

---

# arquivos analíticos de políticas educativas

Revista acadêmica, avaliada por pares,  
independente, de acesso aberto, e multilíngue



aape || epaa

Arizona State University

---

Volume 31 Número 66

6 de junho de 2023

ISSN 1068-2341

---

## A Avaliação Pedagógico-Curricular no Ensino Primário: Uma Revisão Sistemática da Literatura

*Pedro Duarte*

Escola Superior de Educação do Politécnico Porto  
Portugal



*Dora Fonseca*

Departamento de Educação e Psicologia da Universidade de Aveiro  
Portugal

**Citação:** Duarte, P., & Fonseca, D. (2023). A avaliação pedagógico-curricular no ensino primário: Uma revisão sistemática da literatura. *Arquivos Analíticos de Políticas Educativas*, 31(66).  
<https://doi.org/10.14507/epaa.31.7875>

**Resumo:** O artigo corresponde a uma revisão sistemática da literatura, em que se procura analisar como se tem estudado a relação entre a educação primária e a avaliação pedagógico-curricular e quais são os principais resultados dessas investigações. Para a análise, consideraram-se 61 artigos com presença na Scopus e/ou na Web of Science. Enquanto principais resultados, destacam-se três ideias. A investigação privilegia trabalhos de natureza predominantemente técnica, através da validação de instrumentos. Existe uma certa desconsideração ontológica do ensino elementar, fazendo com que a investigação reforce lógicas de fragmentação curricular, deste contexto formativo. As políticas e práticas de investigação parecem não conseguir estabelecer um discurso contra-hegemónico que melhor permite analisar, problematizar e criticar as políticas educativas dominantes.

**Palavras-chave:** avaliação pedagógico-curricular; ensino primário; políticas educativas

### **Pedagogical-curricular assessment in primary education: A systematic review of the literature**

**Abstract:** This article presents a systematic review of the literature, which seeks to analyze how the relationship between primary education and pedagogical-curricular assessment has been studied and what are the main results of these investigations. For the analysis, we considered 61 articles in Scopus and/or Web of Science. As main results, three major findings stand out. The research privileges papers of a predominantly technical nature, through the validation of instruments. There is a certain ontological disregard of elementary education, which makes research reinforce curricular fragmentation logics of this formative context. The policies and practices of research seem to fail to establish a counter-hegemonic discourse that better allows us to analyze, problematize and criticize the dominant educational policies.

**Keywords:** pedagogical-curricular evaluation; primary education; educational policies

### **Evaluación pedagógico-curricular en la educación primaria: Una revisión sistemática de la literatura**

**Resumen:** El artículo corresponde a una revisión sistemática de la literatura, que busca analizar cómo se ha estudiado la relación entre la educación primaria y la evaluación pedagógico-curricular y cuáles son los principales resultados de estas investigaciones. Para el análisis, se consideraron 61 artículos en Scopus y/o Web of Science. Como principales resultados, destacan tres grandes conclusiones. La investigación privilegia los trabajos de carácter predominantemente técnico, a través de la validación de instrumentos. Hay un cierto desconocimiento ontológico de la educación primaria, lo que hace que la investigación refuerce las lógicas de fragmentación curricular de este contexto formativo. Las políticas y prácticas de investigación parecen no establecer un discurso contrahegemónico que permita analizar, problematizar y criticar mejor las políticas educativas dominantes.

**Palabras-clave:** evaluación pedagógico-curricular; educación primaria; políticas educativas

## **A Avaliação Pedagógico-Curricular no Ensino Primário: Uma Revisão Sistemática da Literatura**

Reconhecemos, hoje, um conjunto muito amplo de desafios e singularidades que caracterizam e marcam o tempo contemporâneo, com inegáveis influências nos sistemas e processos educativos. Sem assumirmos o propósito de apresentar um conjunto detalhado destas características, importa, ainda assim, destacar algumas delas, estabelecendo, desde logo, a sua relação com a educação formal. Para a presente introdução, destacamos, inicialmente, a crescente valorização de aspetos relacionados com a eficácia e a eficiência das distintas instituições contemporâneas, generalizando-se a necessidade de entender o funcionamento das organizações de acordo com o funcionamento económico-empresarial, estabelecendo-se “uma *ontologia empresarial*, na qual é simplesmente óbvio que tudo na sociedade deve ser gerido como uma empresa” (Chomsky & Waterstone, 2021, p. 76). Para Morin (2011), tal traduz-se numa *conceção tecnoeconómica* que reduz a reflexão à afirmação de um único modelo de funcionamento social e económico, cego perante as particularidades culturais, sociais e humanas, a partir de um entendimento estandardizado do modo como a sociedade deveria funcionar. Nesse sentido, o pensamento político restringe-se às dimensões técnico-instrumentais e desconsidera, por isso, outras alternativas, outros modos de entender a realidade ou de conceber formas de organização e atuação.

Quando mobilizamos este referencial para as dinâmicas educativas, sobressai um reducionismo técnico de análise e intervenção educativas. Tal postura coincide com a ideia de neutralidade ou tecnocracia das decisões curriculares e didáticas nos diferentes sistemas educativos (Lissovoy, 2022). Como consequência, os docentes são reduzidos às suas funções técnicas, entendidos como funcionários que cumprem, acriticamente, *rituais formalizados*, desconsiderando (i) os alicerces pedagógicos e éticos das opções tomadas, (ii) o modo como os docentes podem implicar-se em processos de investigação e contribuir para a (re)construção do conhecimento educativo, (iii) as potenciais repercussões político-sociais das ações formativas (Giroux, 2022). Nesta lógica de pensamento, surgem, em diferentes contextos geográficos, estruturas que contribuem para a valorização de dimensões como a performatividade docente e a uniformização da sua prática letiva. Tais opções facilitam a formulação de *currículos à prova de professores* e/ou a proliferação de *packs* didáticos, que contribuem para cingir a docência aos seus aspetos técnico-metodológicos, que não necessitam de ser ponderados em relação aos contextos da ação profissional ou às particularidades de professores e alunos (Duarte, 2021).

Como segundo aspeto, mencionamos a legitimação de uma “barbárie fria e gelada da hegemonia do cálculo, do quantitativo, da técnica, do lucro sobre as sociedades e as vidas humanas” (Morin, 2011, p. 29), que dialoga, de forma muito próxima, com os dois outros pontos já referidos. Talvez a ideia mais ilustrativa desta tendência surja pelas palavras de O’Neil (2016): *Armas de Destruição Matemática (Weapons of Math Destruction)*. O’Neil (2016) alerta-nos para a propagação de algoritmos nas diferentes esferas da vivência social, como na regulação financeira, na gestão dos serviços públicos, na divulgação de notícias, na organização das redes sociais, nos processos de recrutamento, entre muitos outros. Por detrás desta lógica está a perspectiva de que é possível estabelecerem-se algoritmos de estudo matemático que possibilitam análises objetivas e rápidas de um número muito elevado de dados, sobretudo para favorecer a eficácia das decisões tomadas. Há, contudo, alguns elementos que importa criticar. Por um lado, muitos destes algoritmos são opacos, isto é, não se torna transparente nem o algoritmo considerado para a análise, nem os dados que foram considerados (ou o modo como foram recolhidos). Por outro lado, a pretensa objetividade destes algoritmos esconde que os modelos matemáticos construídos decorrem sempre de opções, mais ou menos subjetivas, dos seus criadores, e que concorrem, sempre, para resultados que simplificam o fenómeno, porque necessitam de estabelecer aproximações.

A dimensão social remete para fenómenos altamente complexos, com um conjunto muito elevado de variáveis, pelo que as criações de modelos conduzem a explicações mais pobres e simplificadas da realidade – mais ainda, certos algoritmos podem integrar, de forma implícita, algum tipo de preconceito, o que mais os fragiliza. Depois, os mesmos algoritmos necessitam de um número muito elevado de dados e de processos de *feedback*, para que possam ter algum tipo de valor – que podemos, por exemplo, encontrar nas várias plataformas digitais de distribuição de conteúdo, que recolhem várias informações e as confrontam, de forma sistemática, com novos dados que estão constantemente a ser renovados –, algo que dificilmente pode ser mimetizado nas esferas sociais, como é o caso da área da educação. Sobre este contexto, O’Neil (2016) traz-nos dois exemplos que têm particular interesse para o presente trabalho: (i) a avaliação de professores e (ii) a avaliação das organizações escolares, através dos *rankings*. Nas duas situações, a autora mostra que os algoritmos construídos para avaliar professores e instituições educativas conduziram a análises superficiais das dinâmicas formativas e, por esse motivo, a processos que empobreceram as opções pedagógicas, curriculares e organizacionais. A par do indicado, estes modelos ocasionaram um reforço de lógicas como *teaching to the test*, a seleção dos estudantes, a inflação de notas, o florescimento do mercado educativo (como explicadores, editoras, instituições de apoio às candidaturas ao ensino superior, etc.), reforçando a face tecnocrática da avaliação nas organizações educativas, inclusive na dimensão da avaliação pedagógico-curricular, nos diferentes níveis educativos.

Procuramos, através do presente estudo, avançar com um olhar mais detalhado sobre a educação primária<sup>1</sup> e a sua relação com a avaliação. Nesse sentido, estabelecemos a seguinte questão de investigação: *Como se tem estudado a relação entre a educação primária e a avaliação pedagógico-curricular e quais são os principais resultados dessas investigações?*

Assim, o artigo encontra-se organizado em 5 secções distintas. Na primeira e na segunda, avançamos com o posicionamento conceptual sobre a educação primária e a avaliação pedagógico-curricular, respetivamente. Na secção subsequente, esclarecemos os objetivos de investigação considerados, explicamos os alicerces metodológicos associados à revisão da literatura conduzida e identificamos os critérios de validade considerados. Na quarta secção, avançamos com a análise do *corpus* de artigos definido, para que, na quinta e última secção, possamos explicitar os principais resultados alcançados.

## **Educação Primária: Breves Apontamentos**

É importante, desde já, assumirmos que não existe um referencial único associado à educação primária. Mesmo que, como se evidencia pelos dados da OCDE (2021), se note que integra, predominantemente, profissionais femininos – apenas 5 países que compõem esta organização (Dinamarca, Grécia, Índia, Japão e Turquia) têm uma taxa inferior a 75% de professoras neste nível de ensino –, não é adequado assumir arquétipos genéricos sobre este contexto formativo, nas suas dimensões organizacionais, profissionais, curriculares ou pedagógicas. Efetivamente, a educação primária é marcada por uma pluralidade de opções políticas. A título de exemplo, na Bélgica e em Espanha, a educação primária contempla os 6 primeiros anos da educação obrigatória, na Alemanha e em Portugal, por seu turno, fica apenas associada aos 4 primeiros anos – no que se denomina de 1º ciclo do ensino básico –, já em França e Itália contempla 5 anos de escolaridade (Eurydice, 2021). Apesar de assumirmos esta diversidade, apresentamos, nos parágrafos seguintes, algumas ideias que procuram avançar com um (incompleto) retrato deste nível de ensino. Como ponto de partida para esta discussão, sugerimos tomar como referência as indicações assumidas pela Eurydice e pela OCDE:

Os programas a este nível são tipicamente concebidos para fornecer aos estudantes competências fundamentais em leitura, escrita e Matemática (isto é, literacia e numeracia) e estabelecer uma base sólida para a aprendizagem e compreensão de áreas nucleares do conhecimento, desenvolvimento pessoal e social, em preparação para o ensino secundário inferior.

A idade é tipicamente o único requisito de entrada a este nível. A idade habitual ou legal de entrada não é normalmente inferior a 5 anos ou superior a 7 anos. Este nível dura tipicamente seis anos, embora a sua duração possa variar entre quatro e sete anos. (Eurydice, 2021, p. 11, tradução própria)

Concebido para proporcionar uma sólida educação básica em leitura, escrita e Matemática e uma compreensão básica de algumas outras disciplinas. Idade de entrada: entre os 5 e 7 anos. Duração típica: seis anos. (OCDE, 2021, p. 21, tradução própria)

---

<sup>1</sup> Ao longo do texto, tomaremos as ideias de “ensino primário”, “educação primária”, “ensino elementar” e “educação elementar” como sinónimas. Em Portugal, o ensino primário é designado 1º Ciclo do Ensino Básico. Igualmente assumiremos “professor”, “docente” e “educador” como sinónimos.

Assumimos que, de forma genérica, estas concepções convergem com aquelas que são as perspectivas tidas pela generalidade da população. Nele sobressai a ideia de vincular o ensino primário à preparação para o nível educativo subsequente e ao ensino da literacia e na numeracia, como a principal finalidade. Essa perspectiva é, ainda, corroborada pelos próprios dados da OCDE (2021) que esclarecem que, na média dos países que integram esta organização, 25% do horário letivo é dedicado à leitura, escrita e literatura e 17% dedicado à Matemática, sobressaindo de forma evidente em relação às restantes componentes curriculares<sup>2</sup>.

Não subscrevemos, contudo, este entendimento, dado que corresponde a um certo reducionismo das características deste nível de ensino, que, por um lado, não integra aquelas que são as suas efetivas particularidades e, por outro, desconsidera as suas finalidades mais abrangentes e integradoras. Explicando por outras palavras, as citações mobilizadas não evidenciam qualquer reconhecimento das características específicas da educação primária, assumindo, antes, uma definição a partir das suas prioridades curriculares e das faixas etárias que integra. Por esse motivo, será insuficiente assumir os critérios considerados, dado que não mostram o modo como a educação primária se distingue e autonomiza dos níveis escolares subsequentes.

Como ponto de partida para esta reflexão, salientamos o conhecimento que se relaciona com este contexto formativo. Como têm discutido diferentes autores (Bower, 2021; Coltman & Rolls, 2021; Wyse, 2002), o ensino primário integra um conjunto plural de conhecimentos e competências estruturantes e fundacionais que servem de alicerce para aprendizagens posteriores. Contudo, tal alicerce não deve ser confundido com uma aprendizagem simples (ou simplista). Como tal, é necessário reconhecer e valorizar a amplitude e diversidade de saberes que se estabelecem como de suma importância no ensino primário, como a educação plástica, a educação dramática, a educação histórica, a educação no domínio das tecnologias digitais, da literacia informática, entre outros (Campbell, 2005; Wyse, 2002). Este aspeto dialoga com um outro elemento que tem marcado este contexto formativo, o compromisso deste nível de ensino com a educação integral da criança e, por isso, a sua profunda relação com a educação social e ética dos educandos. Hayes (2010) esclarece que os docentes do ensino primário necessitam de tomar decisões que não só favorecem a formação académica dos estudantes, como visam auxiliá-las a “tornarem-se sabias e moralmente perspicazes nas suas decisões” (p. 285). Esta postura converge com as ideias de Coltman e Rolls (2021), que defendem a necessidade de a educação primária se implicar na educação da criança completa (“*whole child*”), integrando aspetos como a valorização da curiosidade, da criatividade, do pensamento complexo, da autonomia, dos valores e do gosto pela aprendizagem e pelo conhecimento.

O que vimos no parágrafo anterior concorre com a ideia de que, na educação primária, importa privilegiarem-se opções pedagógico-curriculares mais integradas, que contribuam para o crescimento e maturação das crianças de forma global. Esta perspectiva facilita o entendimento da educação elementar como um nível de ensino onde há uma complementaridade entre os vários domínios, em que os conhecimentos científicos, artísticos, sociais, estéticos e éticos interagem de forma coerente e enriquecedora (Bower, 2021; Coltman & Rolls, 2021). Apesar desse referencial, reconhecemos que existe, hoje, uma tendência de disciplinarização deste contexto formativo, que tanto se expressa nas opções político-educativas, como na definição dos horários, e consequente prática letiva, ficando o tempo escolar segmentado, em períodos predefinidos para cada uma das componentes curriculares (Campbell, 2005). Esta postura é, contudo, recorrentemente criticada, porque, por um lado, promove uma maior fragmentação curricular e maior dificuldade em promover um conhecimento articulado entre as diferentes áreas (Morin, 2011) e, por outro, descaracteriza este nível de ensino, que mais se aproxima dos formatos e práticas de outros níveis formativos

---

<sup>2</sup> Educação Física assume 10% da carga horária, Artes 9%, Ciências Naturais 7%, aprendizagem de uma segunda língua 7%, Ciências Sociais 6%, entre outros.

(Campbell, 2005; Hayes, 2010; Wyse, 2002). Numa lógica contrária ao identificado, podemos assumir uma postura mais próxima das ideias de articulação ou integração curricular (Duarte, 2021; Torres Santomé, 2015). De acordo com esta perspectiva, o ensino primário é entendido como um contexto formativo onde as diferentes componentes curriculares poderão ser trabalhadas de forma harmoniosa, por vezes em simultâneo, para facilitar uma maior relação entre as distintas áreas da formação humana. De forma complementar, salientamos o valor atribuído ao património e comunidade local como fatores que contribuem para o enriquecimento e complexificação das aprendizagens e, ainda, ao modo como as práticas pedagógicas podem favorecer o desenvolvimento da aprendizagem democrática e da educação em valores.

Para convergir com esses propósitos, salientamos outra característica associada a este nível de ensino: a monodocência ou, pelo menos, a existência de professores generalistas. De acordo com esta opção, em vez de um professor para cada uma das componentes curriculares, um professor que assume a totalidade ou a maioria delas. Mesmo que essa opção não garanta um trabalho efetivo de articulação/integração curricular, pode favorecer condições que facilitam esse tipo de dinâmicas formativas e uma gestão mais flexível do currículo e das atividades pedagógicas. A par do indicado, este perfil profissional fica associado a uma relação continuada entre educadores e educandos, o que contribui, simultaneamente, para o maior envolvimento (que, também, poderá traduzir-se num maior desgaste) afetivo e emocional dos docentes em relação às crianças, aos encarregados de educação e à comunidade local (Hayes, 2010; Wyse, 2002).

Para finalizar, importa avançarmos com um apontamento específico sobre a avaliação, que melhor discutiremos na secção seguinte. Partindo desta ideia de docente generalista, reconhecemos que estes profissionais têm condições que propiciam uma avaliação contínua, através de processos sistemáticos e variados de recolha e análise de dados, que poderão contribuir para dinâmicas avaliativas mais sustentadas, diversificadas e formativas, como portefólios, processos de co-avaliação, trabalhos de projeto, exercícios escritos desenvolvidos em aula, apresentações, entre outros (Duarte, 2021; Wyse, 2002). Talvez por isso, e como refere a *European Commission* (2020), para a realidade europeia, a probabilidade de a avaliação assumir funções formativas é superior no ensino primário, como, igualmente, se assume uma maior predisposição deste nível educativo para acolher mudanças no âmbito da avaliação.

## Enquadramento Metodológico

Como esclarecemos nos parágrafos introdutórios, o presente estudo estabelece-se como uma revisão sistemática da literatura que, como identificam Cohen, Manion e Morrison (2018), corresponde a uma opção metodológica própria, com validade científica reconhecida. Nesta investigação, e com influências de posturas que poderão ser denominadas como *síntese interpretativa* e *síntese crítica* (Suri, 2018), assumimos a seguinte questão de investigação: *Como se tem estudado a relação entre a educação primária e a avaliação pedagógico-curricular e quais são os principais resultados dessas investigações?*

Decorrente desta finalidade, estabelecemos os seguintes objetivos de investigação:

- (i) mapear as opções metodológicas das investigações que articulam a educação primária com a avaliação pedagógico-curricular;
- (ii) compreender as principais finalidades assumidas nos estudos que articulam a educação primária com a avaliação pedagógico-curricular;
- (iii) destacar os principais domínios científicos e alicerces conceptuais nas investigações que articulam a educação primária com a avaliação pedagógico-curricular;

- (iv) perceber os principais resultados das investigações que possibilitem melhor compreender a relação entre educação primária e avaliação pedagógico-curricular.

Para clarificarmos as escolhas metodológicas consideradas, importa problematizar um pouco mais o modo como decorrem as revisões sistemáticas de literatura. De acordo com Denscombe (2014), este tipo de investigação é utilizado para um conjunto amplo de propósitos científicos, sociais e políticos. Como explica Suri (2018), estes estudos foram integrando, paulatinamente, maior diversidade de finalidades investigativas e de opções de análise. Esclarece a autora que, de acordo com as posições *pós-positivistas*, privilegiam-se trabalhos que visam explicar objetivamente o fenómeno em estudo, recorrendo, para isso, a métodos predominantemente estatísticos, que visam estabelecer leis gerais ou relações mensuráveis entre múltiplas variáveis. Porém, outras posturas são, atualmente, consideradas. As *sínteses interpretativas*, afastando-se da ideia de definição objetiva e factual do real, procuram criar condições para um entendimento mais autêntico e aprofundado do fenómeno, reconhecendo que ele é interpretado de modos diferentes por distintos agentes ou grupos e que tal também se reflete na análise realizada. Por seu turno, o paradigma *crítico* contribui para que, através da revisão da literatura, se discutam modos como determinados conhecimentos e normas são privilegiados, tomando em consideração os conceitos e discursos que se encontram presentes, ou ausentes, da investigação, procurando, assim, uma problematização dos modos como se investiga.

A partir desta pluralidade de posicionamentos, foi igualmente possível estabelecer formas mais diversas de conduzir este tipo de estudos. Nesse sentido, se a tradição pós-positivista privilegiava modos predominantemente estatísticos de análise da investigação primária, hoje são já reconhecidas e aceites outras formas de discussão dos dados, validando tanto análises de natureza qualitativa, como análises de natureza mista (Cohen et al., 2018). Para a presente investigação, inspiramo-nos nas propostas de Finfgeld-Connett (2014) e desenvolvemos processos de codificação do *corpus* de artigos considerados para a investigação. Este processo decorreu de forma indutiva, já que os códigos emergiram dos documentos lidos, sem existir uma predefinição de modos de análise, isto é, sem existir o que a autora denomina como *coding template*. Posteriormente, desenvolvemos uma análise mista, integrando tanto modos quantitativos de discussão dos dados (essencialmente estatística descritiva), como modos qualitativos de interpretação, privilegiando o diálogo com outros referenciais teóricos. Esta opção converge com o conceito de *análise narrativa* (Denscombe, 2014), pela qual os investigadores “procuram “contar a História” por trás dos resultados dos vários trabalhos que foram revistos” (p.140, tradução própria).

Importa, aqui, destacar que esta escolha necessita, igualmente, de critérios de validade. Com base no trabalho de Cohen e colaboradores (2018), referimos três modos complementares de procurar garantir a qualidade desta investigação. Como primeiro aspeto, destacamos exatamente a articulação entre abordagens quantitativas e qualitativas que, para os autores, reforça a validade do estudo. Como segundo, salientamos a transparência na identificação da questão e objetivos de investigação e a tentativa de articular, de forma coerente, os processos de seleção e análise do *corpus* de artigos. E, como último esforço de validação do estudo, fazemos referência à validade teórica, associada ao esclarecimento (prévio) dos alicerces conceptuais considerados para a investigação e, ainda, à integração, sistemática, desses contributos teóricos na análise e discussão dos dados.

Em complementaridade aos três elementos referidos, destacamos o uso do software MaxQDA2018. Como refere Finfgeld-Connett (2014), a utilização deste tipo de programas, em particular para amostras com um número significativo de artigos, pode ser uma ferramenta relevante para a qualidade do trabalho. Por fim, assumimos as indicações avançadas por Denscombe (2014), que identifica a importância de se explicitarem as formas de recolher e selecionar os artigos a serem

analisados, identificando, de modo transparente, a quantidade de estudos integrados na investigação e os seus critérios de inclusão e exclusão.

Na sequência do indicado no parágrafo anterior, esclarecemos que, para a composição inicial do *corpus* de texto, foram consideradas duas bases de dados, a *Scopus* e a *Web of Science*. Definimos os seguintes critérios de inclusão:

- os anos de publicação serem 2021 e 2022 (até 30 de agosto de 2022);<sup>3</sup>
- a presença das palavras “elementary” ou “primary” e “assessment” ou “evaluation” nas palavras-chave.

A partir destes critérios de inclusão, definimos as seguintes entradas booleanas, na *Scopus* e na *Web of Science*, respetivamente:

```
(AUTHKEY (“assessment” “primary”) OR AUTHKEY (“assessment”
“elementary”) OR AUTHKEY (“evaluation” “primary”) OR AUTHKEY
(“evaluation” “elementary”)) AND (LIMIT-TO ( SUBJAREA, “SOC1” )) AND
(LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021 )) AND
(LIMIT-TO (DOCTYPE, “ar” ))
```

```
(((((AK=(“evaluation” “elementary”)) OR AK=(“assessment” “elementary”)) OR
AK=(“evaluation” “primary”)) OR AK=(“assessment” “primary”)) AND
WC=(Education & Educational Research)) AND DT=(Article)) AND DOP=(2021-
01-01/2022-08-21)
```

Após essa pesquisa, foram identificados 283 artigos. Deste *corpus*, procedeu-se à remoção dos duplicados, obtendo-se um total de 199 artigos de análise. Como fase seguinte à escolha dos documentos, tradicionalmente procede-se a uma análise dos resumos, para uma seleção mais fina dos artigos a considerar. Contudo, à medida que desenvolvíamos essa leitura, constatamos que a leitura, exclusiva, dos resumos era insuficiente para compreender a pertinência dos documentos, tendo em conta o foco da investigação. Procedemos, então, a uma análise parcelar inicial que, a par do resumo, incluía as decisões metodológicas dos trabalhos. Para tal, consideramos os seguintes critérios (complementares aos anteriormente identificados):

- os estudos têm que corresponder a investigações primárias, isto é, têm que incluir a recolha e análise de dados empíricos;
- os artigos têm que se encontrar escritos em castelhano, inglês ou português;
- a avaliação pedagógico-curricular tem que ser um domínio de investigação, mesmo que não seja o principal;
- a investigação tem que ser realizada especificamente para o ensino primário;
- o estudo enquadra-se no âmbito da investigação em educação.

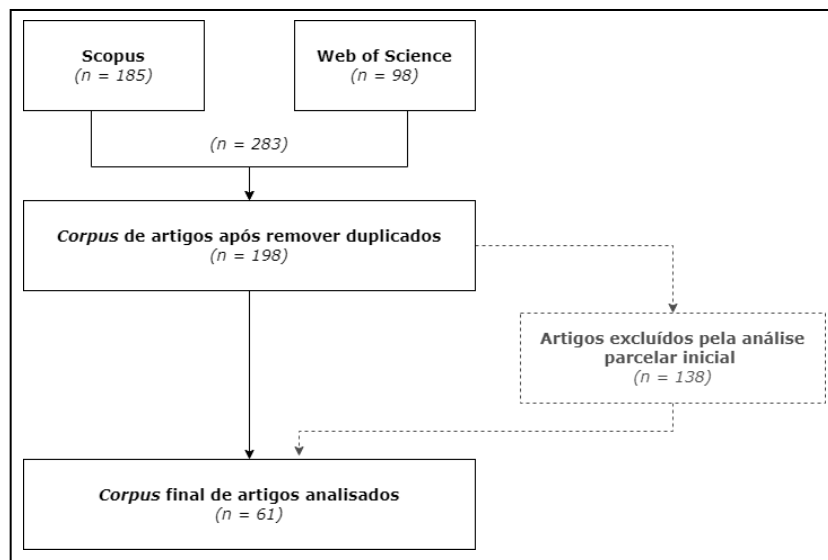
Os critérios acima assinalados possibilitaram a exclusão de 138 artigos, de acordo com a seguinte distribuição (Tabela 1). De forma a sintetizar o procedimento adotado, pode ler-se o Figura 1.

<sup>3</sup> Tomou-se este período de publicação por dois motivos: i) a quantidade de documentos existentes que obrigou à definição de um período relativamente reduzido, para que fosse possível a análise; ii) possíveis relações da avaliação com as alterações com o COVID-19, embora tal, depois, não viesse a evidenciar ser um domínio emergente dos trabalhos analisados.



**Tabela 1***Frequência absoluta e relativa dos critérios de exclusão*

Critérios de Exclusão	Frequência Relativa	Frequência Absoluta
<b>1. Aspectos Formais</b>	<b>13</b>	<b>9,4%</b>
1.1 Trabalho sem dados empíricos	3	2,2%
1.2 Ano de publicação original não corresponde ao período temporal definido	7	5,1%
1.3 O artigo não se encontra escrito nas três línguas consideradas	2	1,4%
1.4 Não contempla as palavras-chave definidas	1	0,7%
<b>2. Diferentes focos de investigação em educação</b>	<b>86</b>	<b>62,3%</b>
2.1 A avaliação não é um eixo de investigação do estudo	25	18,1%
2.2 Não é específico para a educação primária/elementar	14	10,1%
2.3 O estudo encontra-se centrado na formação inicial de professores	13	9,4%
2.4 O estudo encontra-se centrado na formação contínua e/ou no desenvolvimento profissional dos professores	20	14,5%
2.5 Domínios particulares da investigação em educação	13	9,4%
2.5.1 Educação Inclusiva	9	6,5%
2.5.2 Avaliação Institucional	4	2,9%
2.5.3 História da Educação	1	0,7%
<b>3. Diferentes campos científicos</b>	<b>39</b>	<b>28,3%</b>
3.1 Medicina	25	18,1%
3.3 Ecologia	6	4,3%
3.4 Psicologia	6	4,3%
3.5 Ciência Política	1	0,7%
3.6 Sociologia (da intervenção policial)	1	0,7%

**Figura 1***Diagrama dos artigos considerados*

Partindo deste processo, e como já identificamos, foram considerados 61 artigos para análise. Destes, 38 estudos foram publicados em 2021 e 23 publicados em 2022, conforme ilustrado na Tabela 2.

**Tabela 2***Identificação dos artigos analisados, de acordo com o ano de publicação*

Ano de publicação	Artigos analisados
<b>2021</b>	Afandi et al., 2021; Allal, 2021; Amendum et al., 2021; Basu et al., 2021; Buckley-Walker & Lipscombe, 2021; Butler et al., 2021; Chin et al., 2021; Earle, 2021; Fitzgerald et al., 2021; Gane et al., 2021; Gezer et al., 2021; Glušac & Milić, 2021; Herbert, 2021; Hogenboom et al., 2021; Ibourk & Kendrick, 2021; Jong & Tan, 2021; Julião et al., 2021; Kalykbayeva et al., 2021; Kim & Lee, 2021; McFeetors et al., 2021; Mitsiakiet al., 2021; Mohamed et al., 2021; Monteiro et al., 2021; Nicholas et al., 2021; Omasta et al., 2021; Ordaz Villegas & Aclé Tomasini, 2021; Pedasteet al., 2021; Roberts & Brugar, 2021; Sayac & Veldhuis, 2021; Schelling & Rubenstein, 2021; Schmitzet al., 2021; Syzdykbayeva et al., 2021; Tirado-Olivares et al., 2021; Torregrosa et al., 2021; Turner et al., 2021; van den Heuvel-Panhuizen et al., 2021; Wammes et al., 2021; Wikman et al., 2021
<b>2022</b>	Amador et al., 2022; Bellindo-García et al., 2022; Carrillo López & Hortigüela-Alcalá, 2022; Chen et al., 2022; Chin et al., 2022; Chin & Chew, 2022; El-Hamamsy et al., 2022; Eriksson et al., 2022; Gamboa-Araya et al., 2022; Graham et al., 2022; Herwin et al., 2022; Knoester & Meshulam, 2022; Martin et al., 2022; McFadden et al., 2022; Rogers et al., 2022; Ryan et al., 2022; Skar et al., 2022; Takele & Melese, 2022; Tárraga-Mínguez & Tarín-Ibáñez, 2022; Tsai et al., 2022; Unahalekhaka & Bers, 2022; Wammes et al., 2022; Wilson et al., 2022

Relativamente à língua de publicação, 52 foram escritos em inglês, 6 em castelhano e 3 em português. Os estudos realizados dizem respeito a 29 países distintos (confrontar Anexo 1), com particular incidência nos Estados Unidos (Fa=17), na Malásia (Fa=5), na Austrália (Fa=4) e na Espanha e Países Baixos (Fa=3).

O corpus de documentos integra 45 revistas científicas diferentes (confrontar Anexo 2), com destaque para a *Assessment in Education: Principles, Policy&Practice* (Fa= 4) e quatro outras revistas das quais se consideraram 3 artigos de cada: *Computer Science Education*, *Education Sciences*, *Frontiers in Education* e *International Journal of Science and Mathematics Education*.

### Análise dos Dados

Como ponto de partida, focar-nos-emos naquelas que são as finalidades empíricas dos diferentes trabalhos de investigação considerados. Conforme se constata pela Tabela 3, há uma prevalência de estudos que visam melhor compreender as práticas de avaliação consideradas na educação elementar e, ainda, desenvolver algum tipo de trabalho de validação de uma técnica ou instrumento de avaliação – importa assinalar que existem artigos que integram mais do que uma finalidade empírica. Para compreendermos com outro pormenor os focos investigativos destes trabalhos, necessitamos, contudo, de uma análise mais detalhada.

**Tabela 3**

*Frequência absoluta e relativa dos critérios das finalidades investigativas dos artigos analisados*

	Frequência absoluta	Frequência relativa
1. Propósitos empíricos		
1.1. Compreender práticas de avaliação	32	52,5%
1.1.1 Avaliação em sentido lato	16	26,2%
1.1.2 Avaliação Formativa	11	18%
1.1.3 Avaliação Sumativa	2	3,3%
1.1.4 Avaliação Diagnóstica	2	3,3%
1.1.5 Avaliação Autêntica	1	1,6%
1.1.6 Auto-Avaliação	3	4,9%
1.1.7 Co-Avaliação	2	3,3%
1.2 Avaliar ou validar determinadas técnicas ou instrumentos de avaliação	29	47,5
1.3 Compreender perspetivas dos docentes	17	27,9%
1.4 Compreender perspetivas dos estudantes	5	8,2%
1.5 Analisar influência das provas <i>standard</i>	3	4,9%
1.6 Analisar potenciais repercussões da avaliação para o ensino	2	3,3%
1.7 Compreender propostas do currículo apresentado sobre avaliação	1	1,6%

Relativamente à primeira categoria identificada, a mesma corresponde aos artigos que assumiram o propósito de caracterizar práticas de avaliação. Como se evidencia pela tabela anterior (Tabela 2), notam-se focos plurais que mais orientam as finalidades empíricas dos estudos analisados. Na sua maioria (Fa=16), este grupo de documentos visa uma compreensão abrangente dos fenômenos de avaliação, nomeadamente perceber “como os professores avaliavam as aprendizagens dos seus alunos” (Monteiro et al., 2021, p. 1), como se procede à recolha e armazenamento dos dados provenientes da avaliação (van den Heuvel-Panhuizen et al., 2021), quais as práticas (as técnicas e os instrumentos de) avaliação consideradas (Bellindo-Garcia et al., 2022; Omasta et al., 2021), os agentes envolvidos nas estratégias avaliativas (Bellindo-Garcia et al., 2022; Buckley-Walker & Lipscombe, 2021) ou o tempo ou frequência dedicada à avaliação (Bellindo-Garcia et al., 2022; Schelling & Rubenstein, 2021), as dificuldades sentidas pelos docentes (Sayac & Veldhuis, 2021), entre outras. Neste domínio, integram-se, portanto, os estudos que visam entender as experiências avaliativas de forma ampla. Tal postura, pelo menos de forma implícita, parece dialogar com as opções de Rosales López (2014) ou Santos Guerra (2014), por exemplo, que privilegiam uma concepção particularmente ampla da avaliação educativa, reconhecendo que esta se implica em múltiplos domínios, como opções didáticas, estruturas organizacionais ou processos de desenvolvimento profissional.

Não devemos, todavia, descorar as opções que, pretendendo a compreensão dos fenômenos avaliativos, apresentam olhares mais específicos, havendo, inclusive, estudos que, simultaneamente, atentam em dimensões mais particulares da avaliação, como a avaliação formativa e a avaliação sumativa (Gezer et al., 2021) ou processos de auto-avaliação e de co-avaliação (Torregrosa et al., 2021). Sobre a totalidade destes trabalhos, constata-se um particular destaque para os processos de avaliação formativa (Fa=10), que, à semelhança do que esclarece Hayes (2010), tem vindo a ganhar progressiva importância e reconhecimento no ensino primário. Este foco na avaliação formativa adquire características muito diversas, como procurar perceber a relação entre a avaliação formativa e (i) as aprendizagens Matemáticas (Martin et al., 2022), (ii) o domínio do inglês (Mohamed et al., 2021) ou (iii) a dieta das crianças (Carrillo López & Hortigüela-Alcalá, 2021); compreender a relação entre avaliação formativa e avaliação sumativa (Gezer et al., 2021); perceber as práticas que os docentes integram nos processos de avaliação formativa e com que frequência a fazem (Schelling & Rubenstein, 2021) ou os significados atribuídos pelos estudantes a estes processos (Eriksson et al., 2022). Esta amplitude de eixos de estudo evidencia o interesse investigativo na avaliação, ilustrando, por um lado, a complexidade das ações associadas à avaliação formativa e, por outro, a diversidade destas ações nas diferentes organizações escolares. Embora com menor expressão, há, igualmente, estudos focados na autoavaliação (Allal, 2021; Ryan et al., 2022; Torregrosa et al., 2021) e co-avaliação (Allal, 2021; Torregrosa et al., 2021), na avaliação diagnóstica (Wammes et al., 2021, 2022) e na avaliação autêntica (McFeetors et al., 2021), isto é, estratégias de avaliação “concebidas de modo que haja similitude entre a tarefa de aprendizagem e a tarefa de avaliação” (McFeetors et al., 2021, p. 3). Estas opções investigativas convergem com a progressiva importância atribuída às dimensões pedagógicas de avaliação, estabelecendo-se uma mais sólida relação com o pensamento e a ação educativa (Duarte, 2021).

Apesar do reforço da avaliação *para* as aprendizagens, ainda se assume a necessidade de processos de avaliação *das* aprendizagens (Hayes, 2010). Também pelos propósitos dos estudos analisados se reconhece essa complementaridade, através da existência de trabalhos focados na avaliação sumativa (Fa=2), nomeadamente no modo como se articula com a avaliação formativa (Gezer et al., 2021) ou nas formas de escrever as questões nas provas de avaliação (Glušac & Milić, 2021).

O revelado até ao momento induz, pelo menos em parte, que os fenômenos e experiências de avaliação, no ensino elementar, se traduzem em múltiplos interesses investigativos, podendo,

consequentemente, traduzir-se numa compreensão mais plural e multiperspetivada, destacando-se uma postura que faz sobressair a dimensão pedagógica da avaliação. Tais particularidades concorrem para outros dois grupos de categorias que importa mencionar, os propósitos associados à compreensão da perspectiva dos professores ( $F_a=17$ ) e da perspectiva dos estudantes ( $F_a=5$ ). Tomando só em consideração estes valores quantitativos poder-se-ia, de algum modo, assumir as críticas avançadas por Santos Guerra (2014), quando o investigador espanhol alerta para uma certa desconsideração em relação ao modo como os estudantes vivenciam e entendem os processos de avaliação, dada a diferença relativa entre as duas categorias. Contudo, também existem estudos que se centram no modo como os estudantes conceptualizam a avaliação, como analisam as suas consequências (Butler et al., 2021), como compreendem as estratégias avaliativas (Butler et al., 2021; Monteiro et al., 2021) e as práticas de feedback (Eriksson et al., 2022) ou como entendem as emoções provocadas pela existência de provas estandardizadas (*high-stakes standardized tests*) (Knoester & Meshulam, 2022), entre outras. Estas preocupações, ainda que num número mais reduzido, ilustram o reconhecimento, por parte da comunidade científica, da importância de se atentar nas referências sobre a avaliação criadas pelas crianças.

Ainda assim, como já mencionamos, há uma maior prevalência de estudos relacionados com a perspectiva dos docentes face à avaliação. À semelhança do que acontece em relação às crianças, também estes estudos são plurais nas suas intenções ou pontos de interesse. Alguns optam por um foco disciplinar, associado, por exemplo, ao modo como os educadores perspectivam a avaliação no âmbito no ensino do inglês (Mohamed et al., 2021), da educação física (Kim & Lee, 2022) ou da educação Matemática (Gamboa-Araya et al., 2022; Martin et al., 2022; Takele & Melese, 2022); outros estudos há que privilegiam trabalhos de investigação que ora procuram melhor esclarecer o modo como, genericamente, os docentes perspectivam as práticas ou finalidades de avaliação (Bellindo-Garcia et al., 2022; Monteiro et al., 2021), nomeadamente a avaliação diagnóstica (Mitsiaki et al., 2021; Wammes et al., 2022), ora visam analisar como grupos concretos constroem estes referenciais, como professores jovens (Amador et al., 2022). Qualquer uma das opções revela, pelo menos de forma implícita, a centralidade das concepções destes profissionais nas opções de avaliação (Duarte, 2021; Rosales López, 2014; Santos Guerra, 2014), afirmando-se como uma alternativa indireta para a compreensão do fenómeno.

Ainda relativamente à pluralidade de domínios, é interessante observar as outras categorias identificadas, mesmo que, de forma agregada, correspondam a apenas 9,8% dos trabalhos analisados. Um desses estudos (Tárraga-Mínguez & Tarín-Ibáñez, 2022) estabelece uma relação direta entre a avaliação e o currículo apresentado (categoria 1.7) – isto é, os manuais escolares – evidenciando, por isso, que o pensamento e a investigação em educação, no ensino elementar, não se circunscrevem a decisões definidas pelos diferentes agentes escolares.

Em sintonia com este aspeto, salientamos os trabalhos que visaram desenvolver uma análise sobre as provas estandardizadas (categoria 1.5), procurando compreender se os fatores demográficos condicionam os resultados associados à proficiência leitora em provas de larga escala (Amendum et al., 2021), os desafios sentidos pelas crianças bilíngues provocados por tais provas de larga escala (Knoester & Meshulam, 2022) e, ainda, discutir o modo como estes instrumentos de avaliação se implicam nas decisões letivas dos professores no ensino primário (Nicholas et al., 2021). Do último grupo de estudos fazem parte as investigações que se debruçam sobre as potenciais repercussões dos processos de avaliação (categoria 1.6), nomeadamente se determinadas práticas de avaliação diagnóstica influenciam as decisões pedagógicas dos docentes (Mitsiakiet al., 2021) ou como os processos de avaliação se relacionam com a autoestima das crianças (Syzdykbayeva et al., 2021). Face ao exposto, reconhecemos que os estudos que se enquadram nestas três categorias contribuem para ilustrar como a comunidade científica entende que a avaliação se articula com distintas esferas educativas, mesmo que tal pareça não ser um eixo de investigação prioritário.

Apesar do indicado, conseguimos, em simultâneo, reconhecer uma certa tensão no seio dos artigos analisados. Quando tomamos em atenção os estudos focados na validação ou avaliação<sup>4</sup> de técnicas/instrumentos de avaliação (categoria 1.2), verificamos uma expressão muito significativa de investigações (47,5% do *corpus* analisado) que tendem a privilegiar eixos mais próximos de entendimentos burocrático-metodológicos da avaliação. Estes estudos procuram estabelecer a validade de determinadas opções avaliativas, para que possam, depois, ser consideradas pelos agentes em distintos contextos. São exemplo destes estudos investigações que pretendem (i) validar instrumentos no domínio do pensamento computacional, como o *Computational Thinking Test for Elementary School Students* (Tsai et al., 2022) e o *Computerized Adaptive Programming Concepts Test* (Hogenboom et al., 2021), ou a utilização de rubricas de avaliação para o software *Scratch Jr* (Unahalekhaka & Bers, 2022), (ii) discutir estratégias no âmbito da avaliação Matemática, nomeadamente o *Cognitive Diagnostic Assessment* com questões de escolha múltipla (Chin et al., 2021; Chin & Chew, 2022), (iii) analisar opções relacionadas com a aprendizagem da língua, focadas na escrita com o *Writing Motivation Questionnaire* (Graham et al., 2022) ou a *Writing Engagement Scale* (Rogers et al., 2022), (iv) avaliar estratégias orientadas para a aprendizagem científica com uma estratégia desenvolvida num *Web-based Inquiry Science Environment* (Ibourk & Kendrick, 2021) ou para as potencialidades avaliativas do *Science Inquiry Test for Primary Education* (Pedaste et al., 2021), ou, ainda, (v) compreender as potencialidades da gamificação para a avaliação, através da utilização de dispositivos móveis (Herwin et al., 2022) ou com um sistema de avaliação formativo gamificado (Tirado-Olivares et al., 2021). Os exemplos mobilizados não correspondem à totalidade de técnicas ou instrumentos que foram validados ou avaliados, mas permitem evidenciar a diversidade de domínios curriculares estudados e a pluralidade de métodos e instrumentos que podem ser validados. Importa, porém, destacar que este enfoque, como mencionamos já, parece privilegiar um entendimento técnico dos processos avaliativos que, pelo menos parcialmente, resultam num empobrecimento do conceito de avaliação (Santos Guerra, 2014). Para prosseguirmos com a discussão dos dados, podemos tomar como referência Tabela 4.

Numa reação quase imediata à tabela anterior, há um aspeto que se destaca: a reduzida ( $F_a=1$ ) expressão de estudos alicerçados na investigação-ação. Como esclarecem Cohen e colaboradores (2018), a investigação-ação é, em contexto educativo, uma opção metodológica que facilita o desenvolvimento profissional e a emancipação política dos diferentes agentes, dado que legitima um processo de investigação que emana dos diferentes contextos de prática. É, então, uma estratégia de investigação democrática e democratizante, que permite que os diferentes participantes nas organizações educativas se envolvam em práticas de reflexão e investigação, promotoras de melhoria e de novo conhecimento. Ao rejeitar a separação entre investigação e atuação profissional (Denscombe, 2014), esta estratégia metodológica afasta-se dos cânones mais tradicionais da investigação. Talvez esta tensão com os critérios de validade mais clássicos, assim como a exigência de um período de participação e envolvimento efetivo nos contextos de prática, ajudem a explicar a sua reduzida expressão nos artigos analisados.

---

<sup>4</sup> Ainda que a avaliação e a validação sejam focos distintos, dado o propósito do estudo agregaram-se numa única categoria. Nesse sentido, neste domínio tanto se integraram os estudos que visavam, predominantemente, validar determinada estratégia avaliativa como aqueles que procuram refletir ou criticar uma qualquer opção de avaliação.

**Tabela 4**

Relação entre os propósitos empíricos com o design metodológico

2. Design Metodológico	1. Propósitos empíricos						
	1.1 Compreender práticas de avaliação (Fa=32)	1.2 Avaliar ou validar técnicas ou instrumentos de avaliação (Fa=29)	1.3 Compreender perspectiva dos docentes (Fa=17)	1.4 Compreender perspectivas dos estudantes (Fa=5)	1.5 Analisar influência das provas standard (Fa=3)	1.6 Analisar potenciais repercussões da avaliação para o ensino (Fa=2)	1.7 Compreender propostas do currículo apresentado (Fa=1)
2.1 Estudo piloto ou amostral (Fa=31)	10	20	9	1	-	1	-
2.2 Estudo de Caso (Fa=19)	11	7	4	3	3	-	-
2.3 Estudo de Caso Múltiplo (Fa=9)	6	2	4	1	-	1	1
2.4 Investigação quase-experimental (Fa=4)	1	2	-	-	-	1	-
2.5 Design based research (Fa=2)	1	1	-	-	-	-	-
2.6 Grounded theory (Fa=2)	2	-	1	1	-	-	-
2.7 Investigação-ação (participativa) (Fa=1)	1	-	-	-	-	-	-

Em sentido contrário, as opções associadas ao estudo de caso único (Fa=19) e ao estudo de caso múltiplo (Fa=9) estão presentes em 45,90% dos artigos analisados – mesmo que, por vezes, se articulem com outras opções investigativas, como trabalhos de natureza quase-experimental (Syzdykbayeva et al., 2021; Tirado-Olivares et al., 2021) ou opções relacionadas com a *grounded theory* (Eriksson et al., 2022). À semelhança de Cohen e colaboradores (2018), destacamos que os estudos de caso (único ou múltiplos) têm ganho progressivo interesse investigativo e legitimidade científica. Dada a natureza contextual da educação, este tipo de design metodológico permite uma análise mais

aprofundada, tradicionalmente de forma triangular, de um fenómeno concreto. Sem possibilitarem generalizações, estas investigações favorecem análises complexas do que se pretende estudar, contribuindo, assim, para entendimentos progressivamente mais aprimorados das realidades e práticas educativas.

Dadas estas características, não surpreende que este tipo de design tenha sido considerado para finalidades investigativas como compreender as práticas de avaliação ou a perspectiva dos agentes (professores e alunos), coincidentes com a procura de um olhar –qualitativo e multiperspetivado – do que se estuda. Como referência final a esta opção, destacamos que os casos, nos documentos analisados, adquirem diferentes formas, como um município (Schmitzet al., 2021; Skar et al., 2022) ou distrito escolar (Gezer et al., 2021; Sayac & Veldhuis, 2021), uma organização escolar/educativa (Carrillo López & Hortigüela-Alcalá, 2022; Eriksson et al., 2022) ou uma turma (Nicholas et al., 2021), por exemplo.

Apesar do número relevante de estudos de caso, na realidade, a maioria dos trabalhos analisados (50,81%) decorre ou de trabalhos piloto ou de estudos amostrais. Embora correspondam a estudos com particularidades no seu design, a sua recolha de dados decorre da participação de um grupo restrito de indivíduos – selecionados, por conveniência ou de acordo com critérios pré-determinados. Ao contrário do estudo de caso, o foco não é estudar o caso em si, mas investigar o modo como os participantes perspetivam determinado fenómeno que se pretende estudar ou validar determinada técnica, visando, a partir daí, avançar com resultados mais genéricos e não circunscritos aos contextos investigados. Com base nestas características, distintos propósitos foram considerados, como perceber a perspectiva dos docentes (Kim & Lee, 2021; Takele & Melese, 2022) ou dos estudantes (Butler et al., 2021) em relação à avaliação. Ainda assim, e como se evidencia pela Tabela 4, a maioria dos trabalhos tinha como principal foco, através dos participantes, validar algum instrumento ou técnica de avaliação, como, por exemplo, procederam Chin e Chew (2022) ou Tsai e colaboradores (2022). Esse principal interesse torna-se mais evidente quando se constata que, dos 29 artigos que visavam avaliar ou validar alguma técnica ou instrumento, 20 deles (68,9%) integraram um design assente numa amostra ou estudo piloto, de forma isolada ou em conjunto com outras estratégias investigativas.

Para melhor se analisar o corpus de textos estudado, importa discutirmos, de forma mais atenta, as opções metodológicas consideradas. Por esse motivo, será relevante tomar em consideração a Tabela 5.

## Tabela 5

*Frequência absoluta e relativa dos critérios das finalidades investigativas opções metodológicas*

3.1. Opções metodológicas	Frequência absoluta	Frequência relativa
3.1 Constituição do <i>corpus</i> de dados		
3.1.1 Inquéritos por entrevista	22	36,1%
3.1.2 Inquéritos por questionário	21	34,4%
3.1.3 Observação	13	21,3%
3.1.4 Respostas a exercícios didáticos realizados propositadamente	12	19,7%
3.1.5 Repostas dos estudantes a provas de avaliação	10	16,4%
3.1.6 Trabalhos realizados, naturalmente, pelos estudantes	9	14,8%



3.1. Opções metodológicas	Frequência absoluta	Frequência relativa
3.1 Constituição do <i>corpus</i> de dados		
3.1.7 Rubricas, ou instrumentos similares, preenchidas pelos docentes	7	11,5%
3.1.8 Análise documental	13	21,3%
3.1.8.1 Instrumentos de avaliação criados pelos docentes	2	3,3%
3.1.8.2 Currículo apresentado	1	1,6%
3.1.8.3 Documentos curriculares institucionais	1	1,6%
3.1.8.4 Documentos curriculares nacionais	1	1,6%
3.1.8.5 Planos de aula	1	1,6%
3.1.8.6 Portefólios ou documentos similares	1	1,6%
3.1.9 Diálogos ou métodos similares de natureza conversacional informal	6	9,8%
3.1.10. Grupos focais	4	6,6%
3.2. Participantes envolvidos no estudo		
3.2.1 Estudantes	36	59,0%
3.2.2 Professores	33	54,1%
3.2.3 Grupo de Especialistas	6	9,8%
3.2.4 Lideranças Intermédias	1	1,6%

Antes de avançarmos com uma análise mais fina da tabela, é necessário recordar, uma vez mais, que os artigos integram diferentes opções metodológicas, articulando distintas técnicas numa estratégia de avaliação articulada. Essa pluralidade é assumida por vários autores, como se constata nas seguintes citações exemplificativas:

Neste estudo, foi utilizado um desenho de investigação de métodos mistos. Os dados quantitativos e qualitativos foram recolhidos utilizando inquérito por questionário, listas de verificação para a observação, e entrevistas semi-estruturadas. (Jong & Tan, 2021, p. 415, tradução própria)

Foram utilizadas múltiplas formas de dados qualitativos para captar o envolvimento dos estudantes na avaliação de desempenho, incluindo notas de campo, artefactos dos estudantes, e entrevistas. Nós escrevemos notas de campo durante e após as aulas para registar as observações do pensamento dos alunos e interações enquanto resolviam a avaliação de desempenho. Os artefactos produzidos por os estudantes durante toda a intervenção de avaliação foram simultaneamente fontes de dados. (McFeetors et al., 2021, p. 6, tradução própria)

Face ao citado, e em articulação com os dados presentes na Tabela 4, conseguimos verificar que, para a investigação sobre a avaliação no ensino elementar, é possível a integração de uma diversidade efetiva de opções metodológicas. Esta ideia converge com a perspectiva de Cohen e colaboradores (2018), quando os autores esclarecem que não existe um método ideal, na investigação em educação, sendo antes necessário estabelecerem-se estratégias que, de forma coerente, garantam a adequação entre os propósitos de investigação e os modos de recolha de dados considerados.

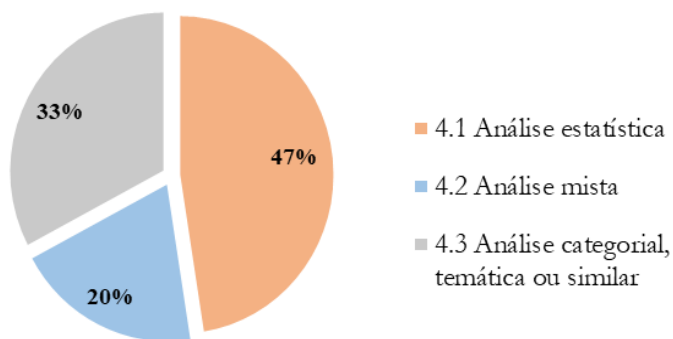
A par do identificado, também se deteta que as investigações assumiram eixos particulares de estudo, com alguns trabalhos a incidir nos documentos naturais, emergentes das organizações escolares e da prática letiva, como textos curriculares das escolas (Schmitzet al., 2021), ou da planificação curricular docente (Amador et al., 2022), nos manuais escolares (Tárraga-Mínguez & Tarín-Ibáñez, 2022), em exercícios realizados naturalmente pelos estudantes ou em instrumentos de avaliação (Glušac & Milić, 2021) e processos de feedback (Ibourk & Kendrick, 2021) considerados pelos docentes. Outros estudos seguiram opções diferentes, favorecendo uma recolha de dados específicos para as suas investigações, provenientes, por exemplo, de instrumentos de avaliação criados e realizados propositadamente para os projetos (Chen et al., 2022; Turner et al., 2021), ou ferramentas próprias de apoio à avaliação desenvolvidas para professores (Gezer et al., 2021; McFadden et al., 2022). Importa fazer referência a um conjunto relevante de artigos que integraram inquéritos por entrevista (Fa=22) ou inquéritos por questionário (Fa=21), dialogando com práticas mais abrangentes, tanto no âmbito da investigação educativa (Cohen et al., 2018), como na investigação em Ciências Sociais, em sentido mais amplo (Denscombe, 2014).

Como última nota, referimos o facto de as investigações parecerem privilegiar as dimensões de avaliação associadas ao que acontece em sala de aula. Tal pode ser entendido quando se toma em consideração a reduzida expressão de estudos (sobre avaliação pedagógico-curricular) que integram esferas de análise mais amplas, como as opções político-curriculares nacionais (van den Heuvel-Panhuizen et al., 2021), as decisões organizacionais (Schmitzet al., 2021) ou as estratégias dos grupos editoriais (Tárraga-Mínguez & Tarín-Ibáñez, 2022). A este propósito, como nos alertam Rosales López (2014) e Santos Guerra (2014), importa reconhecer que, no âmbito da avaliação, distintas esferas e campos de intervenção dialogam de forma sistemática e ecológica, em processos de influência mútua. Por esse motivo, para uma compreensão mais sustentada do fenómeno, é importante integrar uma análise destas interações de articulação entre distintos domínios (como a avaliação organizacional, os processos de controlo nacional, práticas de avaliação profissional docente, opções dos livros didáticos, entre outros).

Também quando observamos os participantes envolvidos, constatamos que só um envolve as lideranças intermédias, sobressaindo, em contrapartida, a recolha de dados junto de professores e crianças (com alguns estudos a incluírem, ainda, grupos de especialistas predominantemente para a validação de instrumentos ou recursos). Esta especificidade, pelo menos em parte, parece desviar-se das conceções de Rosales López (2014) e Santos Guerra (2014), que salientam as dimensões sistemáticas e ecológicas associadas à avaliação educativa, não reduzidas ao âmbito letivo. Prosseguindo com a análise, podemos tomar como referência o Gráfico 1.

### Gráfico 1

*Frequência relativa dos processos de análise considerados*



O gráfico é especialmente elucidativo: há uma prevalência evidente da análise estatística, verificada em quase metade de todos os projetos de investigação lidos. Na verdade, como apontam Cohen e colaboradores (2018), a gênese da investigação em educação foi marcada por uma certa valorização de trabalhos de natureza (predominantemente) quantitativa. Nesta linha de pensamento, podemos assumir que “a influência da psicometria, a valorização das escalas quantificáveis, a valorização dos dados estatísticos como evidência científica mais legítima são exemplos da integração deste enquadramento metodológico” (Duarte, 2021, p. 125) nos estudos educativos. De algum modo, o predomínio identificado no corpus analisado parece, pelo menos em parte, fazer ecos desta história. De modo a prosseguirmos com a discussão, tomemos em atenção a Tabela 6.

**Tabela 6**

*Relação entre os propósitos empíricos com o design metodológico*

4. Processos de análise 2. Principais domínios conceptuais/científicos	4.1 Análise estatística (Fa=29)	4.2 Análise mista (Fa=12)	4.3 Análise qualitativa (Fa=20)
5.1 Avaliação (Fa=20)	6	4	10
5.2 Instrumentos e técnicas de avaliação concretos (Fa=8)	6	-	2
5.3 Educação Matemática (Fa=16)	7	3	6
5.4 Educação de Língua Materna (Fa=9)	6	1	2
5.5 Educação em pensamento computacional (Fa=6)	6	-	-
5.6 Educação Científica (Fa=5)	1	-	4
5.7 Estudos curriculares (Fa=6)	1	2	3
5.8 Tecnologia educativa (Fa=4)	3	-	1
5.9 Dimensões psicossociais (Fa=4)	3	1	-
5.10 Educação de uma segunda língua (Fa=3)	-	2	1
5.11 Educação Física (Fa=2)	1	-	1
5.12 Educação Bilingue (Fa=1)	-	-	1
5.13 Educação em Estudos Sociais (Fa=1)	-	1	-
5.13 Educação Dramática (Fa=1)	-	-	1
5.14 Gamificação (Fa=1)	-	1	-
5.15 Inquiry-basedlearning (Fa=1)	1	-	-

A tabela evidencia que, com exceção dos trabalhos associados à tecnologia educativa e daqueles associados ao desenvolvimento do pensamento computacional, os domínios científicos predominantes no artigo não nos permitem predizer quais são as opções de análise. Nesse sentido, identificaram-se, por exemplo trabalhos no âmbito da educação da língua materna que optaram por análises estatísticas e outros que optaram por análises qualitativas ou mistas. O mesmo sucede para a educação Matemática e para a educação científica. Mais ainda, com a análise desenvolvida foi possível identificar projetos de investigação que visam articular distintos domínios, como a avaliação e os estudos curriculares, a educação Matemática, a Educação Científica ou a Educação em Ciências Sociais com dimensões psicossociais (Roberts & Brugar, 2021). Tal pluralidade de estratégias permite-nos constatar que há um conjunto muito significativo de domínios conceptuais que não limitam as formas de análise aos investigadores.

A propósito dos eixos científicos ou domínios conceptuais de maior incidência, o Anexo 3 apresenta um retrato mais detalhado daqueles que foram os principais alicerces teóricos considerados nos estudos. Estes domínios foram identificados através da leitura dos artigos, correspondendo ora aos campos teóricos com especial destaque no enquadramento conceptual, ora aos focos centrais estabelecidos nas questões de investigação e/ou objetivos empíricos. Como se ilustra nesse mesmo anexo, diferentes áreas investigação em educação são mobilizadas pelos documentos, como domínios relacionados com a docência, com a formação inicial de professores, com o desenvolvimento curricular ou com as didáticas específicas. Esse fator elucida a amplitude de âmbitos que podem ser associados à avaliação educativa como, igualmente, permitem evidenciar a complexidade do fenómeno, que pode ser analisado à luz de distintos referenciais.

Contudo, importa salientar que os principais domínios considerados ilustram, de forma implícita, as prioridades dos investigadores. Nesse sentido, mencionamos quatro aspetos. O primeiro remete para o facto de existir um conjunto significativo de projetos ( $Fa=20$ ) nos quais a avaliação se estabelece como um eixo relevante de posicionamento teórico, com maior incidência da avaliação formativa ( $Fa=6$ ) em relação à avaliação sumativa ( $Fa=1$ ). Como segundo aspeto, emerge a centralidade da educação Matemática nos projetos recolhidos, estando presente em mais de um quarto dos documentos estudados (26,2%), o que ilustra uma prioridade atribuída a este domínio conceptual/científico, que melhor exploramos em parágrafos seguintes. Como terceira nota, apontamos a existência de 8 trabalhos que desenvolvem um enquadramento teórico amplo sobre instrumentos ou técnicas avaliativas concretas, o que, uma vez mais, concorre para a ideia de se atribuir à avaliação uma dimensão maioritariamente técnica. Como última indicação, e apesar de existirem alguns trabalhos que desenvolvem uma discussão teórica associada às ciências da educação, em sentido lato, nenhum dos trabalhos analisados apresenta um posicionamento específico sobre o ensino elementar. Embora surjam referências à educação primária, como se constata pelos exemplos abaixo, estas decorrem de um posicionamento mais particular sobre áreas como as didáticas específicas ou as tecnologias educativas:

Com base no Currículo Nacional de Inglês desenvolvido pelo Ministério da Educação, as escolas devem prever quatro aulas de 45 minutos por semana, para estudantes do terceiro até ao sexto ano. O objetivo do ensino primário de inglês é ajudar os alunos desenvolverem competências básicas de audição, oralidade, leitura e escrita através de várias atividades letivas e cultivar o interesse dos estudantes em aprender inglês. (Butler et al., 2021, p. 434, tradução própria)

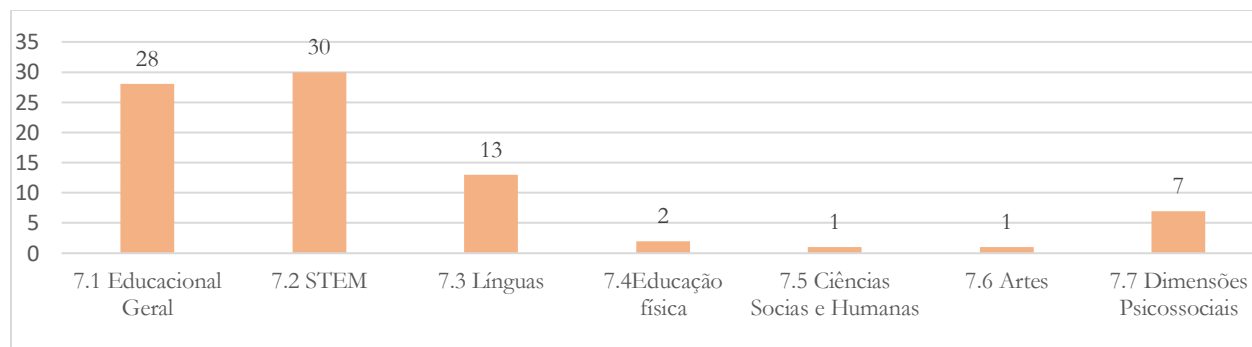
O conceito de tempo é geralmente introduzido no currículo primário de Matemática no domínio da medição (National Council of Teachers of Mathematics, 2000). De acordo com Harris (2008), o conceito de aprendizagem do tempo em Matemática

compreende duas competências-chave, nomeadamente contar o tempo e determinar as durações. (Chin et al., 2021, p. 822, tradução própria)

Em sintonia com o citado, reconhecemos que os estudos analisados integram menções particulares ao ensino primário. Contudo, esse posicionamento teórico não emerge de uma reflexão específica para este nível educativo, antes deriva da adaptação de ideias ou conceitos relacionados com a educação geral – como a avaliação – ou com o ensino particular das disciplinas, como ilustrado nas duas citações anteriores. Nesse sentido, parece ser possível considerarmos que, implicitamente, as investigações espelham tendências mais genéricas de uma certa desconsideração da ontologia particular no ensino primário (Campbell, 2005). Um outro elemento que evidencia uma certa reprodução, através das investigações, de tendências político-educativas mais amplas, prende-se com a natureza destes domínios. Agrupando os eixos conceptuais identificados, podemos obter o Gráfico 2.

## Gráfico 2

*Frequência relativa por áreas de saber*



A imagem anterior é elucidativa do modo como, também no corpus analisado, há uma prevalência das áreas STEM (Science, Technology, Engineering and Mathematics) em relação a outros domínios do conhecimento. Nesse sentido, constatamos que a investigação reproduz a valorização/desvalorização curricular. À semelhança do que se identifica a propósito das opções curriculares internacionais, há uma reduzida expressão das artes – aliás, a educação musical e as artes visuais encontram-se ausente dos estudos analisados –, das Ciências Sociais e da Educação Física. Em contrapartida, as áreas associadas às STEM são aquelas com maior representação. Por esse motivo, parece-nos ser possível afirmar que, mais do que estabelecer-se como alternativa ou com a finalidade de crítica sistemática, a investigação que articula a avaliação e o ensino elementar assume, predominantemente, uma função reprodutora daquelas que são as correntes contemporâneas com maior influência à escala global, como têm identificado Duarte (2021) e Giroux (2022).

Para concluirmos a análise de dados, evidenciaremos os principais resultados dos estudos analisados. Dialogando com a distribuição identificada no gráfico anterior, e dada a pluralidade de âmbitos de investigação, apresentamos os resultados de acordo com essas temáticas.

No domínio psicossocial, destacamos duas ideias. Em primeiro, os estudos desenvolvidos mostram que a utilização dos instrumentos de avaliação cognitiva – Cognitiv Diagnostic Assessment, no âmbito da educação Matemática, parece contribuir para uma avaliação sustentada do perfil de competências dos estudantes, nomeadamente em cálculos e resolução de problemas relacionados com o tempo (Chin et al., 2021, 2022; Chin & Chew, 2022). Também se reconhecem evidências de que o Elementary Social Behavior Assessment (ERSA) “tem um importante valor

prático, na medida em que pode identificar áreas que necessitam de melhoria no que diz respeito ao comportamento escolar prosocial” (Wikman et al., 2021, p. 6). Num outro sentido, salientamos os estudos que analisaram o desenvolvimento de competências psicossociais. Sobre este tema, a investigação clarifica que os processos de autoavaliação condicionam o modo como os estudantes desenvolvem a sua autoestima (Kalykbayeva et al., 2021), em especial quando os estudantes se envolvem de forma reflexiva e participam em todo o processo, nomeadamente na definição dos critérios (Syzdykbayeva et al., 2021). De forma complementar, a existência de questões abertas, em processos de avaliação, pode contribuir para o desenvolvimento do pensamento crítico das crianças, porque permitem aos estudantes “envolverem-se com os textos e construir conhecimento” (Roberts & Brugar, 2021, p. 146).

No âmbito da educação artística, só se estudou um único artigo, associado à educação dramática no ensino elementar (Omasta et al., 2021). Sobre este estudo, há distintas ideias que surgem como os seus principais resultados, pelo que referimos três. Em primeiro, salientamos a existência de tensões na perspetiva dos docentes, que, por um lado, criticam a existência de estratégias de avaliação heterónomas, não definidas por eles, mas, por outro, compreendem a importância da avaliação legitimar esta área curricular. De seguida, notamos que os professores do ensino elementar (na vertente da educação dramática) valorizam predominantemente uma avaliação centrada nas competências dos estudantes, em oposição a uma abordagem centrada, sobretudo, na avaliação de conteúdos aprendidos. Por fim, destacamos que os professores do ensino elementar (de Drama) parecem valorizar uma avaliação não circunscrita a uma única componente curricular, valorizando aspetos transversais, como a participação, a maturação emocional e o desenvolvimento de competências sociais, mesmo que, na sua prática, nem sempre o evidenciem nas estratégias de avaliação consideradas.

Relativamente ao trabalho de Ciências Sociais, já fizemos referência parcial, no âmbito do desenvolvimento de competências de pensamento crítico. Aprofundando um pouco mais os aspetos relacionados com as Ciências Sociais, os autores (Roberts & Brugar, 2021) esclarecem que a avaliação pode captar o modo como as crianças pensam criticamente e constroem significado sobre os diferentes assuntos lidos. O estudo induz, igualmente, que o envolvimento das crianças na leitura de textos pode contribuir para uma problematização mais sustentada das informações, enquadrando-as historicamente e pode, ainda, facilitar maior participação cívica, dado que contraria uma certa passividade das crianças e promove uma compreensão multiperspetivada da realidade social contemporânea. Como explicam os autores, este trabalho pedagógico contribuiu para que as crianças tomassem posições morais que indicavam que estavam a aplicar os seus sentidos de justiça em relação às dificuldades dos outros, em outros lugares e épocas, à medida que leem. Isto, mais uma vez, é o início de formas de análise do passado e de uma visão da História não como um conjunto de factos objetivos, mas como uma interpretação de uma série de escolhas humanas com consequências para pessoas reais. (Roberts & Brugar, 2021, p. 146, tradução própria)

Relativamente aos trabalhos no domínio da Educação Física, um dos estudos (Carrillo López & Hortigüela-Alcalá, 2022) estabelece que, para o ensino primário, há uma correlação entre a avaliação formativa e a qualidade da dieta e o peso das crianças, evidenciando que as práticas de avaliação, no ensino primário no contexto de Educação Física, não devem reduzir o seu foco às dimensões técnicas ou às competências escolares, mas assumir um olhar sobre diferentes aprendizagens, nomeadamente aquelas que podem ser mobilizadas para o quotidiano das crianças. A outra investigação (Kim & Lee, 2021) centra-se no modo como os professores perspetivam a avaliação em Educação Física, no ensino primário. Dado esse foco, o trabalho concluiu que “a avaliação em Educação Física corresponde a ações matizadas e complexas, que ultrapassam um entendimento

prático” (p.11), dado que integra contributos da experiência e da teoria. De acordo com este estudo, os professores não se limitam a replicar as suas experiências enquanto estudantes, antes apresentam preocupações formativas e pedagógicas mais abrangentes. Apesar do indicado, os autores alertam que a maioria dos professores da educação primária revelam não ter conhecimento suficiente sobre como avaliar em Educação Física, o que tende a contribuir para estratégias de avaliação assentes na medição e classificação, definidas unilateralmente pelos docentes, sem o real envolvimento de estudantes ou de outros profissionais.

Prosseguindo com a análise a propósito do ensino das línguas, há diferentes aspetos que importa destacar. Alguns dos trabalhos dedicaram-se ao estudo de ferramentas de avaliação, evidenciando contributos na utilização de distintos recursos como o *Automated Essay Scoring* (AES; Chen et al., 2022), o *Norwegian Early Writers Signal* (NEWS; Skar et al., 2022), os *Padlet* (Jong & Tan, 2021) no apoio aos docentes para a avaliação da escrita, facilitando o processo, sem substituir o envolvimento destes profissionais. Também com o propósito associado à avaliação da escrita, salientamos a pluralidade de recursos promotores de avaliação formativa como o *Automated Writing Evaluation* (AWE; Wilson et al., 2022) e o *Writing Engagement Scale* (WES; Rogers et al., 2022), este último com igual potencialidade em dinâmicas de autoavaliação. O *Writing Motivation Questionnaire* (WMQ) foi também estudado, constatando-se a sua validade para a avaliação da motivação (intrínseca e extrínseca) e dos processos de autorregulação das crianças norte-americanas que frequentam o quarto e o quinto ano de escolaridade (Graham et al., 2022). A propósito do *Reflexive Writing Instrument* (RWI), destacamos como este se pode estabelecer como um recurso de autoavaliação e de meta-reflexão, no âmbito da escrita, e, em articulação com práticas de avaliação formativa, pode contribuir para o desenvolvimento da autonomia das crianças. No âmbito da leitura, salientamos a necessidade de se considerarem diferentes instrumentos e critérios de avaliação, por exemplo, para não se reduzir a compreensão leitora à fluência oral da leitura (Amendum et al., 2021). Essa ideia é reforçada pelo trabalho de Nicholas e colaboradores (2021), que conclui a necessidade de não se estabelecerem provas estandardizadas de leitura como as únicas e principais ferramentas de avaliação, sendo necessárias, em contraciclo, práticas sistemáticas que permitam avaliar a leitura. Também a investigação de Allal (2021) evidencia que as práticas de avaliação formativa, especialmente quando os estudantes se envolvem na co-construção dos instrumentos de avaliação, contribuem para práticas mais recorrentes de revisão do texto escrito e para a melhoria de competências de escrita. Num foco mais particular, a investigação de Mitsiaki e colaboradores (2021) possibilitou o reconhecimento do *Greek Diagnostic Language Assessment* (GDLA) como uma ferramenta de diagnóstico pertinente para o ensino primário – em particular em comunidades migrantes onde o grego é uma língua adicional –, estabelecendo-se como um potencial recurso para uma compreensão mais sustentada das aprendizagens dos estudantes e, igualmente, para uma reformulação das práticas formativas, reorientando-as para uma ação mais coerente com as (novas) opções curriculares nacionais.

Os estudos no âmbito STEM são particularmente plurais, com trabalhos no domínio da educação Matemática, da educação científica, da tecnologia educativa e do pensamento computacional, pelo que analisaremos algumas das suas ideias principais. Um número relevante de trabalhos estuda a validação de instrumentos de avaliação no domínio da educação STEM. Neste âmbito, diferentes projetos consagram a pertinência de instrumentos particulares como *Base de Orientación no Lineal* (BONL; Torregrosa et al., 2021), *Competent Computational Thinking Test* (cCT; El-Hamamsy et al., 2022), *Cognitive Diagnostic Assessment* (CDA) com questões de escolha múltipla (Chin et al., 2021, 2022; Chin & Chew, 2022), *Computerized Adaptive Programming Concepts Test* (CAPCT; Hogenboom et al., 2021), *Mathematical Modeling Student Assessment* (MMSA; Turner et al., 2021), *Science Inquiry Test for Primary Education* (SIT-PE test; Pedaste et al., 2021), a utilização de um sistema gamificado de avaliação formativa (Tirado-Olivares et al., 2021) ou de um sistema de

rubricas para a avaliação Matemática (Ordaz Villegas & Acle Tomasini, 2021). Estes trabalhos legitimaram a adequação de múltiplas ferramentas que podem ser integradas em estratégias avaliativas mais amplas, a serem consideradas nos contextos educativos, a propósito da educação STEM. Existem, porém, outros resultados que importa assinalar. Iniciamos pelo reconhecimento da correlação, na aprendizagem Matemática, entre a avaliação formativa e os resultados de avaliação sumativa, em particular para estudantes que apresentam resultados mais baixos (Gezer et al., 2021), assim como salientamos a avaliação autêntica na promoção de aprendizagens Matemáticas (McFeetors et al., 2021). Também no âmbito do ensino do pensamento computacional estabeleceu-se como relevante a relação entre a prática letiva e a avaliação, como elementos interdependentes (Gane et al., 2021). No âmbito do ensino das ciências, destaca-se o modo como os docentes integram os dados provenientes da avaliação formativa na redefinição das práticas, nomeadamente na alteração de aulas ou planificações futuras (Earle, 2021). Num sentido contrário, as propostas de avaliação contidas nos manuais escolares, parecem privilegiar problemas de menor interesse pedagógico, correspondendo a exercícios predominantemente fáceis, por vezes de resolução mecânica (Tárraga-Mínguez & Tarín-Ibáñez, 2022). Complementando o indicado, um dos estudos evidencia a dificuldade de os professores da educação primária desenvolverem práticas de avaliação, no âmbito das competências e dos conhecimentos técnicos, apresentando alguns preconceitos de género em relação à proficiência no domínio da leitura e da Matemática (Wammes et al., 2022), assim como parece haver evidência de que ainda são utilizados, por alguns profissionais, “métodos de avaliação que continuam a promover a etiquetagem e o nivelamento dos alunos, de acordo com um único teste de avaliação instantâneo” (Fitzgerald et al., 2021, p. 104), evidenciando que os conceitos prévios dos docentes influenciam o modo como perspetivam a avaliação em Matemática (Sayac & Veldhuis, 2021). A este propósito, existem igualmente evidências de que os docentes, no início da sua carreira, valorizam os conhecimentos e práticas de avaliação, embora revelem alguma dificuldade em diversificar as estratégias de avaliação na prática educativa, sendo essa dificuldade mais notória nas aulas de ciências do que nas aulas de Matemática (Amador et al., 2022).

Para finalizar, discutiremos os estudos associados aos domínios gerais da educação, alguns já referidos em parágrafos anteriores. Sobre este aspeto, Julião e colaboradores (2021) são transparentes ao indicar que, para a realidade brasileira, “na prática diária (...), ainda predomina uma pedagogia do exame, que utiliza as provas como instrumentos disciplinadores e transforma o aproveitamento escolar em notas” (p. 77). Porém, outros estudos avançam com resultados com outras nuances, embora assumam uma certa conservação das práticas avaliativas, dado que as estratégias de avaliação mudam devagar (Monteiro et al., 2021). Um dos aspetos que importa destacar associa-se ao modo como processos contínuos de avaliação formativa, no ensino primário, facilitam a aprendizagem dos estudantes em diferentes domínios curriculares (Kalykbayeva et al., 2021; Van Den Heuvel-Panhuizen et al., 2021) e contribuem para melhorar os resultados da avaliação sumativa (Gezer et al., 2021), em particular quando assumem um entendimento individualizado dos processos avaliativos (Nicholas et al., 2021), facilitam o envolvimento e a reflexão dos alunos (McFeetors et al., 2021; Rogers et al., 2022) e resultam em feedback coerentes e compreendidos pelas crianças (Eriksson et al., 2022). Essa perspetiva é, aliás, partilhada pelos docentes que, ainda sem apresentarem uma conceção uniforme, assumem a pertinência de processos contínuos e sistemáticos de avaliação, de acordo com a sua finalidade formativa, sem que tal se tenha que traduzir em estratégias formais e/ou inflexíveis (Martin et al., 2022). Neste contexto, a utilização de rubricas ou outros recursos de correção e avaliação, que possibilitem um entendimento mais agregador e completo das aprendizagens dos estudantes, dos seus perfis e das suas dificuldades, parece contribuir para práticas de ensino e de avaliação com particular relevância formativa (Chin et al., 2021, 2022; Chin & Chew, 2022; Ordaz Villegas & Acle Tomasini, 2021). Também a formação pedagógica dos docentes, assim como as suas conceções sobre avaliação e ensino (Bellindo-Garcia et



al., 2022; Schmitz et al., 2021) e a sua perspectiva de justiça (Buckley-Walker & Lipscombe, 2021) condicionam o modo como estes profissionais agem, e, inevitavelmente, como desenham as diferentes estratégias de avaliação das crianças, ainda que não se estabeleçam como as únicas variáveis (Takele & Melese, 2022). Em sintonia com o mencionado, um outro aspeto relaciona-se com a existência de evidências que permitem assumir que a implicação da avaliação nos processos de desenho e desenvolvimento curricular é um fator fundamental para a melhoria das práticas pedagógicas e, conseqüentemente, para as aprendizagens dos estudantes (Afandi et al., 2021). Explicando por outras palavras, esses estudos mostram que a articulação sistemática e cíclica entre decisões curriculares, prática pedagógica e opções de avaliação são estruturantes nos diferentes processos educativos no ensino elementar. Aliás, é desse modo que os docentes entendem a sua prática, esclarecendo que fazem uso dos dados emergentes da avaliação para modificarem a sua prática e as suas planificações (Earle, 2021), não circunscrevendo a avaliação a domínios meramente cognitivos ou conceptuais (Omasta et al., 2021). Ainda assim, estão conscientes que estas decisões se encontram em diálogo com outros fatores, como a existência de provas estandardizada (Knoester & Meshulam, 2022), as decisões políticas ou a cultura organizacional de cada escola, que por vezes valorizam as dimensões mais burocráticas (Schelling & Rubenstein, 2021).

### Considerações Finais

Como indicamos ao longo do trabalho, a presente investigação tem como propósito compreender *como se tem estudado a relação entre a educação primária e a avaliação pedagógico-curricular e quais são os principais resultados dessas investigações*.

Assim – retomando o primeiro objetivo de investigação *mapear as opções metodológicas das investigações que articulam a educação primária com a avaliação pedagógico-curricular* –, foi possível constatar que há uma efetiva diversidade no modo como os estudos são realizados. Embora exista uma prevalência de trabalhos de natureza estatística, decorrente de estudos amostrais/piloto, não será adequado estabelecer um qualquer arquétipo de projeto que representa as investigações desenvolvidas nos últimos dois anos relativamente à articulação entre ensino primário e avaliação. Todavia, este trabalho permite assumir que há uma primazia de estudos dedicados a propósitos de validação ou avaliação de determinados instrumentos ou técnicas de avaliação. Esse aspeto pode, então, ser explicado por dois motivos. Por um lado, uma tendência internacional que, genericamente, faz sobressair as dimensões técnicas do funcionamento social, privilegiando um olhar instrumental sobre a ação de cada um (Morin, 2011). Por outro lado, faz ecos das concepções mais convencionais da avaliação que (sobre) valorizam as características formais e metodológicas, quase reduzindo a avaliação a estruturas burocráticas, por vezes ritualizadas, que visam medir as aprendizagens dos estudantes (Santos Guerra, 2014). Num sentido complementar, a pesquisa realizada igualmente demonstra que, à semelhança daquelas que têm sido as opções políticas educativas internacionais, os estudos têm atribuído maior destaque aos domínios STEM. A par do indicado, percebem-se reduzidas preocupações com as particularidades ontológicas do ensino primário, com diminuta, ou nula, reflexão teórica sobre aquilo que caracteriza este nível de ensino, do ponto de vista organizacional, curricular ou humano.

O identificado no parágrafo anterior – e atentando no segundo objetivo de investigação: *compreender as principais finalidades assumidas nos estudos que articulam a educação primária com a avaliação pedagógico-curricular* – permite-nos aferir que a investigação realizada assume, predominantemente, uma função reprodutora das tendências internacionais no âmbito da educação e dos aspetos sociais mais abrangentes, como discutimos nas páginas iniciais. Como tal, estes projetos de investigação parecem não conseguir estabelecer-se como efetivas alternativas ao discurso dominante, apropriando-se dos seus referências, focos e prioridades formativas. Parece, então, ser importante

continuar a analisar o modo como a academia reproduz as tendências políticas e como a investigação tem dificuldade em estabelecer uma postura crítica e/ou construir um discurso contra hegemônico. Essa análise poderá, por exemplo, problematizar como o financiamento científico é realizado, quais as margens de autonomia (política) dos centros de investigação e quais as chamadas para artigos das várias revistas.

Em articulação com o terceiro objetivo de investigação: *destacar os principais domínios científicos e alicerces conceptuais nas investigações que articulam a educação primária com a avaliação pedagógico-curricular*, constatamos que essa opção conduz à reprodução das hierarquias disciplinares, com uma expressão muito significativa de domínios científicos com maior relevância económica em detrimento de áreas como as expressões estético-artísticas e as Ciências Sociais e Humanas (Duarte, 2021). A par do mais, continua a legitimar-se a fragmentação do conhecimento (Torres Santomé, 2015), em sentido lato, e a disciplinarização do currículo, neste nível de ensino, dada a dificuldade na criação de projetos de investigação conducentes com as lógicas de integração curricular que lhe são características. Efetivamente, os estudos parecem corresponder a análises associadas a áreas particulares, como as didáticas específicas ou as tecnologias educativas, para as quais a educação primária ou a avaliação são um dos objetos de estudo, mas outras de natureza similar podiam concretizar-se para outros níveis formativos.

Dado o indicado, os estudos analisados originam uma diversidade muito significativa de resultados, marcados por um amplo domínio de saberes científicos, finalidades de investigação e alicerces teóricos. Torna-se difícil, portanto, estabelecer um corpo coeso e integrado que evidencie aquelas que são as suas principais conclusões, como se visava concretizar com o quarto objetivo de investigação: *perceber os principais resultados das investigações que possibilitem melhor compreender a relação entre educação primária e avaliação pedagógico-curricular*. Explicando por outras palavras, a amplitude de domínios científicos parece, pelo menos de forma parcial, conduzir a uma pulverização do conhecimento construído, o que nem sempre dialoga ou permite estabelecer um referencial teórico articulado e integrado. Ainda assim, podemos salientar três ideias. Em primeiro, a centralidade da avaliação formativa no ensino primário, como promotora da melhoria dos processos educativos da aprendizagem das crianças. Em segundo, a existência de inúmeros estudos sobre ferramentas/técnicas de avaliação contribui para a legitimação de estratégias diversificadas e criativas, cientificamente enquadradas, que os diferentes agentes podem mobilizar. Em terceiro, ainda que com menor representatividade face aos anteriores, a centralidade atribuída ao envolvimento e participação das crianças, através de dinâmicas de auto avaliação, de processos de co-avaliação ou de implicação na construção dos instrumentos de avaliação.

Visando, agora, apresentar uma discussão integrada dos resultados, destacamos três ideias agregadoras. Em primeiro, a forma como os estudos realizados evidenciam e reproduzem tendências políticas mais abrangentes. Aqui, destaca-se um diálogo evidente com a supremacia curricular atribuída às STEM, a centralidade técnico-burocrático da avaliação, desvalorizando ou desconsiderando as dimensões éticas que lhe estão associadas e, igualmente, a acentuação de lógicas de validação instrumental, que, pelo menos em parte, contribuem para a ideia de que é possível definir práticas educativas passíveis de serem mobilizadas de forma transversal em diferentes contextos, assumindo, pelo menos de forma implícita, processos educativos standardizados. Explicando por outras palavras, as investigações parecem carecer de uma postura efetivamente crítica, evidenciando dificuldades em estabelecer linguagens e conceitos divergentes do discurso (político) dominante, integrando-o e, de certo modo, legitimando-o. Como segunda ideia, e decorrente da anterior, salientamos como os trabalhos de investigação analisados, ao assumirem uma certa vocação instrumental de *verificação* de instrumentos ou técnicas, contribuem para perfilhar um entendimento aplicacionista da ciência, reduzindo, por um lado, o foco e as potencialidades do trabalho científico e, por outro, relegando os profissionais de educação a uma espécie de

consumidores da ciência produzida por investigadores externos às realidades escolares. Esta exteriorização da investigação, em relação aos agentes e contextos escolares, concorre para o último aspeto que importa destacar, a dificuldade da investigação (sobre avaliação) dialogar, de forma efetiva e sistemática, com as particularidades do ensino elementar, traduzindo-se numa espécie de descaracterização curricular e pedagógica deste nível de ensino.

Para finalizar, esclarecemos que esta investigação apresenta algumas limitações, nomeadamente a circunscrição a três línguas e a análise de estudos apenas publicados em 2021 e 2022. Outros trabalhos poderão contribuir para alargar a reflexão sobre este campo de investigação, integrando um número maior de estudos na constituição do *corpus*. Apesar disso, e como já identificamos, as pesquisas que relacionam a avaliação com o ensino primário privilegiam um olhar sobre este contexto educativo, e não a construção de conhecimento através das suas particularidades e práticas. Nesse sentido, parece relevante continuar a estudar este fenómeno, procurando, por exemplo, valorizar outros *designs* metodológicos que contribuam para outros tipos de análise, através de estudos de caso ou de práticas de investigação-ação, por exemplo. Esta opção pode conduzir a que a avaliação no ensino elementar seja investigada com propósitos de construção de conhecimentos praxiológicos, favorecendo um saber construído a partir da experiência de docentes e estudantes.

## Referências<sup>5</sup>

- \*Afandi, M., Wahyuningsih, S., & Mayasari, L. I. (2021). Does elementary school teacher's performance matter? *Jurnal Cakrawala Pendidikan*, 40(1), 242–252. <https://doi.org/10.21831/cp.v40i1.35284>
- Allal, L. (2021). Involving primary school students in the co-construction of formative assessment in support of writing. *Assessment in Education: Principles, Policy & Practice*, 28(5–6), 584–601. <https://doi.org/10.1080/0969594X.2021.1951164>\*
- \*Amador, J. M., Rogers, M. A. P., Hudson, R., Phillips, A., Carter, I., Galindo, E., & Akerson, V. L. (2022). Novice teachers' pedagogical content knowledge for planning and implementing mathematics and science lessons. *Teaching and Teacher Education*, 115, 1–14. <https://doi.org/10.1016/j.tate.2022.103736>
- \*Amendum, S. J., Conradi Smith, K., & Liebfreund, M. D. (2021). Explaining reading variance by student subgroup: should we move beyond oral reading fluency? *Journal of Research in Reading*, 44(4), 757–786. <https://doi.org/10.1111/1467-9817.12371>
- \*Basu, S., Rutstein, D. W., Xu, Y., Wang, H., & Shear, L. (2021). A principled approach to designing computational thinking concepts and practices assessments for upper elementary grades. *Computer Science Education*, 31(2), 169–198. <https://doi.org/10.1080/08993408.2020.1866939>
- \*Bellindo-Garcia, R., Cruzata-Martinez, A., Ponce Marin, M., & Rejas Borjas, L. (2022). Implicit Theories on Learning Assessment and the Use of Triangulation as a Means of Qualitative Validity and Reliability. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2022.4231>
- Bower, V. (2021). The purpose of education: the purpose of primary education. In V. Bower (Ed.), *Debates in primary education* (pp. 1–14). Routledge.
- \*Buckley-Walker, K., & Lipscombe, K. (2021). Validity and the design of classroom assessment in teacher teams. *The Australian Educational Researcher*, 49(2), 425–444. <https://doi.org/10.1007/s13384-021-00437-9>

<sup>5</sup> Colocou-se um asterisco em todas as referências nos trabalhos que constituíram o *corpus* de documentos analisados.

- \*Butler, Y. G., Peng, X., & Lee, J. (2021). Young learners' voices: Towards a learner-centered approach to understanding language assessment literacy. *Language Testing*, 38(3), 429–455. <https://doi.org/10.1177/0265532221992274>
- Campbell, R. (2005). Primary education or primary schooling? *Education 3-13*, 33(1), 3–6. <https://doi.org/10.1080/03004270585200021>
- \*Carrillo López, P. J., & Hortigüela-Alcalá, D. (2021). Impacto de la evaluación formativa en el estado de peso y calidad de la dieta en escolares de primaria. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity*, 8(1), 19–39. <https://doi.org/10.17979/sportis.2022.8.1.8671>
- \*Chen, D., Hebert, M., & Wilson, J. (2022). Examining human and automated ratings of elementary students' writing quality: A multivariate generalizability theory application. *American Educational Research Journal*, 000283122211067. <https://doi.org/10.3102/00028312221106773>
- \*Chin, H., & Chew, C. M. (2022). Online cognitive diagnostic assessment with ordered multiple-choice items for word problems involving 'time.' *Education and Information Technologies*, 27(6), 7721–7748. <https://doi.org/10.1007/s10639-022-10956-2>
- \*Chin, H., Chew, C. M., & Lim, H. L. (2021). Development and validation of online cognitive diagnostic assessment with ordered multiple-choice items for 'multiplication of time.' *Journal of Computers in Education*, 8(2), 289–316. <https://doi.org/10.1007/s40692-020-00180-7>
- \*Chin, H., Chew, C. M., Lim, H. L., & Thien, L. M. (2022). Development and validation of a cognitive diagnostic assessment with ordered multiple-choice items for addition of time. *International Journal of Science and Mathematics Education*, 20(4), 817–837. <https://doi.org/10.1007/s10763-021-10170-5>
- Chomsky, N., & Waterstone, M. (2021). *As consequências do capitalismo: Produção de descontentamento e resistência*. Editorial Presença.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge.
- Coltman, P., & Rolls, L. (2021). Nurturing compassionate citizens of the future: Weaving together pedagogy and curriculum. In J. Biddulph & J. Flutter (Eds.), *Inspiring primary curriculum design* (pp. 27–43). Routledge.
- Denscombe, M. (2014). *The Good Research Guide: for small-scale social research projects* (5th ed.). Open University Press.
- Duarte, P. (2021). *Pensar o desenvolvimento curricular: Uma reflexão centrada no ensino*. Instituto Politécnico do Porto. Escola Superior de Educação.
- \*Earle, S. (2021). Formative decision-making in response to primary science classroom assessment: What to do next? *Frontiers in Education*, 5, 1–7. <https://doi.org/10.3389/educ.2020.584200>
- \*El-Hamamsy, L., Zapata-Cáceres, M., Barroso, E. M., Mondada, F., Zufferey, J. D., & Bruno, B. (2022). The competent Computational Thinking Test: Development and validation of an unplugged computational thinking test for upper primary school. *Journal of Educational Computing Research*, 1818-1866. <https://doi.org/10.1177/07356331221081753>
- \*Eriksson, E., Boistrup, L. B., & Thornberg, R. (2022). “You must learn something during a lesson”: How primary students construct meaning from teacher feedback. *Educational Studies*, 48(3), 323–340. <https://doi.org/10.1080/03055698.2020.1753177>
- European Commission. (2020). *Prospective report on the future of assessment in primary and secondary education*. Directorate-General for Education, Youth, Sport and Culture.
- Eurydice. (2021). *The structure of the European education systems 2021/22: Schematic diagrams*. Author.
- Finfgeld-Connett, D. (2014). Use of content analysis to conduct knowledge-building and theory-generating qualitative systematic reviews. *Qualitative Research*, 14(3), 341–352. <https://doi.org/10.1177/1468794113481790>

- \*Fitzgerald, L., Hunter, J., & Hunter, R. (2021). Shifting teacher practices in relation to grouping: Gap gazing or strengths focused approaches. *Mathematics Teacher Education and Development*, 23(3), 97–110.
- \*Gamboa-Araya, R., Hidalgo-Mora, R., & Castillo-Sánchez, M. (2022). La implementación de los programas de estudio de matemática en primaria desde la visión de la persona docente. *Uniciencia*, 36(1), 1–31. <https://doi.org/10.15359/ru.36-1.11>
- \*Gane, B. D., Israel, M., Elagha, N., Yan, W., Luo, F., & Pellegrino, J. W. (2021). Design and validation of learning trajectory-based assessments for computational thinking in upper elementary grades. *Computer Science Education*, 31(2), 141–168. <https://doi.org/10.1080/08993408.2021.1874221>
- \*Gezer, T., Wang, C., Polly, A., Martin, C., Pugalee, D., & Lambert, R. (2021). The relationship between formative assessment and summative assessment in primary grade mathematics classrooms. *International Electronic Journal of Elementary Education*, 13(5), 673–685. <https://doi.org/10.26822/iejee.2021.220>
- Giroux, H. A. (2022). *Pedagogy of resistance: Against manufactured ignorance*. Bloomsbury Academic.
- \*Glušac, T., & Milić, M. (2021). Quality of written instructions in teacher-made tests of English as a foreign language. *English Teaching & Learning*, 46(1), 39–57. <https://doi.org/10.1007/s42321-021-00079-1>
- \*Graham, S., Harbaugh-Schattenkirk, A. G., Aitken, A., Harris, K. R., Ng, C., Ray, A., Wilson, J. M., & Wdowin, J. (2022). Writing motivation questionnaire: validation and application as a formative assessment. *Assessment in Education: Principles, Policy & Practice*, 29(2), 238–261. <https://doi.org/10.1080/0969594X.2022.2080178>
- Hayes, D. (2010). *Encyclopedia of primary education*. Routledge.
- \*Herbert, S. (2021). Overcoming challenges in assessing mathematical reasoning. *Australian Journal of Teacher Education*, 46(8), 17–30. <https://doi.org/10.14221/ajte.2021v46n8.2>
- \*Herwin, H., Nurhayati, R., & Dahalan, S. C. (2022). Mobile assessment to improve learning motivation of elementary school students in online learning. *International Journal of Information and Education Technology*, 12(5), 436–442. <https://doi.org/10.18178/ijiet.2022.12.5.1638>
- \*Hogenboom, S. A. M., Hermans, F. F. J., & van der Maas, H. L. J. (2021). Computerized adaptive assessment of understanding of programming concepts in primary school children. *Computer Science Education*, 1–30. <https://doi.org/10.1080/08993408.2021.1914461>
- \*Ibourk, A., & Kendrick, M. (2021). Elementary students' explanation of variation of traits and teacher's feedback using an online embedded assessment tool. *International Journal of Science Education*, 43(8), 1173–1192. <https://doi.org/10.1080/09500693.2021.1903618>
- \*Jong, B., & Tan, K. H. (2021). Using Padlet as a technological tool for assessment of students writing skills in online classroom settings. *International Journal of Education and Practice*, 9(2), 411–423. <https://doi.org/10.18488/journal.61.2021.92.411.423>
- \*Julião, B. de O., Piza, C. T., & Mello, C. B. de. (2021). Escala de sondagem do desenvolvimento neuropsicomotor e de habilidades acadêmicas para o Ensino Fundamental I: Percepção de professores. *Ensaio: Avaliação e Políticas Públicas Em Educação*, 29(112), 761–782. <https://doi.org/10.1590/s0104-40362020002802036>
- \*Kalykbayeva, A., Satova, A., Autayeva, A., Ospanova, A., Suranchina, A., & Elmira, U. (2021). Using self-assessment instruction to develop primary school students' self-esteem in inclusive practice in Kazakhstan. *Cypriot Journal of Educational Sciences*, 16(4), 1631–1642. <https://doi.org/10.18844/cjes.v16i4.6049>
- \*Kim, Y., & Lee, O. (2021). Exploring the phenomenological meanings of assessment in physical education through primary teachers' lived experiences. *Physical Education and Sport Pedagogy*, 1–13. <https://doi.org/10.1080/17408989.2021.1911983>

- \*Knoester, M., & Meshulam, A. (2022). Beyond deficit assessment in bilingual primary schools. *International Journal of Bilingual Education and Bilingualism*, 25(3), 1151–1164. <https://doi.org/10.1080/13670050.2020.1742652>
- Lissovoy, N. D. (2022). *Capitalism, pedagogy, and the politics of being*. Bloomsbury Academic.
- \*Martin, C. L., Mraz, M., & Polly, D. (2022). Examining elementary school teachers perceptions of and use of formative assessment in mathematics. *International Electronic Journal of Elementary Education*. 417-425. <https://doi.org/10.26822/iejee.2022.253>
- \*McFadden, J., Jung, K., Robinson, B., & Tretter, T. R. (2022). Teacher-developed multi-dimensional science assessments supporting elementary teacher learning about the Next Generation Science Standards. *Journal of Science Teacher Education*, 33(1), 55–82. <https://doi.org/10.1080/1046560X.2021.1905331>
- \*McFeeters, P. J., Marynowski, R., & Candler, A. (2021). Generative unit assessment: Authenticity in mathematics classroom assessment practices. *Education Sciences*, 11(7), 1-22. <https://doi.org/10.3390/educsci11070366>
- \*Mitsiaki, M., Kyriakou, N., Kyprianou, D., Giannaka, C., & Hadjitheodoulou, P. (2021). Washback effects of diagnostic assessment in Greek as an SL: Primary school teachers' perceptions in Cyprus. *Languages*, 6(4), 1-28. <https://doi.org/10.3390/languages6040195>
- \*Mohamed, M., Abd Aziz, M. S., & Ismail, K. (2021). “Assessment for learning” practices amongst the primary school English language teachers: A mixed methods approach. *Pertanika Journal of Social Sciences and Humanities*, 29(3), 1875 – 1900. <https://doi.org/10.47836/pjssh.29.3.21>
- \*Monteiro, V., Mata, L., & Santos, N. N. (2021). Assessment conceptions and practices: Perspectives of primary school teachers and students. *Frontiers in Education*, 6, 1-15. <https://doi.org/10.3389/feduc.2021.631185>
- Morin, E. (2011). *A Via: para o futuro da humanidade*. Edições Piaget.
- \*Nicholas, M., Rouse, E., & Paatsch, L. (2021). Child-centred teaching: Helping each child to reach their full potential. *Education Sciences*, 11(6), 1-19. <https://doi.org/10.3390/educsci11060280>
- OCDE. (2021). *Education at a glance 2021*. <https://doi.org/10.1787/b35a14e5-en>
- \*Omasta, M., Murray, B., McAvoy, M., & Chappell, D. (2021). Assessment in elementary-level drama education: Teachers' conceptualizations and practices. *Arts Education Policy Review*, 122(4), 265–279. <https://doi.org/10.1080/10632913.2020.1746710>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishers.
- \*Ordaz Villegas, G., & Acle Tomasini, G. (2021). Desempeñomatemático. *Perfiles Educativos*, 43(173), 76-93. <https://doi.org/10.22201/iisue.24486167e.2021.173.59772>
- \*Pedaste, M., Baucal, A., & Reisenbuk, E. (2021). Towards a science inquiry test in primary education: development of items and scales. *International Journal of STEM Education*, 8(1), 1-19. <https://doi.org/10.1186/s40594-021-00278-z>
- \*Roberts, K. L., & Brugar, K. A. (2021). Using verbal protocol to examine construction of meaning from social studies texts. *The Journal of Social Studies Research*, 46(2), 135–151. <https://doi.org/10.1016/j.jssr.2021.05.005>
- \*Rogers, P. M., Marine, J. M., Ives, S. T., Parsons, S. A., Horton, A., & Young, C. (2022). Validity evidence for a formative writing engagement assessment in elementary grades. *Assessment in Education: Principles, Policy & Practice*, 29(2), 262–284. <https://doi.org/10.1080/0969594X.2022.2054942>
- Rosales López, C. (2014). *Evaluar es reflexionar sobre la enseñanza* (3ª ed.). Narcea Ediciones.
- \*Ryan, M., Khosronejad, M., Barton, G., Myhill, D., & Kervin, L. (2022). Reflexive writing dialogues: Elementary students' perceptions and performances as writers during classroom experiences. *Assessing Writing*, 51, 1-20. <https://doi.org/10.1016/j.asw.2021.100592>

- Santos Guerra, M. Á. (2014). *La evaluación como aprendizaje: Cuando la flecha impacta en la Diana* (3ª ed.). Narcea Ediciones.
- \*Sayac, N., & Veldhuis, M. (2021). Mathematics assessment practices of primary school teachers in France. *International Journal of Science and Mathematics Education*, 1-16. <https://doi.org/10.1007/s10763-021-10229-3>
- \*Schelling, N., & Rubenstein, L. D. (2021). Elementary teachers' perceptions of data-driven decision-making. *Educational Assessment, Evaluation and Accountability*, 33(2), 317-344. <https://doi.org/10.1007/s11092-021-09356-w>
- \*Schmitz, H., Almeida, T. G. S., & Souza, B. N. S. (2021). What are the assessment instruments chosen by elementary school teachers and what are the reasons to choose them? *Revista Tempos e Espaços Em Educação*, 14(33), 1-14. <https://doi.org/10.20952/revtee.v14i33.14289>
- \*Skar, G. B., Kvistad, A. H., Johansen, M. B., Rijlaarsdam, G., & Aasen, A. J. (2022). Identifying texts in the warning zone. *Writing and Pedagogy*, 13(1-3), 267-310. <https://doi.org/10.1558/wap.22095>
- Suri, H. (2018). Meta-analysis, systematic reviews and research syntheses. In L. Cohen, M. Lawrence, & Morrison. Keith (Eds.), *Research methods in education* (8th ed., pp. 427-439). Routledge.
- \*Syzydkbayeva, A. D., Izmagambetova, R. K., Amirova, A. S., Bainazarova, T. B., & Sadykova, M. K. (2021). Formation of primary school children self esteem on the basis of criteria-based assessment. *The Education and Science Journal*, 23(7), 147-169. <https://doi.org/10.17853/1994-5639-2021-7-147-169>
- \*Takele, M., & Melese, W. (2022). Primary school teachers' conceptions and practices of assessment and their relationships. *Cogent Education*, 9(1), 1-16. <https://doi.org/10.1080/2331186X.2022.2090185>
- \*Tárraga-Mínguez, R., & Tarín-Ibáñez, J. (2022). Mathematical problem solving in elementary education. An analysis of facilitator guides. *Revista de Educación*, 396, 235-259. <https://doi.org/10.4438/1988-592X-RE-2022-396-536>
- \*Tirado-Olivares, S., González-Calero, J. A., Cózar-Gutiérrez, R., & Toledano, R. M. (2021). Gamificando la evaluación: Una alternativa a la evaluación tradicional en educación primaria. *REICE. Revista Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, 19(4), 125-143. <https://doi.org/10.15366/reice2021.19.4.008>
- \*Torregrosa, A., Albarracín, L., & Deulofeu, J. (2021). Orientación y coevaluación: Dos aspectos clave para la evolución del proceso de resolución de problemas. *Bolema: Boletim de Educação Matemática*, 35(69), 89-111. <https://doi.org/10.1590/1980-4415v35n69a05>
- Torres Santomé, J. (2015). Sin muros en las aulas: El currículum integrado. In *Ensayos sobre el currículum: teoría y práctica* (pp. 148-159). Ediciones Morata.
- \*Tsai, M.-J., Chien, F. P., Wen-Yu Lee, S., Hsu, C.-Y., & Liang, J.-C. (2022). Development and validation of the Computational Thinking Test for Elementary School Students (CIT-ES): Correlate CT competency with CT disposition. *Journal of Educational Computing Research*, 60(5), 1110-1129. <https://doi.org/10.1177/07356331211051043>
- \*Turner, E. E., Roth McDuffie, A., Bennett, A. B., Aguirre, J., Chen, M.-K., Foote, M. Q., & Smith, J. E. (2021). Mathematical modeling in the elementary grades: Developing and testing an assessment. *International Journal of Science and Mathematics Education*. 1-23. <https://doi.org/10.1007/s10763-021-10195-w>
- \*Unahalekhaka, A., & Bers, M. U. (2022). Evaluating young children's creative coding: Rubric development and testing for ScratchJr projects. *Education and Information Technologies*, 27(5), 6577-6597. <https://doi.org/10.1007/s10639-021-10873-w>

- \*van den Heuvel-Panhuizen, M., Sangari, A. A., & Veldhuis, M. (2021). Teachers' use of descriptive assessment in primary school mathematics education in Iran. *Education Sciences*, 11(3), 100. <https://doi.org/10.3390/educsci11030100>
- \*Wammes, D., Slof, B., Schot, W., & Kester, L. (2021). Pupils' prior knowledge about technological systems: design and validation of a diagnostic tool for primary school teachers. *International Journal of Technology and Design Education*. 1-33. <https://doi.org/10.1007/s10798-021-09697-z>
- \*Wammes, D., Slof, B., Schot, W., & Kester, L. (2022). Teacher judgement accuracy of technical abilities in primary education. *International Journal of Technology and Design Education*. 1-24. <https://doi.org/10.1007/s10798-022-09734-5>
- \*Wikman, C., Allodi, M. W., & Ferrer-Wreder, L. (2021). Psychometric properties of the elementary social behavior assessment in Swedish primary school: A teacher rated index of students' prosocial school behaviors. *Frontiers in Education*, 6, 1-6. <https://doi.org/10.3389/educ.2021.681873>
- \*Wilson, J., Myers, M. C., & Potter, A. (2022). Investigating the promise of automated writing evaluation for supporting formative writing assessment at scale. *Assessment in Education: Principles, Policy & Practice*, 29(2), 183–199. <https://doi.org/10.1080/0969594X.2022.2025762>
- Wyse, D. (2002). *Becoming a primary school teacher*. Routledge Falmer.



## Anexo 1

## Países Associados aos Estudos Analisados

País associado ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
Estados Unidos da América	17	27,9%
Malásia	5	8,2%
Austrália	4	6,6%
Espanha	3	4,9%
Países Baixos	3	4,9%
Brasil	2	3,3%
Indonésia	2	3,3%
República do Cazaquistão	2	3,3%
Suíça	2	3,3%
Suécia	2	3,3%
Canadá	1	1,6%
China	1	1,6%
Chipre	1	1,6%
Coreia	1	1,6%
Costa Rica	1	1,6%
Espanha	1	1,6%
Estónia	1	1,6%
Etiópia	1	1,6%
França	1	1,6%
Hong Kong	1	1,6%
Inglaterra	1	1,6%
Irão	1	1,6%
México	1	1,6%
Noruega	1	1,6%
Nova Zelândia	1	1,6%
Peru	1	1,6%
Portugal	1	1,6%
Sérvia	1	1,6%
Taiwan	1	1,6%

## Anexo 2

## Revistas Associadas aos Estudos Analisados

Revistas associadas ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
<i>Assessment in Education: Principles, Policy &amp; Practice</i>	4	6,6%
<i>Computer Science Education</i>	3	4,9%
<i>Education Sciences</i>	3	4,9%
<i>Frontiers in Education</i>	3	4,9%
<i>International Journal of Science and Mathematics Education</i>	3	4,9%
<i>Australian Journal of Teacher Education</i>	2	3,3%
<i>Education and Information Technologies</i>	2	3,3%
<i>International Electronic Journal of Elementary Education</i>	2	3,3%
<i>International Journal of Technology and Design Education</i>	2	3,3%
<i>Journal of Educational Computing Research</i>	2	3,3%
<i>American Educational Research Journal</i>	1	1,6%
<i>Arts Education Policy Review</i>	1	1,6%
<i>Assessing Writing</i>	1	1,6%
<i>Bolema</i>	1	1,6%
<i>Cakrawala Pendidika</i>	1	1,6%
<i>Cogent Education</i>	1	1,6%
<i>Cypriot Journal of Educational Sciences</i>	1	1,6%
<i>Educational Assessment, Evaluation and Accountability</i>	1	1,6%
<i>Educational Studies</i>	1	1,6%
<i>English Teaching &amp; Learning</i>	1	1,6%
<i>Ensaio</i>	1	1,6%
<i>International Journal of Science Education</i>	1	1,6%
<i>International Journal of Bilingual Education and Bilingualism</i>	1	1,6%
<i>International Journal of Education and Practice</i>	1	1,6%
<i>International Journal of Information and Education Technology</i>	1	1,6%

Revistas associadas ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
<i>International Journal of STEM Education</i>	1	1,6%
<i>Journal Computational Education</i>	1	1,6%
<i>Journal of Research in Reading</i>	1	1,6%
<i>Journal of Science Teacher Education</i>	1	1,6%
<i>Language Testing</i>	1	1,6%
<i>Languages</i>	1	1,6%
<i>Mathematics Teacher Education and Development</i>	1	1,6%
<i>Perfiles Educativos</i>	1	1,6%
<i>Physical Education and Sport Pedagogy</i>	1	1,6%
REICE	1	1,6%
<i>Revista de Educación</i>	1	1,6%
<i>Revista tempos e espaços em educação</i>	1	1,6%
<i>Social Sciences &amp; Humanities</i>	1	1,6%
<i>Sportis</i>	1	1,6%
<i>Teaching and Teacher Education</i>	1	1,6%
<i>The Education and Science Journa</i>	1	1,6%
<i>The Journal of Social Studies Research</i>	1	1,6%
<i>The Qualitative Report</i>	1	1,6%
<i>Uniciencia</i>	1	1,6%
<i>Writing &amp; Pedagogy</i>	1	1,6%

## Anexo 3

## Totalidade de Domínios Conceptuais Abordados nos Estudos Analisados

Revistas associadas ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
1.1 Avaliação	23	37,7%
1.2 Avaliação sumativa	3	4,9%
1.3 Avaliação autêntica	3	4,9%
1.4 Auto-avaliação	3	4,9%
1.5 Avaliação diagnóstica	1	1,6%
1.6 Co-avaliação	1	1,6%
1.7 Conceptualização da avaliação	13	21,3%
1.7.1 Perspetivas clássicas	6	9,8%
1.7.2 Perspetiva dos docentes	4	6,6%
1.7.3 Perspetiva dos estudantes	4	6,6%
1.8 Avaliação como um processo de melhoria dos processos educativos	6	9,8%
1.9 Instrumentos e técnicas de avaliação	21	34,4%
1.9.1 Validação de instrumentos/técnicas de avaliação	8	13,1%
1.9.2 Utilização de rúbricas de avaliação	6	9,8%
1.9.3 Instrumentos de escolha múltipla	4	6,6%
1.9.4 Linguagem considerada nos instrumentos de avaliação	1	1,6%
1.10 Opções avaliativas nacionais	3	4,9%
1.11 Contributos da avaliação para a compartimentalização curricular	1	1,6%
1.12 Dimensão cognitiva da avaliação	1	1,6%
<b>2. Educação matemática</b>	<b>18</b>	<b>29,5%</b>
2.1 Domínios de aprendizagem	11	18,0%
2.1.1 Raciocínio matemático	5	8,2%
2.1.2 Resolução de problemas	5	8,2%
2.1.3 Tempo	2	3,3%
2.1.4 Metacognição	1	1,6%
2.1.5 Modelação matemática	1	1,6%
2.1.6 Sentido de número	1	1,6%
2.2 Práticas de avaliação (matemática)	7	11,5%
2.3 Instrumentos e técnicas de avaliação	4	6,6%
2.4 Orientações curriculares nacionais (de Matemática)	3	4,9%
2.5 Formação e conhecimento profissional docente	2	3,3%
2.6 Influência do currículo apresentado no ensino de Matemática	1	1,6%
2.7 Auto-regulação na aprendizagem matemática	1	1,6%
2.7.1 Base de Orientación no Lineal (BONL)	1	1,6%
<b>3. Desenvolvimento curricular</b>	<b>16</b>	<b>26,2%</b>
3.1 Avaliação como uma dimensão do desenvolvimento curricular	6	9,8%

Revistas associadas ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
3.2 Planificação curricular	5	8,2%
3.3 Práticas de 'teach to the test'	4	6,6%
3.4 Finalidades educativas/ curriculares	3	4,9%
3.5 Opções curriculares nacionais	3	4,9%
3.6 Data-driven decision-making (DDDM)	1	1,6%
<b>4. Instrumentos de avaliação a larga escala</b>	<b>12</b>	<b>19,7%</b>
4.1 Influências negativas dos testes estandarizados	1	1,6%
4.2 Instrumentos de avaliação associados à leitura e à escrita	3	4,9%
4.3 Programme for International Student Assessment (PISA)	2	3,3%
4.4 Trends in International Mathematics and Science Study (TIMSS)	2	3,3%
4.5 Instrumentos de avaliação associados à Matemática	1	1,6%
4.6 Greek Diagnostic Language Assessment (GDLA)	1	1,6%
4.7 National Assessment of Educational Progress (NAEP)	1	1,6%
4.8 National Assessment Program -Literacy and Numeracy (NAPLAN)	1	1,6%
4.9 Virtual Performance Assessment	1	1,6%
<b>5. Educação de língua materna</b>	<b>10</b>	<b>16,4%</b>
5.1 Avaliação da escrita	6	9,8%
5.2 Avaliação da leitura	1	1,6%
5.3 Relação da literacia com o (in)sucesso escolar	4	6,6%
5.4 Práticas de avaliação (da língua materna)	3	4,9%
5.5 Opções curriculares nacionais (no âmbito do ensino da língua materna)	2	3,3%
5.6 Instrumentos e técnicas de avaliação	2	3,3%
5.7 Motivação na aprendizagem da língua	1	1,6%
5.8 Metacognição na aprendizagem da língua	1	1,6%
6. Educação de pensamento computacional e similares	6	9,8%
<b>6.1 Instrumentos e técnicas de avaliação</b>	<b>5</b>	<b>8,2%</b>
6.2 Relação da resolução de problemas com a aprendizagem de pensamento computacional	5	8,2%
6.3 Práticas de avaliação na educação do pensamento computacional	3	4,9%
6.3 Relevância curricular do pensamento computacional	2	3,3%
6.3 Integração curricular do pensamento computacional, na educação primária	1	1,6%
6.3 Educação do pensamento computacional como preparação laboral	1	1,6%
<b>7. Tecnologias educativas</b>	<b>6</b>	<b>9,8%</b>
7.1 Práticas de avaliação	5	8,2%
7.2 Ferramentas de avaliação	2	3,3%
7.2.1 Padlet	1	1,6%
7.2.2 Telemóveis	1	1,6%
7.3 Relação das tecnologias educativas com o currículo	1	1,6%

Revistas associadas ao estudo	Frequência Absoluta (Fa)	Frequência Relativa (Fr)
7.4 Relação das tecnologias educativas com o ensino de competências técnicas	1	1,6%
7.5 Relação das tecnologias educativas com o ensino da língua	1	1,6%
7.6 Relação das tecnologias educativas com a motivação	1	1,6%
<b>8. Docência</b>	<b>5</b>	<b>8,2%</b>
8.1 Formação contínua e desenvolvimento profissional docente	3	4,9%
8.2 Formação inicial docente	1	1,6%
8.2 Relação das concepções pedagógicas com a prática docente	2	3,3%
8.3 Relação do conhecimento profissional com a prática docente	1	1,6%
<b>9. Educação científica</b>	<b>4</b>	<b>6,6%</b>
9.1 Opções curriculares nacionais (no âmbito da educação científica)	2	3,3%
9.2 Práticas de avaliação	2	3,3%
9.3 Domínios de aprendizagem	1	1,6%
9.3.1 Pensamento científico	1	1,6%
9.4 Conhecimento profissional docente	1	1,6%
9.5 Instrumentos e técnicas de avaliação	1	1,6%
<b>10. Educação de língua não materna</b>	<b>4</b>	<b>6,6%</b>
10.1 Técnicas e instrumentos de avaliação	3	4,9%
10.2 Práticas de avaliação	1	1,6%
10.3 Relação com a aprendizagem da escrita	1	1,6%
10.3 Relação com estudantes migrantes	1	1,6%
<b>11. Educação Física</b>	<b>2</b>	<b>3,3%</b>
11.1 Opções curriculares nacionais (no âmbito da educação física)	2	3,3%
11.2 Práticas de avaliação	1	1,6%
11.3 Relação da Educação Física com COVID-19	1	1,6%
11.4 Relação da Educação Física com a obesidade	1	1,6%
<b>12. Gamificação</b>	<b>2</b>	<b>3,3%</b>
12.1 Centralidade do sistema de pontos	2	3,3%
12.2 Valorização da competição	1	1,6%
<b>13. Educação bilingue</b>	<b>1</b>	<b>1,6%</b>
13.1 Instrumentos e técnicas de avaliação	1	1,6%
<b>14. Educação dramática</b>	<b>1</b>	<b>1,6%</b>
14.1 Educação dramática no ensino generalista	1	1,6%
14.2 Instrumentos e técnicas de avaliação	1	1,6%
<b>15. Inquiry-based learning</b>	<b>1</b>	<b>1,6%</b>
15.1 Instrumentos e técnicas de avaliação	1	1,6%
15.2 Fases do Inquiry-based learning	1	1,6%
15.2 Relação do Inquiry-based learning com a educação científica	1	1,6%

*Nota:* O processo de codificação previa que o mesmo artigo ou código pudesse ficar associado a mais do que um domínio conceptual.

## Sobre o Autores

### Pedro Duarte

Escola Superior de Educação do Politécnico do Porto

pedropereira@ese.ipp.pt

ORCID: <https://orcid.org/0000-0002-3048-6959>

Pedro Duarte é doutorado em Educação pela Universidade de Santiago de Compostela e docente na Escola Superior de Educação do Instituto Politécnico do Porto. Na sua investigação, o autor tem-se focado em diferentes domínios educativos, nomeadamente: o currículo; as políticas educativas e a administração escolar; a educação cidadã. Encontra-se, atualmente, a desenvolver o seu pós-doutoramento em Políticas Educativas e Administração Escolar, na Universidade de Aveiro. É investigador integrado do inED e investigador colaborador do CITCEM.

### Dora Fonseca

Universidade de Aveiro

dorafonseca@ua.pt

ORCID: <https://orcid.org/0000-0003-4408-9716>

Dora Fonseca é doutorada em Ciências de Educação e pós-doutorada em Administração Educacional pela Universidade de Aveiro, instituição onde, atualmente, é docente. É investigadora integrada do Centro de Investigação em Didática e Tecnologia na Formação de Formadores (CIDTFF). Em termos de investigação, os seus interesses situam-se na área das Políticas Educativas e Administração Educacional, com foco nos processos de descentralização, avaliação organizacional e regulação educativa e, ainda, processos de internacionalização no ensino superior e rankings universitários.

---

# arquivos analíticos de políticas educativas

Volume 31 Número 66

6 de junho 2023

ISSN 1068-2341

---



Este artigo pode ser copiado, exibido, distribuído e adaptado, desde que o(s) autor(es) e *Arquivos Analíticos de Políticas Educativas* sejam creditados e a autoria original atribuídos, as alterações sejam identificadas e a mesma licença CC se aplique à obra derivada. Mais detalhes sobre a licença Creative Commons podem ser encontrados em <https://creativecommons.org/licenses/by-sa/4.0/>. *Arquivos Analíticos de Políticas Educativas* é publicado pela Mary Lou Fulton Teachers College, Arizona State University. Os artigos que aparecem na AAPE são indexados em CIRC (Clasificación Integrada de Revistas Científicas, España) DIALNET (Espanña), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), PubMed, QUALIS A1 (Brazil), Redalyc, SCImago Journal Rank, SCOPUS, Socolar (China).

Sobre o Conselho Editorial: <https://epaa.asu.edu/ojs/about/editorialTeam>

Para erros e sugestões, entre em contato com [Fischman@asu.edu](mailto:Fischman@asu.edu)

EPAA Facebook (<https://www.facebook.com/EPAAAPE>) Twitter feed @epaa\_aape.