narrativos produzidos pelo mesmo participante.

Análise de dados

A análise estatística dos dados foi realizada com recurso ao *software* estatístico IBM SPSS (v.24, SPSS Inc., Chicago, IL). Primeiramente, para garantir a validade dos resultados, foi avaliado o acordo interjuízes, através do cálculo do Coeficiente de Correlação Intraclasses (ICC). Numa segunda fase, realizou-se o estudo da estrutura interna, através de uma Análise Fatorial Exploratória, e validação da aplicabilidade e adequação da MACE, através da ANOVA Unidirecional e do estudo da validade de conteúdo. Consideraram-se efeitos estatisticamente significativos valores $p < 0,05$.

Resultados

Acordo interjuízes

Para assegurar a fiabilidade do processo de análise, 56 textos narrativos (10%) foram avaliados através da MACE pelo painel de quatro especialistas e o Coeficiente de Correlação Intraclasses (ICC) foi calculado. O valor médio de ICC foi de 0,87, $p < 0,001$ indicando um acordo interjuízes "quase perfeito" (LANDIS; KOCH, 1977). Na Tabela 1, estão representados os valores de acordo atingido entre os quatro juízes para cada critério da MACE e para a escala-total.

Tabela 1 – Acordo interjuízes para cada critério e escala-total

Critério	Coefficiente de Correlação Intraclases
Estrutura/Organização	0,91**
Conteúdo	0,87**
Vocabulário	0,84**
Audiência	0,81**
Convenções Gramaticais	0,88**
Originalidade	0,84**
Escala total	0,95**

** $p < 0,001$

Fonte: Os autores (2019).

Análise fatorial exploratória

Para estudar a estrutura interna da MACE, foi realizada uma análise fatorial exploratória, considerando o total de 560 textos produzidos pelos participantes. Num primeiro momento, foram verificados os pressupostos da normalidade para cada critério. Adicionalmente, foram realizados o teste de Kaiser-Meyer-Olkin ($KMO = 0,87$) e o teste de esfericidade de Bartlett [$\chi^2(15) = 1546,3$, $p < 0,001$], indicando adequação da MACE à realização de uma análise fatorial exploratória (FIELD, 2013).

Para definir o número de fatores a reter foi utilizado o critério de Kaiser. Este critério é preciso para $N > 250$ e média de comunalidades superior a 0,60 (FIELD, 2013), o que se verifica nesta pesquisa ($N = 560$; $M_{comunalidades} = 0,62$). De acordo com o critério utilizado, foi extraído um fator único com 61,56% de variância explicada. Ainda, verificou-se que os critérios da MACE apresentam, entre si, correlações positivas de força moderada a forte $r > 0,30$ (COHEN, 1988) conforme Tabela 2. Em decurso da análise realizada, foram mantidos todos os critérios da MACE.

Tabela 2 – Estatísticas descritivas dos critérios, cargas fatoriais e correlações critério-critério e critério-escala total

	Média (Desvio-padrão)	Carga fatorial	1	2	3	4	5	6
1. Estrutura/Organização	3,24 (0,78)	0,75	-					
2. Conteúdo	3,01 (0,77)	0,81	0,57	-				
3. Vocabulário	2,91 (0,72)	0,77	0,41	0,56	-			
4. Audiência	2,79 (0,75)	0,84	0,53	0,60	0,59	-		
5. Convenções Gramaticais	2,97 (0,77)	0,70	0,49	0,49	0,42	0,56	-	
6. Originalidade	2,84 (0,84)	0,82	0,55	0,60	0,63	0,65	0,41	-
7. Escala total	17,77 (3,61)	-	0,76	0,81	0,76	0,83	0,71	0,82

Nota: todas as correlações são significativas com $p < 0,01$ (bilateral).

Fonte: Os autores (2019).

A consistência interna da MACE, enquanto medida unidimensional, foi estudada através do *alfa de Cronbach* (α). Obteve-se um coeficiente α de 0,87 para a escala total (isto é, considerando os seis critérios), o que indica uma consistência interna global “boa” (FIELD, 2013). Adicionalmente, todos os valores de α “se o item for excluído” são similares e inferiores ao α para a escala total, o que sugere uma consistência dos critérios considerados na MACE. O reporte completo do estudo de consistência interna é apresentado na Tabela 3.

Tabela 3 – Estudo de consistência interna com estatísticas item-total

Itens	Média da escala se o item for excluído	Variância de escala se o item for excluído	Correlação de item total corrigida	Alfa de Cronbach se o item for excluído
Estrutura/Organização	14,53	9,40	0,64	0,86
Conteúdo	14,76	9,16	0,71	0,85
Vocabulário	14,86	9,62	0,66	0,86
Audiência	14,98	9,12	0,75	0,84
Convenções Gramaticais	14,79	9,67	0,58	0,86
Originalidade	14,92	8,78	0,72	0,84

Fonte: Os autores (2019).

Precisão teste-reteste

O cálculo da precisão teste-reteste foi realizado a fim de garantir a estabilidade da avaliação de textos narrativos com recurso à MACE. Os dois textos de cada participante foram avaliados pelo mesmo juiz com um intervalo aproximado de 45 dias. O coeficiente r de Pearson foi calculado por forma a avaliar a força de associação das cotações totais atribuídas, ao mesmo participante, nos dois momentos. Os resultados demonstraram uma correlação positiva, significativa e forte, com $r = 0,69$, $p < 0,001$ (COHEN, 1988).

Validade de conteúdo

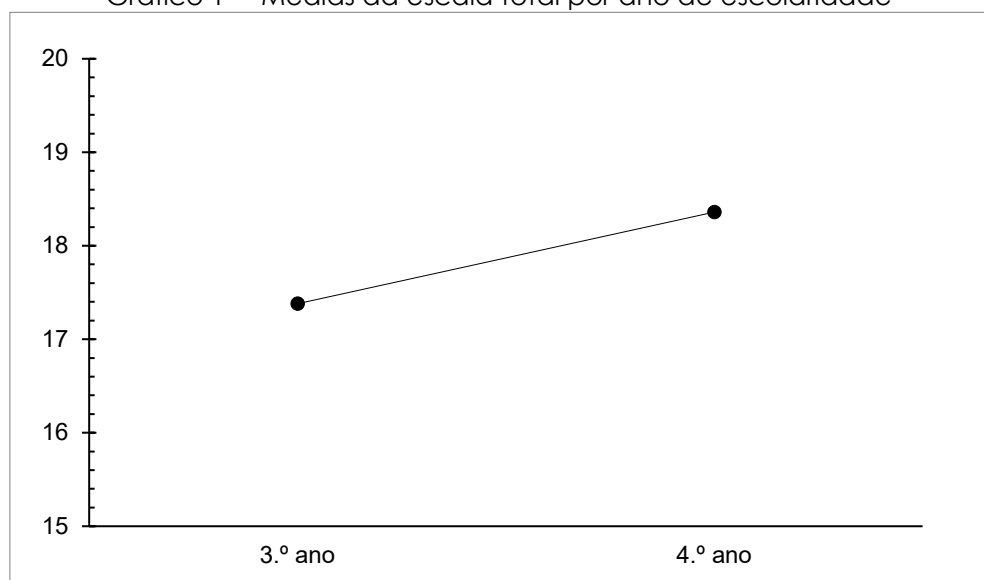
As correlações entre cada critério da MACE e a escala total foram calculadas através do coeficiente r de Pearson (conforme Tabela 2). Todos os coeficientes de correlação item-total foram significativos com $p < 0,001$ (entre $r = 0,71$ para *Convenções gramaticais* e $r = 0,83$ para *Audiência*). As correlações inter-item foram também significativas na sua totalidade, variando entre $r = 0,65$, $p < 0,001$ (*Audiência/Originalidade*) e $r = 0,41$, $p < 0,001$ (*Convenções gramaticais/Originalidade* e *Estrutura e Organização/Vocabulário*).

Todas as correlações verificadas são positivas e, na sua maioria, fortes (ou seja, $r > 0,50$) de acordo com os critérios de Cohen (1998).

Atributos observáveis

Para estimar a diferença na avaliação através da MACE, de textos narrativos produzidos em diferentes anos de escolaridade (ou seja, 3º e 4º ano), foi realizada uma ANOVA Unidirecional. Foi encontrada uma diferença estatisticamente significativa [$F(1, 558) = 10,02, p = 0,002$] entre a média obtida pelos alunos do 3º ($M = 17,38, DP = 3,63$) e do 4º ano ($M = 18,36, DP = 3,51$). O Gráfico 1 representa as médias de cada ano de escolaridade.

Gráfico 1 – Médias da escala total por ano de escolaridade



Fonte: Os autores (2019).

Discussão

A presente investigação pretendeu contribuir para a construção de uma medida válida e precisa, que permitisse a identificação de necessidades específicas na aprendizagem da escrita da língua portuguesa nos alunos do 1º ciclo. Para tal, realizou-se um conjunto de análises estatísticas com vista à validação da MACE. Os resultados obtidos evidenciaram níveis elevados de robustez e precisão na avaliação da composição escrita e adequação da MACE ao contexto educativo para o qual foi construído. Especificamente, foram registadas correlações significativas, positivas e de força moderada a forte entre critérios e entre critério-escala total, sugerindo que os vários indicadores da MACE contribuem para a avaliação da mesma

competência (isto é, composição escrita). No mesmo sentido, a análise fatorial exploratória evidenciou a unidimensionalidade da MACE, o que reforça os resultados anteriormente obtidos. Consequentemente, numa perspetiva funcional da aprendizagem e com vista à monitorização e compreensão do processo de escrita, a MACE revelou-se uma medida útil e completa, permitindo não só uma avaliação global do texto narrativo, como também uma leitura mais detalhada através da análise de cada critério individualmente. Considerando o desenvolvimento humano como um processo não linear, que inclui avanços e retrocessos, esta característica da MACE adquire grande importância para o acompanhamento individual dos alunos.

Adicionalmente e tal como noutras investigações (DUNSMUIR et al., 2015; SCHOOL CURRICULUM AND STANDARDS AUTHORITY, 2015), a MACE revelou-se sensível ao nível de desempenho dos alunos na realização de composições escritas, contribuindo para aferir mudanças no sentido da evolução na aprendizagem. A computação da ANOVA Unidirecional permitiu verificar resultados significativamente superiores nos alunos do 4º ano em todos os critérios avaliados (ou seja, estrutura/organização, conteúdo, vocabulário, audiência, convenções gramaticais e originalidade da composição escrita), por comparação com os alunos do 3º ano.

Os resultados encontrados suportam a pertinência da utilização de instrumentos de avaliação analítica que considerem critérios que vão para além dos aspetos formais da língua, contribuindo para uma avaliação abrangente e compreensiva da competência de escrita, como tem vindo a ser destacado também no contexto internacional (BRAGA; PIMENTEL; AGUIAR, 2016; LEÓN et al., 2016). Desta forma, atribui-se um peso equivalente aos aspetos mecânicos e aos aspetos de conteúdo da escrita (MCGRANE; HUMPHRY; HELDSINGER, 2018). Os dados apresentados permitiram ainda reforçar a utilidade da MACE na avaliação funcional da composição escrita no 1º ciclo do Ensino Básico e, consequentemente, o contributo desta medida para o planeamento estratégico de intervenções psicoeducativas no âmbito da escrita (WAGNER et al., 2018), tendo em conta o Perfil do Aluno (APRENDIZAGENS, 2018a).

A construção de medidas de avaliação da composição escrita deve, pois, espelhar a complexidade do processo de escrita, que implica uma série de etapas e operações inter-relacionadas entre si e a níveis distintos (ALLEN; POCH; LEMBKE, 2018;

BRAGA; PIMENTEL; AGUIAR, 2016; CARVALHO; BARBEIRO, 2013). Esta visão da escrita enquanto atividade complexa exige que os indivíduos, para além de estarem conscientes das competências a mobilizar, sejam capazes de as autorregular em todo o processo. Neste sentido, ao permitirem um *feedback* de desempenho (VEIGA SIMÃO, 2008; BORGES; SILVA, 2017), as medidas analíticas revelam-se um instrumento de apoio importante para o conhecimento compreensivo destas competências, tanto pelo profissional de educação, como pelo próprio aluno (STRUTHERS; LAPADAT; MACMILLAN, 2013; VIANA et al., 2017). Assim, este tipo de medidas, nomeadamente a MACE, podem ser utilizadas desde cedo, preferencialmente nos primeiros anos de escolaridade, de forma a facilitar uma aprendizagem autorregulada do processo de escrita (PISCALHO; VEIGA SIMÃO, 2014; VEIGA SIMÃO, 2017). Contudo, esta abordagem de ensino da escrita pressupõe um ensino centrado nos processos e exige dos profissionais de educação o domínio da estratégia a ensinar pelo que, ele próprio, deve ser capaz de adotar uma postura reflexiva face às suas práticas de escrita (CARLOS; PEREIRA, 2011) e adaptar constantemente os métodos de ensino e aprendizagem (ALLEN; POCH; LEMBKE, 2018).

Por outro lado, a coavaliação dos produtos escritos (ou seja, a avaliação do profissional de educação e do aluno, em conjunto), um dos métodos mais utilizados para ajudar os alunos na tomada de consciência dos seus progressos, pode beneficiar bastante do *feedback* fornecido através das medidas analíticas. Este tipo de informação permite ao aluno e ao profissional de educação, perceber se os objetivos predefinidos estão a ser alcançados ou se é necessário fazer ajustes no percurso de aprendizagem. Especificamente, a estratégia de fornecimento de *feedback* sobre o texto narrativo tem mostrado ser uma das mais eficazes, sobretudo quando o processo auxilia o aluno a identificar os elementos em falta ou os que podem ser melhorados (BORGES; SILVA, 2017). Estas estratégias são essenciais para a promoção da autoavaliação. De facto, quando o aluno aciona os seus processos metacognitivos e de autocontrolo sobre o seu desempenho, toma consciência dos erros, da maneira como se confronta com os obstáculos, regulando os processos de ensino e aprendizagem (KIM et al., 2014; VIANA et al., 2017). Desta forma, a utilização de medidas como a MACE, que permitem uma avaliação abrangente e compreensiva do desempenho na composição escrita e contribuem para discriminar os pontos fortes e fracos e a eficácia das estratégias adotadas, em complemento de uma reflexão guiada pelo profissional de educação, congregam potencialidades

para a promoção da aprendizagem autorregulada com vista à autonomia dos alunos (VEIGA SIMÃO, 2008).

Considera-se, então, que a MACE deve ser utilizada numa perspetiva de avaliação formativa. Sendo a avaliação um processo complexo que compreende normalmente a definição prévia de critérios, a tomada de decisões pertinentes, assim como a comunicação clara e sistemática dos resultados, a avaliação através da MACE deverá olhar a composição escrita, não só enquanto produto, mas fundamentalmente enquanto processo (MALPIQUE; VEIGA SIMÃO, 2019). De facto, a avaliação formativa tem como principal função informar o aprendente e o profissional de educação sobre a posição alcançada face ao objetivo estabelecido, contribuindo para adequar as práticas educativas (VEIGA SIMÃO, 2008). A versatilidade da MACE contribui, pois, para todos os aspetos levantados, sendo a intencionalidade pedagógica do profissional da educação que define qual a componente a dar mais importância em cada situação, seja o *feedback*, seja o resultado da avaliação numa lógica de evolução do aprendente. É, deste modo, fundamental salientar o duplo valor regulador da avaliação, em particular através de medidas analíticas: por um lado, ela está diretamente relacionada com o papel formativo e regulador do ensino; mas por outro lado, também desempenha um papel regulador da própria aprendizagem. Assim encarada, a avaliação funciona como ferramenta para a adaptação constante das formas de ensino às características dos alunos e para facilitar o assumir, cada vez maior, do controlo e da responsabilidade por parte do aluno sobre o seu próprio processo de aprendizagem (VEIGA SIMÃO, 2008).

Considerações finais

A presente investigação permitiu validar a MACE enquanto medida válida e precisa para a avaliação da composição escrita no 3º e 4º anos de escolaridade. Os principais resultados do estudo da validação da MACE sustentam o seu contributo enquanto medida de avaliação funcional da composição escrita, permitindo a avaliação global do texto narrativo bem como uma avaliação compreensiva através da análise por critério (ou seja, estrutura/organização, conteúdo, vocabulário, audiência, convenções gramaticais e originalidade). Assim sendo, a MACE representa um instrumento particularmente útil na perspetiva funcional da

aprendizagem, numa lógica de monitorização da evolução do aprendente que pode ser utilizado como indicador para a melhoria da composição escrita, mas também na promoção dos processos autorregulatórios subjacentes ao processo de escrita (ou seja, planear, monitorizar e rever).

Não obstante, esta pesquisa comporta algumas limitações que devem ser consideradas, nomeadamente no que concerne a algumas opções metodológicas. Especificamente, foi utilizada uma amostra de conveniência condicionando a generalização dos resultados obtidos. No mesmo sentido, considerou-se uma amostra geográfica e culturalmente circunscrita. Assim, e tendo em conta que a validação é um processo contínuo e dependente do contexto, sugere-se que investigação futura conduza novos estudos de validação da MACE com recurso a amostras diversificadas e representativas, estudando a sua adequação a outras faixas etárias e contextos educativos. Por fim, apesar de não ter sido possível nesta investigação, futuramente deve também contemplar-se um estudo de validação concorrente e divergente da MACE.

Referências

ALLEN, A.; POCH, A.; LEMBKE, E. An exploration of alternative scoring methods using curriculum-based measurement in early writing. *Learning Disability Quarterly, Califórnia*, v. 41, n. 2, p. 85-99, maio 2018.

AMADO, J. *Manual de investigação qualitativa em educação*. 2. ed. Coimbra: Imprensa da Universidade de Coimbra, 2014.

APRENDIZAGENS essenciais: articulação com o perfil dos alunos: português, 4.º ano, 1.º ciclo do ensino básico. Lisboa: Ministério da Educação e Ciência, 2018a.

Disponível em:

http://www.dge.mec.pt/sites/default/files/Curriculo/Aprendizagens_Essenciais/1_ciclo/portugues_1c_4a_ff.pdf. Acesso em: 4 jan. 2019

APRENDIZAGENS essenciais: articulação com o perfil dos alunos: português, 6.º ano, 2.º ciclo do ensino básico. Lisboa: Ministério da Educação e Ciência, 2018b.

Disponível em:

http://www.dge.mec.pt/sites/default/files/Curriculo/Aprendizagens_Essenciais/2_ciclo/6_portugues.pdf. Acesso em: 4 jan. 2019.

BORGES, S.; SILVA, A. C. A revisão como um instrumento para a melhoria da qualidade de textos narrativos. *Revista de Estudios e Investigación en Psicología y Educación*, Coruña, n. 12, p. 331-336, dez. 2017.

BRAGA, I.; PIMENTEL, S. R. G.; AGUIAR, G. S. Avaliação do desempenho de leitura e escrita dos alunos do 2º e 3º anos do ensino fundamental I do colégio Pedro II - Campus Humaitá I: uma meta-avaliação. *Revista Meta: Avaliação*, Rio de Janeiro, v. 8, ed. especial, p. 21-36, 2016.

BUESCU, H. et al. *Programa e metas curriculares de português do ensino básico*. Lisboa: Ministério da Educação e Ciência, 2015. Disponível em:

https://www.dge.mec.pt/sites/default/files/Basico/Metas/Portugues/pmcpeb_julho_2015.pdf. Acesso em: 5 dez. 2019.

CAPRISTANO, C. Por uma concepção heterogênea da escrita que se produz e que se ensina na escola. *Cadernos de Educação*, Pelotas, v. 35, p. 171-193, jan./abr. 2010.

CARLOS, L. C.; PEREIRA, M. de F. C. R. A escrita como dispositivo de formação em educação. *Cadernos de Educação*, Pelotas, v. 39, p. 175-192, maio/ago. 2011.

CARVALHO, J. A. B.; BARBEIRO, L. F. Reproduzir ou construir conhecimento? Funções da escrita no contexto escolar português. *Revista brasileira de educação*, Rio de Janeiro, v. 18, n. 54, p. 609-628, jul./set. 2013.

COHEN, J. *Statistical power analysis for the behavioral sciences*. 2. ed. Hillsdale: Lawrence Erlbaum Associates, 1988.

COLE, J. et al. Written expression reviewed. *Research in the Schools*, [s. l.], v. 4, n. 1, p. 17-34, 1997.

DUNSMUIR, S. et al. An evaluation of the Writing Assessment Measure (WAM) for children's narrative writing. *Assessing Writing*, Amsterdam, v. 23, p. 1-18, jan. 2015.

EPIN, C. A.; WEISSENBURGER, J. W.; BENSON, J. B. Assessing the writing performance of students in special education. *Exceptionality*, Londres, v. 12, p. 55-66, jun. 2004.

FIELD, A. *Discovering statistics using IBM SPSS statistics*. 4. ed. Londres: Sage Publications, 2013.

GONÇALVES, M. D. Avaliação funcional e monitorização das aprendizagens. In:

VEIGA SIMÃO, A. M.; FRISON, L.; ABRAHÃO, M. H. (org.). *Autorregulação da aprendizagem e narrativas autobiográficas: epistemologia e práticas*. Natal: EDUFRN; Porto Alegre: EDIPUCRS; Salvador: EDUNEB, 2012. p. 229-257.

INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P. [IAVE]. *Provas de Aferição 1º ciclo: português e estudo do meio*. Lisboa: Ministério da Educação e Ciência, 2019. Disponível em: <<http://iave.pt/index.php/avaliacao-de-alunos/arquivo-de-provas-exames/provas-de-afericao-1-ciclo>>. Acesso em: 9 jan. 2019.

JACOBS, H. et al. *Testing esl composition: a practical approach*. Londres: Newbury House Publishers, Inc., 1981.

KIM, Y. S. et al. Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research*, Rockville, v. 57, n. 1, p. 199-211, fev. 2014.

LANDIS, J. R.; KOCH, G. The measurement of observer agreement for categorical data. *Biometrics*, Washington DC, v. 33, n. 1, p. 159-174, mar. 1977.

LEITE, L. et al. A prática da avaliação como instrumento de colaboração e autoria dos discentes. *Meta: Avaliação*, Rio de Janeiro, v. 10, n. 28, p. 235-246, jan./abr. 2018.

LEÓN, C. et al. Como avaliar a escrita? Revisão de instrumentos a partir das pesquisas nacionais. *Revista Psicopedagogia*, São Paulo, v. 33, n. 102, p. 331-345, maio 2016.

MALPIQUE, A.; VEIGA SIMÃO, A. M. 'Does it work?': Adapting evidence-based practices to teach argumentative writing. *Journal of Writing Research*, Antuérpia, v. 10, n. 3, p. 527-567, fev. 2019.

MARQUES, M.; SILVESTRE, A.; GONÇALVES, M. D. Avaliação da composição escrita em alunos do 1.º ciclo. In: GONÇALVES, M. D. (org.). *Encontros IDEA: investigação de dificuldades para a evolução na aprendizagem (Livro I)*. Óbidos: Sinapsis Editores, 2012. p. 195-210.

MCGRANE, J.; HUMPHRY, S.; HELDSINGER, S. Applying a thurstonian, two-stage method in the standardized assessment of writing. *Applied Measurement in Education*, Oxford, v. 31, n. 4, p. 297-311, out. 2018.

OHTA, R.; PLAKANS, L.; GEBRIL, A. Integrated writing scores based on holistic and multi-trait scales: a generalizability analysis. *Assessing Writing*, Amsterdam, v. 38, p. 21-36, out. 2018.

PISCALHO, I.; VEIGA SIMÃO, A. M. Promover competências autorregulatórias da aprendizagem nas crianças dos 5 aos 7 anos: perspectivas de investigadores e docentes. *Interações*, Santarém, BA, v. 10, n. 30, p. 72-109, jun. 2014.

SCHOOL CURRICULUM AND STANDARDS AUTHORITY. *Online literacy and numeracy assessment (OLNA): writing guide*. Cannington: School Curriculum and Standards Authority, 2015. Disponível em: https://senior-secondary.scsa.wa.edu.au/__data/assets/pdf_file/0008/76409/OLNA_writing_guide.pdf. Acesso em: 4 jan. 2019.

STRUTHERS, L.; LAPADAT, J. C.; MACMILLAN, P. D. Assessing cohesion in children's writing: development of a checklist. *Assessing Writing*, Amsterdam, v. 18, n. 3, p. 187-201, jul. 2013.

VIANA, F. L. et al. O ensino explícito da compreensão da leitura: análise do impacto de um programa de intervenção. *Revista Brasileira de Educação*, Rio de Janeiro, v. 22, n. 71, p. 1-30, dez. 2017.

VEIGA SIMÃO, A. M. Reforçar o valor regulador, formativo e formador da avaliação das aprendizagens. In: ALVES, Maria; MACHADO, Eusébio (org.). *Avaliação com sentido(s): contributos e questionamentos*. Santo Tirso: De Facto Editores, 2008. p. 125-151.

VEIGA SIMÃO, A. M. et al. *Criativo: promoção de estratégias de autorregulação na escrita*. Lisboa: Faculdade de Psicologia da Universidade de Lisboa e Câmara Municipal de Lisboa, 2017.

WAGNER, K. et al. Exploration of new complexity metrics for curriculum-based measures of writing. *Assessment for Effective Intervention*, Califórnia, v. 44, n. 4, p. 256-266, maio 2018.

WECHSLER, D. *Wechsler Individual Achievement Test (WIAT-II UK)*. London: Pearson Education Ltd, 2005.

ANEXOS

Anexo 1- Medida de Avaliação da Composição Escrita.

Nível	Critério	Descrição
Avaliação da qualidade do texto definida pela estrutura e organização, conteúdo, vocabulário, audiência, convenções gramaticais e originalidade.		
Estrutura e Organização		
4	Muito boa	Texto bem estruturado (exemplo: introdução, desenvolvimento e conclusão adequados).
3	Boa	Texto com estrutura parcelar (exemplo: organização incompleta, linear, ou ordem cronológica linear).
2	Suficiente	Texto só com uma estrutura aparente (exemplo: conteúdos surgem aglutinados, não sendo enumerados, sem fio condutor explícito).
1	Inadequada	Texto sem sequência lógica, coerente nem compreensível.
Conteúdo		
4	Muito bom	Inclui frequentemente informação coerente, articulada e completa para a definição de personagens, locais, tempo e acontecimentos.
3	Bom	Inclui alguma informação coerente, articulada e completa para a definição de personagens, locais, tempo e acontecimentos (exemplo: informação específica e parcelar).
2	Suficiente	Inclui pouca (ou ocasionalmente) informação coerente, articulada e completa para a definição de personagens, locais, tempo e acontecimentos (exemplo: domínio de aspetos acessórios, marginais, irrelevantes).
1	Insuficiente	Não inclui informação coerente, articulada e completa para a definição de personagens, locais, tempo e acontecimentos (exemplo: informação incoerente, desarticulada, incompleta e aspetos que não se adequam ao tema escolhido).
Vocabulário		
4	Muito Rico	Inclui frequentemente vocabulário temático muito diversificado (exemplo: termos menos vulgares).
3	Rico	Inclui vocabulário temático diversificado (exemplo: com tentativas de uso menos comum).
2	Corrente	Inclui pouco (ou ocasionalmente) vocabulário temático e diversificado (exemplo: vocabulário adequado mas previsível e comum, geralmente sem repetições).
1	Limitado	Não inclui algum vocabulário temático e diversificado (exemplo: pouco diversificado, muito pobre e repetitivo).
Audiência		
4	Muito boa	Demonstra um forte sentido de audiência (i.e., adapta a escrita às características do leitor), revelando uma boa consciência das expectativas do leitor e proporcionando o seu envolvimento com o texto.
3	Boa	Demonstra sentido de audiência, revelando uma boa consciência das expectativas do leitor.
2	Suficiente	Demonstra um reduzido sentido de audiência, revelando alguma consciência das expectativas do leitor.
1	Insuficiente	Não demonstra evidência sobre o sentido de audiência (texto muito curto).
Convenções gramaticais		
4	Muito boa	Respeita com aparente facilidade as convenções gramaticais, os mecanismos e uso do Português escrito, podendo apresentar alguns erros menores.
3	Boa	Respeita as convenções gramaticais, mecanismos e uso do Português escrito, podendo apresentar alguns erros que não interferem na construção de sentido.
2	Suficiente	Apresenta ocasionalmente erros graves ou com alguma frequência erros menores na construção gramatical, nos mecanismos e uso do Português escrito, que podem interferir na construção de sentido.
1	Insuficiente	Apresenta erros graves na construção gramatical, nos mecanismos e uso do Português escrito, que interferem na construção de sentido.
Originalidade		
4	Elevada	Inclui frequentemente ideias e palavras de forma muito original, com recursos inovadores (exemplo: recursos estilísticos (e.g., comparações, metáforas), humor, envolvimento emocional, entre outras).
3	Moderada	Inclui algumas ideias e palavras de forma original com recursos inovadores (e.g., comparações, metáforas), humor, envolvimento emocional, entre outras).
2	Limitada	Inclui poucas (ou ocasionalmente) ideias e palavras de forma original com recursos inovadores (e.g., comparações, metáforas), humor, envolvimento emocional, entre outras).
1	Inexistente	Não inclui ideias e palavras de forma original nem recursos inovadores (e.g., comparações, metáforas), humor, envolvimento emocional, entre outras).
Comentários ou anotações:		