



Examining how generative AI tools benefit and challenge teachers' research-informed practice

Stephen Sowa, Chris Brown, Tae-Hee Choi & Rachele Newman

To cite this article: Stephen Sowa, Chris Brown, Tae-Hee Choi & Rachele Newman (09 Jun 2025): Examining how generative AI tools benefit and challenge teachers' research-informed practice, European Journal of Teacher Education, DOI: [10.1080/02619768.2025.2518193](https://doi.org/10.1080/02619768.2025.2518193)

To link to this article: <https://doi.org/10.1080/02619768.2025.2518193>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 09 Jun 2025.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

Examining how generative AI tools benefit and challenge teachers' research-informed practice

Stephen Sowa , Chris Brown , Tae-Hee Choi  and Rachele Newman

Southampton Education School, University of Southampton, Southampton, UK

ABSTRACT

Teachers' utilisation of research evidence, known as research-informed educational practice (RIEP), has yet to be widely adopted in schools due to barriers such as time constraints and the complex language used in scientific publications. However, the introduction of generative artificial intelligence (AI) tools like ChatGPT offers potential solutions to these challenges. This study explores the benefits and challenges associated with teachers' use of generative AI tools for their RIEP. A systematic review was conducted, guided by a theory of RIEP. From 790 initial studies, 19 were included after full-text screening. Our findings reveal a range of practical, pedagogical, and psychological benefits and costs associated with teachers' use of generative AI tools in their RIEP. Additionally, our review indicates that while these tools can enhance teachers' engagement in RIEP, they may also have mixed impacts on their professional identity – both challenging and enriching the theoretical framework employed in this study.

ARTICLE HISTORY

Received 30 August 2024

Accepted 31 May 2025

KEYWORDS

Teaching; research-informed practice; generative AI

Introduction

This study examines teachers' engagement in research-informed educational practices and how generative artificial intelligence (AI) tools are being used to support such practices. For the purposes of this paper, we define research-informed educational practice (RIEP) as the utilisation of research evidence by teachers and school leaders to enhance aspects of their teaching, decision-making, leadership, or ongoing professional development (Brown 2020). Encouraging RIEP can be important for multiple reasons. A growing body of literature suggests that when educators engage with research evidence to make or alter decisions, embark on new initiatives, or develop novel practices, this can positively impact both teaching and learning outcomes (Cain 2016; Cordingley 2015; Godfrey 2016). Moreover, because education involves a certain kind of intervening in the lives of students, there may be a moral responsibility for teachers to draw on high-quality research evidence to secure benefits and minimise harms.

CONTACT Stephen Sowa  s.j.sowa@soton.ac.uk  Southampton Education School, University of Southampton, Highfield Campus, University Road, Southampton SO17 1BJ, UK

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/02619768.2025.2518193>

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Despite these considerations, RIEP as a ‘business-as-usual’ approach has not yet been widely adopted in most schools (Biesta 2007; Cain 2016). For instance, studies across various country contexts indicate that, typically, the majority of teachers (often more than 60%) rarely or never use research evidence in their practice (Cain 2016). A lack of adoption or disconnects between research and practice can occur due to educators encountering difficulties enacting various aspects of RIEP. Facilitating educational improvements through research evidence typically involves educators, either collectively or individually, engaging in five key steps: (1) accessing academic research; (2) comprehending academic research; (3) critically engaging with research evidence by understanding its strengths and weaknesses and justifying its truth claims; (4) relating research evidence to existing knowledge and understanding; (5) making or altering decisions, embarking on new initiatives, or developing new practices when required (Brown et al. 2022). With regard to these five steps, a frequently cited issue is the difficulty educators face in accessing academic research, which is often behind paywalls (Cordingley 2015). Educators can additionally struggle with understanding the esoteric language commonly used in academic research (Cain et al. 2019).

Another common issue is that research is either too context-independent or highly specific, making it difficult for teachers to apply findings effectively to their own settings (Biesta 2007; Cain et al. 2019; Wrigley 2018). Teachers and school leaders often report a lack time to engage with research, learn from it, or implement it to develop new practices (Brown 2020; Cain 2016). In summary, while educators’ engagement with research could significantly enhance teaching and learning outcomes, various conceptual and practical barriers have hitherto prevented RIEP from being meaningfully realised. However, with the dawn of new generative AI technologies, there could be scope to address many of these barriers.

Generative AI and teachers’ RIEP

The release of OpenAI’s ‘ChatGPT’ in 2022 stimulated a tidal wave of engagement with generative AI tools, with around 42–72% of teachers reporting their use of these tools Department for Education (2024); Laak and Aru (2024). This technology represents a possible paradigm shift for RIEP as we know it. No longer are teachers reliant on academic researchers/research. Instead, by entering prompts or questions into generative AI platforms such as ChatGPT or Gemini, teachers can instantly summarise and incorporate the latest research evidence into anything, from lesson-plans to school improvement policies – with these tools providing accessible and actionable results (Celik et al. 2022). Further, AI chatbots can provide interactive means for educators to ask questions about specific research findings, methodologies, or concepts, receiving explanations in plain language on the strengths and weaknesses of studies. Platforms using AI to aggregate open-access research papers (e.g. CORE) can also bypass paywalls by directing educators to freely available research.

To help facilitate the practical application of research findings, generative AI tools can also be used to generate accompanying implementation plans, including timelines and resource requirements. With this level of utility, generative AI tools would thus appear to have the potential to solve a number of the issues traditionally associated with RIEP (as mentioned above), such as relevance or actionability of research findings (Brown 2020).

But while these technologies may offer opportunities, they will not support improved teaching quality if not used effectively. As noted by some researchers, there are several potential risks in this area (Chounta et al. 2022). For instance, generative AI tools can sometimes produce outputs containing inaccuracies, including fictitious sources (Gustilo, Ong, and Rose Lapinid 2024). Missing or outdated information can also present a misleading account of a particular topic area (Davis and Jik Lee 2023). AI outputs therefore require validation through critical appraisal to ensure rigour, recency, context specificity, and the presence of a meaningful theory of action. AI-generated outputs should also be combined with an effective way of appraising impact, to assess whether they are delivering desirable improvements in teaching and learning outcomes.

While there is a growing body of research exploring educators' use of generative AI tools in their teaching (Celik et al. 2022; Chounta et al. 2022), little attention has been given to how teachers are using these tools to find and integrate research findings into their practice. Accordingly, this study sets out to make an important contribution by investigating how generative AI tools are contributing to and challenging teachers' RIEP. Based on these findings, we suggest reconceptualising RIEP theory to meet the evolving technological affordances and demands raised by generative AI. In so doing, we re-examined and extended a leading theory of RIEP (see Brown et al. 2022). This theory is summarised next.

A theoretical frame for exploring how generative AI can support RIEP

Research in the field of RIEP has been criticised for being 'under-theorised' (Brown et al. 2022). This criticism is significant because it suggests that researchers may not fully consider or address the complex range of factors that contribute to the research-practice gap. To address this issue and provide a theoretical foundation for our analysis, we utilised a novel framework first proposed by Brown et al. (2022). This framework draws upon Baudrillard's (1968) semiotic theory of consumption. Brown et al. (2022) theoretical framework enabled us to conceptualise teachers' use of research evidence via the medium of generative AI as a function of three interrelated factors:

- (1) *Benefits*: The primary question for educators is whether utilising research evidence is likely to yield positive benefits for their leadership, teaching practices, or professional learning. Additionally, educators must consider whether the perceived benefits of using research evidence outweigh those of alternative methods. These alternatives might include professional development courses, interactions with trusted colleagues, or resources available through social media.
- (2) *Costs*: The costs of utilising research can be multifaceted, extending beyond financial expense – such as subscription fees or paywalls – to include the time invested in searching for, understanding, and implementing research findings. There are also mental costs involved, since engaging with research can be a cognitively demanding process. As with benefits, such costs are relative to those associated with other sources of information, which may be more affordable, accessible, or easier to comprehend. Educators often assess costs in terms of whether they outweigh the potential benefits derived from using research evidence.

- (3) *Signification*: The concept of signification, as applied in this study, refers to the extent to which educators perceive research use as desirable. This form of desirability is distinct from the benefits associated with research use. Instead, it pertains to specific actions or behaviours that educators wish to be identified with. In the context of research use, desirability concerns whether educators want to be associated with engaging in academic research. This desire can be influenced by external expectations from colleagues or an internally motivated sense of professional identity, where engaging with research enhances an educator's professional self-conception.

As a consequence of the accelerating development and growing adoption of generative AI tools by teachers, there are reasons to reconsider whether such theories of RIEP still offer a valuable lens to understand teachers' use of research evidence. For instance, the perceived benefits/costs associated with finding and using academic research may change in response to the near-instant availability and accessibility of research outputs from AI tools. The theory may also face difficulties in explaining how the use of generative AI tools may undermine the perceived significance of research-informed teaching practices. This study aimed to shed light on the problem of how generative AI tools are contributing to and challenging teachers' RIEP and the theoretical issues connected to this. A two-part investigation was conducted. Part one reviewed the existing body of literature documenting teachers' use of generative AI tools to find and incorporate research evidence into their teaching practice. Part two then drew on insights from part one to propose modifications to RIEP theory in order to align with technological developments. Through this investigation two research questions were addressed:

- (1) How do teachers use generative AI tools to find and incorporate research evidence into their practice, and what reasons, benefits, and challenges are associated with this research use?
- (2) What are the practical and theoretical implications of teachers' use of generative AI tools for RIEP?

Materials and methods

To address the first research question a systematic review was carried out. This type of review involved an exhaustive search of the literature, helping to collate available evidence on teachers' use of generative AI tools across different contexts. A mixed studies review approach was employed to find and analyse both qualitative and quantitative evidence pertinent to the research question (Grant and Booth 2009). Following a pre-prepared protocol aligned with PRISMA guidelines, various steps were taken to reduce biases and the possibility of overlooking relevant studies.

Search strategy

Sources of literature

A range of electronic databases were searched to identify relevant studies across different disciplines. The following academic databases were searched: ERIC;

Education Database (via ProQuest); Australian Education Index; British Education Index; Education Abstracts; Education Source; Educational Administration Abstracts; Teacher Reference Center; and Web of Science (Core Collection). These databases were chosen on the basis that they are some of the most comprehensive sources of academic literature reporting on teaching and educational technology use. To supplement these databases, a search was carried out on Google Scholar. The reference lists of past literature reviews on teachers' use of generative AI tools were also scanned to find any further studies pertinent to the research questions. In addition, various sources of grey literature were searched. These included studies and reports published by governments and intergovernmental organisations (e.g. OECD, UNESCO, UK, and US governments), as well as those published by non-governmental/research organisations (e.g. National Foundation for Educational Research, European Educational Research Association). It was important to search reputable sources of grey literature given the recent proliferation of generative AI tools in education and the high probability that some relevant research would be published outside of traditional academic outlets due to the shorter time lag to publication.

Inclusion criteria

For studies to be included in the review they needed to meet the following inclusion criteria: 1) report novel empirical evidence concerning teachers' use of generative AI tools to find and/or incorporate research evidence into their practice; 2) detail relevant qualitative and/or quantitative evidence relating to pre-service or in-service teachers working at any level of education in any country; 3) be published in the English language at any time point prior to 1 June 2024 (date the literature search commenced); and 4) explicitly report the research aims, context, participants, study design, and findings. Studies reporting on teachers' use of digital technologies not typically classified as generative AI (e.g. standard Google search) were excluded from the review. Given the recency of generative AI tools and limited amount of empirical research on teachers' RIEP using these tools, the criteria were formulated to include as much available evidence as possible. To increase the scope of the search, we used a broad definition of teachers as those who are qualified or in the process of qualifying as teachers.

Search terms

To achieve a high level of both sensitivity and precision, various search terms, including multiple synonyms and international spelling variations, Boolean operators, and special characters were tested and finalised for each database. For example, different synonyms of generative AI and teacher (e.g. instructor, educator, etc.) were tested and in some cases disregarded because they did not increase sensitivity significantly and lowered precision substantially by more than doubling the number of results. When searching on EBSCOhost, the following search terms were used on the advanced search function: teachers OR teaching OR educators OR school leaders AND generative AI OR ChatGPT AND research OR evidence. The following search options were also selected: language = English; apply related words; and apply equivalent subjects. Search terms were largely consistent across the databases, with minor changes made to achieve an optimal balance of sensitivity and precision and/or to fit within the functionality of each platform.

Screening search results

An independent double screening procedure was used to identify eligible studies for the review. All 790 search results were screened by the lead researcher against the inclusion criteria. Independent screening was also carried out by three other senior academics who each screened approximately one-third of the 790 studies each. The inter-rater reliability score, Cohen's K, was calculated as 0.59 (classified as moderate agreement). The lead reviewer and second reviewer reached the same decision in 92% of cases. Meanwhile, the third and fourth reviewers agreed with the lead reviewer's decision in 88% and 90% of cases, respectively. Differences in reviewer decisions were resolved through in-person discussion. At the title and abstract screening stage, after removing duplicates, 790 studies were identified. Following abstract screening, 129 studies were considered for full-text screening. In total, 19 studies met the inclusion criteria after full-text screening. The full screening process is depicted in [Figure 1](#). A full record keeping log of the search and screening process can be accessed in the specified data repository.

Quality assessment and data analysis

To assess the methodological quality of the 19 studies included in the review, the Mixed Methods Appraisal Tool was used. This tool has been extensively tested and includes a checklist with specific criteria for assessing the quality of quantitative, qualitative, and mixed methods studies (see Hong et al. [2018](#)). After extracting relevant data from the studies, a thematic analysis was conducted. This method was chosen because it allows for the aggregation and analysis of varying sources of data (Braun and Clarke [2006](#)), capturing the study participants' experiences with generative AI tools to uncover repeated patterns of meaning. Using a deductive approach based on the categories specified in Brown et al. ([2022](#)) theory, initial codes were first developed to highlight various pedagogical, practical, and professional identity issues emerging from the data. Codes were subsequently grouped into provisional themes with accompanying data extracts. Identified themes were reviewed and refined in relation to the entire dataset (Braun and Clarke [2006](#)).

To mitigate bias during the analysis process, critical conversations took place between the study researchers after reading the included studies. These discussions ensured all codes and themes were developed and scrutinised by more than one researcher, taking into consideration the studies' differing contexts, reliability of the evidence, and the most plausible theoretical interpretations. Brown et al. ([2022](#)) framework was utilised to inform the development of the themes and evaluate the benefits and challenges of teachers' use of generative AI tools in their RIEP. The thematic analysis was partially guided by core elements of the framework, including the benefits, costs, and signification associated with teachers' use of research evidence. However, insights from the inductively generated codes and themes were also used to critique and highlight potential limitations in the theory. Documents detailing the entire coding process and development of the themes can be found in the specified data repository.

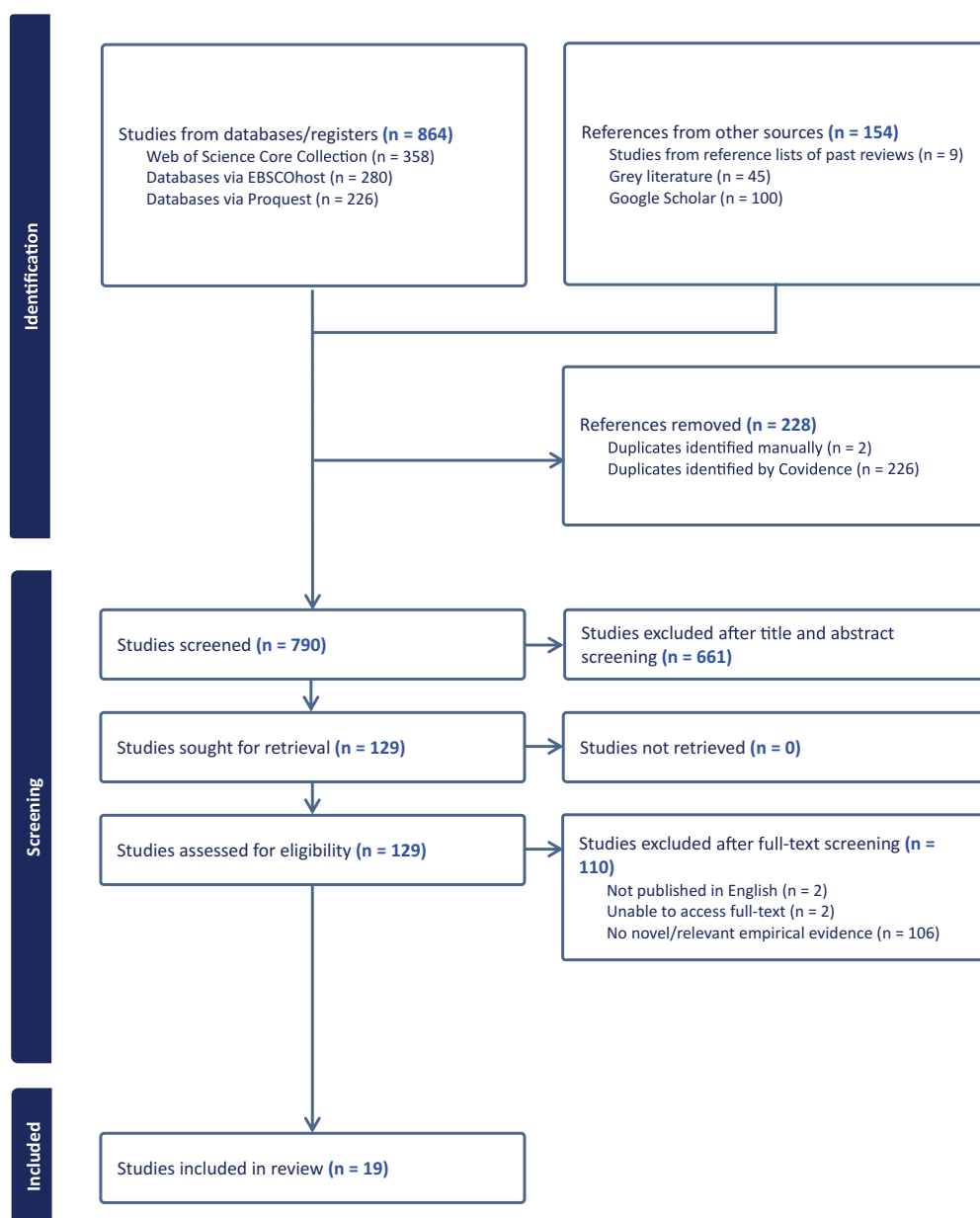


Figure 1. Overview of the literature search and screening process.

Study characteristics

Before reporting findings from the thematic analysis, information on the design and context of the 19 included studies is provided in Table 1. Most studies included in the review were conducted in a higher education context (62.96%). Meanwhile, a smaller number of studies involved teachers working at secondary (18.52%), primary (14.81%) and/or preschool level (3.7%). Although these research studies are not concentrated in

Table 1. Contextual information and design of the studies included in the review.

Country	Level of education	Subject area	Study design & method	Sample size
India (<i>n</i> = 2)	Higher education	Multiple (<i>n</i> = 8)	Qual – interviews/focus	10 or less
Oman (<i>n</i> = 2)	(<i>n</i> = 17)	English as a foreign	groups (<i>n</i> = 16)	(<i>n</i> = 4)
Turkey (<i>n</i> = 2)	Secondary (<i>n</i> = 5)	language (<i>n</i> = 5)	Qual – participant journals	11 to 25
Saudi Arabia (<i>n</i> = 2)	Primary (<i>n</i> = 4)	Computer science (<i>n</i> = 1)	(<i>n</i> = 3)	(<i>n</i> = 4)
Thailand (<i>n</i> = 1)	Preschool (<i>n</i> = 1)	Dental education (<i>n</i> = 1)	Qual – survey (<i>n</i> = 2)	26 to 50
Spain (<i>n</i> = 1)		Management and finance	Quan – survey (<i>n</i> = 7)	(<i>n</i> = 4)
Philippines (<i>n</i> = 1)		(<i>n</i> = 1)		51–200
Multiple: USA, Chile,		Teacher education (<i>n</i> = 1)		(<i>n</i> = 4)
Egypt (<i>n</i> = 1)		Mathematics (<i>n</i> = 1)		201–500
United Arab Emirates		Language education		(<i>n</i> = 3)
(<i>n</i> = 1)		(<i>n</i> = 1)		
China (<i>n</i> = 1)				
South Korea (<i>n</i> = 1)				
Iran (<i>n</i> = 1)				
Turkey (<i>n</i> = 1)				
England/UK (<i>n</i> = 1)				
Vietnam (<i>n</i> = 1)				
Hong Kong (<i>n</i> = 1)				

a narrow band of countries, a larger proportion of studies were conducted in countries across Asia (42.11%), the Middle East (31.58%), and Europe (21.05%). Many studies involved teachers who teach across different subject areas (42.11%). Meanwhile, five studies exclusively focused on teachers who are teaching English as a foreign language (26.32%). Smaller proportions of studies involved teachers of computer science, dental education, management and finance, education, mathematics, and language education (5.26% for each subject).

Most studies employed a qualitative design using interviews/focus groups (57.14%) and/or participant journals/narratives (10.71%). A smaller proportion of studies used a quantitative design with survey data collection methods (25%). Approximately 21.05% of studies used a sample size of 10 or less. Another 42.11% of studies used a sample size between 11 and 50 participants. The remaining studies used a sample size between 51 and 500 participants (36.84%). Nine out of the 19 studies included both male and female teachers, with one study involving female only participants (9 other studies did not report the participants' gender). Most studies did not disclose the socio-economic background of the participants.

Results from the Mixed Methods Appraisal Tool revealed that 11 out of 12 qualitative studies met all five criteria used to assess the quality of qualitative research. In contrast, only one quantitative study met all five criteria used to assess the quality of descriptive quantitative research (two other studies met three of the five criteria). These studies did not use a sampling strategy suited to the research questions (e.g. did not use a probability sampling method) and they did not ensure the sample was representative of the target population. Of the four mixed methods studies, three studies achieved three out of the five quality criteria. Overall, the quality of the included studies was moderately high, with the quantitative evidence being of a lower standard.

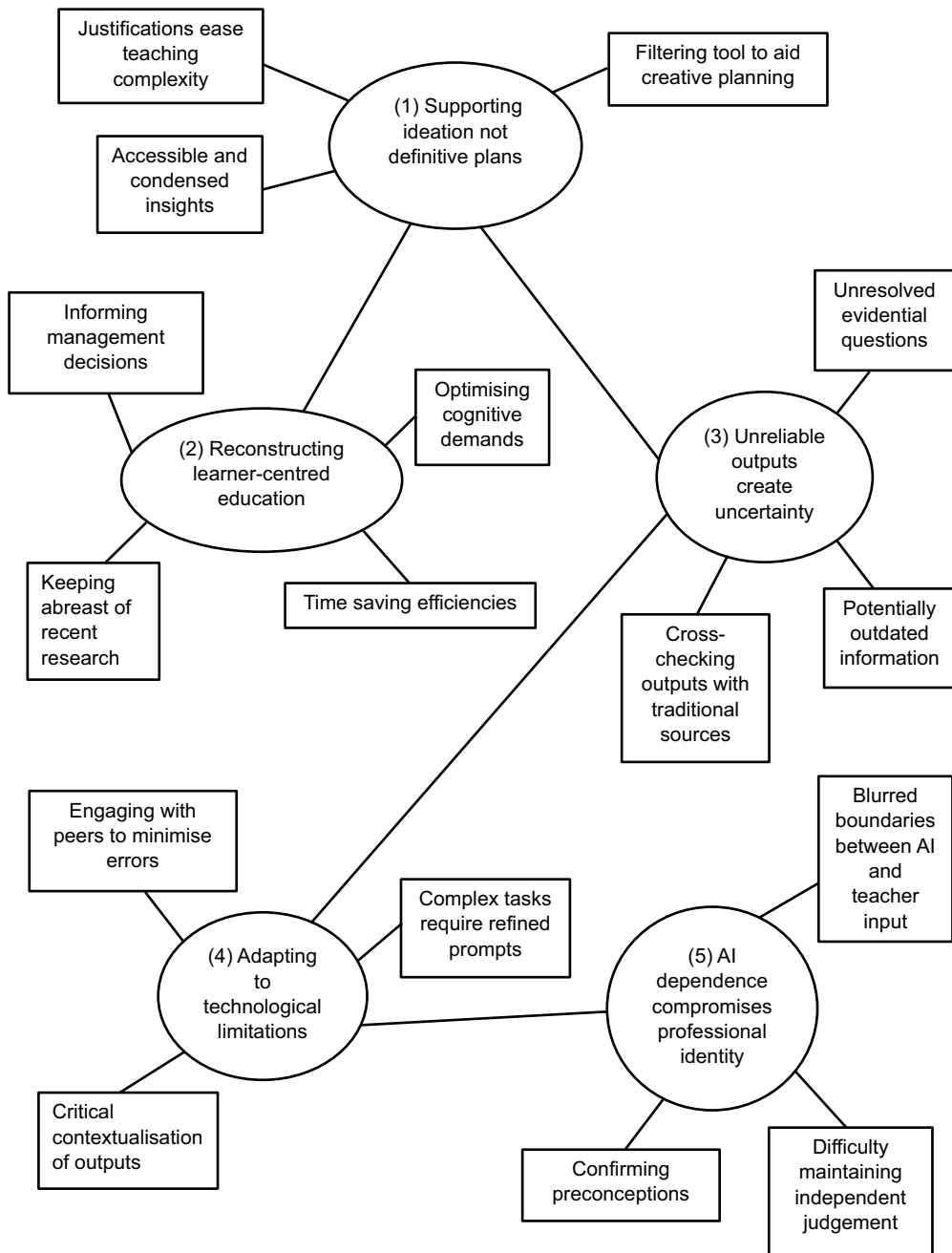


Figure 2. Thematic map of findings from the systematic review.

Results

Findings from the thematic analysis of the review evidence are depicted in [Figure 2](#). Five themes and various sub-themes were identified. Each of the five themes is elucidated in the following sub-sections with accompanying extracts/

evidence from the included studies. Fifteen studies relate to Theme 1; six studies to Theme 2; nine studies to Theme 3; nine studies to Theme 4; and four studies to Theme 5.

Theme 1: supporting ideation not definitive plans

In addressing the first research question, several benefits were identified concerning teachers' use of generative AI tools in their RIEP. A recurrent finding across multiple studies was teachers' use of generative AI tools (e.g. ChatGPT) to find research insights to inform their ideation processes rather than to retrieve or identify definitive teaching plans/solutions (Al-Mughairi and Bhaskar 2024; ElSayary 2024; Hasanein and Sobaih 2023; Prakasha, Sanskriti, and Ishani 2024; Tapan-Broutin 2024). In research conducted by Tapan-Broutin (2024), 10 pre-service mathematics teachers used ChatGPT to consult for research ideas and seek out suggestions for practice. Instead of treating the AI outputs as ready-to-implement solutions, teachers often employed ChatGPT as a filtering tool to eliminate unviable ideas and to consider 'appropriate suggestions to create a new, original, and personalised lesson plans according to their own schemes' (Tapan-Broutin 2024, 171). Suggestions from generative AI tools often served as a source of creative inspiration for teachers. For instance, in a study involving grade 6–12 teachers working in the United Arab Emirates, one interviewee remarked that unlike with Google Search, when using ChatGPT 'I [can more effectively] spend time combining information and tailoring them to develop creative ideas in my lesson plan' (ElSayary 2024, 938).

Another unique advantage of generative AI tools over a standard Google Search was reported to be the supporting AI-generated explanations as to how and why a lesson or approach could be tailored in a certain way (ElSayary 2024; Prakasha, Sanskriti, and Ishani 2024; Tapan-Broutin 2024). These accompanying justifications were purported to help clarify complex links between theory and practice, giving teachers a greater sense of confidence that they are developing well-informed and actionable teaching approaches. Further to these benefits, various studies indicate that teachers value the accessible and condensed research insights from generative AI tools (Bhaskar and Rana 2024; Nguyen Thi Thu 2023; Uribe et al. 2024). A survey of 280 educators in the UK found that teachers used generative AI tools to research a topic or concept, as well as for summarising articles and books (Department for Education 2024). The comprehensible and compact outputs produced by the AI technology allowed teachers to readily engage with 'brief summaries of contemporary research topics' (Bhaskar and Rana 2024, 8).

Theme 2: reconstructing learner-centred education

The second theme emerging from the review findings refers to teachers' use of generative AI tools to find evidence to support the development of learner-centred activities and practices. Several studies reported teachers' use of ChatGPT to create or combine evidence-based learner-centred activities to meet the unique learning needs of students (ElSayary 2024; Govindarajan and Christuraj 2024; Tapan-Broutin 2024). But rather than merely using the technology to discover a specific learner-centred approach to directly apply in their practice (e.g. problem-based learning), some teachers entered questions/prompts into the generative AI tools to reformulate their learner-centred plans and

activities to align with research-informed theories, such as Bloom's taxonomy (ElSayary 2024). In this way, teachers looked to reconstruct a learner-centred education from the ground up, recreating and scaffolding materials and plans to be more optimised for the different cognitive capacities and interests students possess (Govindarajan and Christuraj 2024; Kartal 2024).

To facilitate their learner-centred practice, teachers were found using generative AI tools to access the latest research evidence and pedagogical techniques (Bhaskar and Rana 2024; Tapan-BROUTIN 2024). Teachers took advantage of these tools by bypassing historic barriers, such as searching articles individually or navigating paywalls. A study with 48 teachers in India reported that AI tools have enabled teachers to overcome barriers associated with traditional literature searching and collate multiple findings at once (Bhaskar and Rana 2024). Another frequently reported benefit of generative AI tools was the time-saving efficiencies they afforded educators. Teachers in multiple studies used tools such as ChatGPT to streamline the search process and efficiently generate summaries of educational research (Bhaskar and Rana 2024; Derakhshan and Ghiasvand 2024). This made it more practically feasible for teachers to develop research-informed resources and approaches within their limited working hours. Furthermore, review findings also indicated that generative AI tools were used by some teachers to support classroom- and school-level decision-making (Chiu 2023; Kartal 2024). Research involving 88 primary and secondary schoolteachers in China found participants utilised ChatGPT to find different sources of evidence for school management problems, including formulating school policy on attention deficit hyperactivity disorder (Chiu 2023).

Theme 3: unreliable outputs create uncertainty

While these findings indicate a range of practical, pedagogical, and mental benefits could contribute to teachers' use of research evidence via generative AI tools, according to RIEP theory, perceived costs may counteract teachers' research consumption (Brown et al. 2022). For instance, although AI tools may in certain regards raise teachers' confidence by providing supporting justifications for teaching plans, this feeling was juxtaposed by some teachers experiencing uncertainty due to the occasionally unreliable AI outputs and unresolved evidential questions (Al-Mughairi and Bhaskar 2024; Govindarajan and Christuraj 2024). In survey research undertaken amongst 100 teachers and researchers at a university in the Philippines, 25% of respondents expressed concerns over the reliability, authenticity, and validity of the outputs (Gustilo, Ong, and Rose Lapinid 2024). Participants explained that some sources were not searchable on Google or other websites. As such, teachers were sometimes left in a predicament as to whether believe or disbelieve the authenticity and validity of the outputs. Teachers from Thailand further explained that ChatGPT occasionally left them with unanswered questions which cast doubt on what the right research-informed approach was to take (Ulla, Perales, and Busbus 2023).

Whilst generative AI tools were used by teachers to find recent research and pedagogical techniques, in contradistinction, some teachers encountered the problem that the technologies sometimes produced outdated or inaccurate information. Davis and Jik Lee

(2023) reported that Korean schoolteachers retrieved outdated or incorrect information when using ChatGPT to produce evidence-based lesson plans and activities. At the time the tool's training data was limited to September 2021. To overcome problems of potentially inaccurate and outdated information or fabricated sources, participants in several studies explained that they would cross-check the outputs with traditional sources of information (Kartal 2024; Ulla, Perales, and Busbus 2023). For example, EFL pre-service teachers in Turkey explained that to verify the accuracy of AI outputs, they would cross-check the outputs against sources such as textbooks, articles, and course materials (Kartal 2024).

Theme 4: adapting to technological limitations

In contending with the occasionally unreliable outputs produced by generative AI tools, evidence suggests that some teachers would intentionally adapt their approach to compensate for the limitations of the tools (Al-Mughairi and Bhaskar 2024; Alammari 2024; Moorhouse and Kohnke 2024). Tapan-Broutin (2024) reported that upon noticing deficiencies or inaccuracies in the outputs produced by ChatGPT, teacher participants would often try to rephrase their question/prompt or directly request corrections from the AI tool to realign the output with their requirements. However, several teachers found that this still led to irrelevant responses, encountering difficulties in retrieving evidence-informed insights related to specific classroom management, pedagogical, and curricular challenges they faced (Tapan-Broutin 2024). In some studies, teachers adjusted their use of the generative AI tool based upon the complexity of the teaching task (ElSayary 2024; Galindo-Domínguez et al. 2024). A schoolteacher based in the United Arab Emirates explained that on more complex teaching tasks, they would sometimes obtain less meaningful responses and need to refine their language or employ multiple prompts to gradually extract relevant insights from the tool (ElSayary 2024).

Despite various teachers claiming generative AI tools could aid them in developing contextually relevant teaching approaches, some participants nonetheless suggested that outputs needed to be approached more critically and cautiously due to the potentially generic or uncertain nature of the outputs. A teacher working in a higher education institution in Oman explained: 'I need to be cautious and critically evaluate the information generated by ChatGPT' (Al-Mughairi and Bhaskar 2024, 7). To avoid uncritically accepting the outputs produced by generative AI tools, teachers in several studies reported engaging with peers to confirm the research-based suggestions, discuss discrepancies, and consider limitations with the practical application of the output (Chiu 2023; Kartal 2024). As one EFL teacher from Turkey remarked: 'I always make sure to compare the content generated by ChatGPT with other sources and engage in discussions with my colleagues to ensure its accuracy and credibility' (Kartal 2024, 631).

Theme 5: AI dependence compromises professional identity

As well as adapting to the complex nexus of costs and benefits associated with using generative AI tools, utilisation of these tools also held significance for teachers' professional identity (ElSayary 2024; Kartal 2024; Moorhouse and Kohnke 2024).

Moorhouse and Kohnke's (2024) study with teacher educators in Hong Kong found that some educators were uncertain how to reconcile their professional role with the technology. For some teachers, including several EFL teachers from a public university in Turkey, their immersion in the AI platform and co-creation of research-informed plans meant that the boundaries between the AI's contributions and their own personal contributions became somewhat blurred (Kartal 2024). As Kartal (2024) explained, the difficulty of clearly distinguishing the teachers' own thought from the AI-generated information raised doubts among some participants about their professional value and autonomous judgement. The teachers were conscious of becoming overly dependent on generative AI tools. They were mindful of the risk of failing to achieve a balance between AI-generated ideas and their own independent research and thinking to maintain their professionalism and critical thinking capacities (Kartal 2024).

Conversely, teachers displayed behaviours with generative AI tools which suggest they were seeking to validate their preconceptions (ElSayary 2024; Tapan-BROUTIN 2024). Findings documented by Tapan-BROUTIN (2024) revealed that some pre-service mathematics teachers used ChatGPT to reaffirm their own ideas and to validate their pre-existing plans. In this case, teachers were not necessarily intending to use AI tools to uncover new insights, but to socially signal the value of their professional judgement and teaching identity by finding evidence to justify the teaching practices they claimed personal ownership of. These findings present a challenge to Brown et al. (2022) theory on RIEP by showing that teachers' identification with research use, via the medium of generative AI, may include a combination of positive and negative connotations.

Discussion

So far, this study has detailed existing evidence on the benefits and challenges presented by teachers' use of generative AI tools to find and incorporate research findings into their practice. In the following discussion, findings from the review are critically evaluated in relation to the three components of Brown et al. (2022) theory of RIEP. Practical and theoretical implications for RIEP are also discussed.

Benefits and costs of using generative AI for RIEP

Findings from the thematic analysis revealed that many teachers found value in using generative AI tools to find research evidence to support their ideation processes rather than for obtaining definitive plans. As prior studies suggest, engaging with research findings in this open-ended way may serve to promote teachers' divergent thinking and lead to positive outcomes for teaching and learning (Cain 2016; Godfrey 2016). Teachers could assimilate new AI-generated research ideas into their existing schemas of teaching (Tapan-BROUTIN 2024). This can enable teachers to draw on more elaborate schemas to imagine, develop, and trial new teaching practices. However, when integrating research-informed ideas into their ideation processes, there is a risk teachers may assign equal weighing to all sources of information packaged in the AI outputs. It is thus important for teachers to critically appraise each research-informed output, to consider its

methodological strengths and limitations, and to judge its merits with regard to the ethical and practical constraints of the educational context (Biesta 2007; Wrigley 2018).

Findings from this review additionally showed that teachers working across different contexts valued the accessible and condensed research-informed outputs produced by generative AI tools (Bhaskar and Rana 2024; Nguyen Thi Thu 2023). Because technologies such as ChatGPT can collate different research findings and convey the insights using comprehensible language, they can save teachers time in moving between multiple websites to locate information and help to demystify complex scientific terminology. These benefits could be crucial in promoting the wider adoption of RIEP given that teachers and school leaders have often reported a lack of time and esoteric language as key barriers preventing their engagement with research (Brown 2020; Cain 2016).

A further benefit uncovered in this review was teachers' use of generative AI tools to find research in support of their construction of learner-centred activities and approaches (ElSayary 2024; Govindarajan and Christuraj 2024). Instead of setting out to identify and apply an existing learner-centred approach, some teachers entered prompts/questions into the generative AI tool to recreate their learner-centred plans and activities to align with research-informed theories (ElSayary 2024). A bottom-up approach such as this may help teachers overcome a common challenge associated with RIEP. That is the adoption of a generic research-derived practice without due consideration of or adaptation to the educational context and unique characteristics of the students (Biesta 2007; Wrigley 2018). This contributes new insights to RIEP theory by highlighting how teachers can engage with evidence during a process of pedagogical reconstruction, rather than only using research as a means to validate an existing practice or to justify a shift between practices.

Findings from the thematic analysis also revealed that there were several costs associated with teachers' use of generative AI in their RIEP. Multiple studies reported that teachers sometimes found the AI-generated outputs to be unreliable. In some cases, this included teachers encountering made up sources, outdated or inaccurate information, as well as unanswered evidential questions. For some teachers, such experiences fostered a degree of uncertainty and cast doubt on what would be the most suitable teaching approach to take (Ulla, Perales, and Busbus 2023). As explained by Brown et al. (2022), costs of this sort can place a burden on teachers' time and mental resources as they may need to engage in extra work to verify the AI outputs using external sources. Some teachers did indeed look to minimise these costs by cross-checking outputs with traditional sources of evidence such as textbooks or articles (Kartal 2024). Nonetheless, according to RIEP theory, should these mental costs accumulate to the point where they are perceived to exceed the costs of alternative sources of information (e.g. social media), this could significantly limit teachers' utilisation of research via generative AI tools (Brown et al. 2022).

Practical implications

To help pre-service and in-service teachers overcome the psychological and time costs they may encounter while using generative AI tools in their RIEP, schools and teacher education provisions can provide practical guidance on crafting more sophisticated

prompts to efficiently retrieve relevant research findings. Moreover, to equip teachers to discern fictitious sources or misleading research summaries in generative AI outputs, teacher educators can guide pre-service teachers through worked examples that require cross-checking AI-generated outputs and references against academically reliable sources. With support from teacher educators, pre-service teachers can continue to practice and benefit from reconstructing learner-centred plans and activities with generative AI tools. This pedagogical expertise can be fostered through opportunities for experimentation and reflective practice during the early stages of teachers' careers. Teacher educators can also encourage apprentice teachers to collaborate with experienced peers to learn how to critically comprehend the contextual and practical application of research-informed AI outputs (Chiu 2023; Kartal 2024).

Signification of using generative AI for RIEP

According to RIEP theory set out Brown (2020), signification is a distinct concept referring to specific actions, values, or behaviours educators want to be identified with through their engagement with academic research. Findings from this review partly challenge this conception by illustrating how a teacher's professional identity may, in some circumstances, be compromised as a result of their engagement with research via generative AI. Evidence revealed that some teachers experienced issues distinguishing their personal contributions from those of the generative AI tool when developing research-informed teaching plans (Kartal 2024). The relative value or identifiable contribution of the teacher was, for some, becoming uncertain. This may be attributable to the perception among educators that their creativity and judgement had become less prominent during the search process and development of research-informed teaching practices using generative AI tools. As generative AI tools can independently generate novel research summaries, appraisals, and research-informed teaching plans, the perceived or signified value a teacher might normally gain from developing research-informed practices using traditional means may be diminished (i.e. not personally finding publications and independently crafting teaching plans).

Moreover, some teachers expressed concerns that the regular consumption and application of research using AI may create a dependency that undermines their sense of professional identity (Kartal 2024). Several teachers were thus conscious of maintaining a balance between research insights generated by generative AI tools and their own independent research and thought to preserve their sense of professionalism and critical thinking capacities. However, in other respects, teachers' use of generative AI tools may also serve to enhance their professional conception by offering a source of validation for their existing teaching beliefs and plans. Findings from this review indicated that some teachers approached the AI tools with the intent of affirming their prior judgements and plans using AI-generated evidence-based recommendations (Tapan-BROUTIN 2024). This finding echoes the seminal work of Weiss (1979) who identified symbolic uses of research evidence – which in certain instances includes using research to justify an existing decision. In this case, however, research and AI may be providing a symbolic, legitimising function: with practitioners identifying research and AI as signifying a stronger power or offering a greater currency than their professional experience alone.

Theoretical implications

The themes uncovered in this study reveal new interconnections between the three components of Brown et al. (2022) theory of RIEP (see lines in Figure 2). Connections between themes one and two suggest that benefits acquired from generative AI tools for RIEP, such as support for ideation and reconstructing learner-centred practices, can be mutually reinforcing. Teachers' AI-informed ideation can feed into their reconstruction of learner-centred practices. Likewise, teachers' active engagement in reformulating their learner-centred practices with generative AI can promote further ideation processes. Concurrently, linkages between themes 1 and 3 show that RIEP benefits from generative AI can be inhibited by significant costs. Teachers' ideation processes can be disrupted by occasionally unreliable outputs produced by generative AI tools. In contrast to predictions of RIEP theory (Brown et al. 2022), this interplay between RIEP benefits and costs might be taking on a more fluid and intensified form with generative AI. While wide-ranging research-informed ideas can be immediately obtained and applied using these technologies, teachers' RIEP can now be quickly derailed by encountering unreliable outputs and new levels of technological uncertainty.

This evolving dynamic has created an imperative for educators to take on a more adaptive role. Links between themes three and four illustrate that teachers are in some cases adapting to the limitations of generative AI tools by cross-checking sources and engaging with peers. However, in using and adapting to generative AI tools, findings connecting themes 4 and 5 suggest that teachers could encounter new complexities for their professional identity. RIEP theory has historically recognised that research consumption can afford positive signification to teachers (Brown et al. 2022). However, teachers' retrieval and use of research via the medium of generative AI may now entail divergent consequences not previously theorised. This includes the consequence that by increasingly offloading their research-informed planning and ideation processes to generative AI, and by adapting their behaviours to compensate for outputs produced by an external 'intelligence', teachers may lose a sense of their professional identity as autonomous and imaginative practitioners. In light of these findings, future developments in RIEP theory will need to reconceive how teachers, research, and advance digital technologies intersect and create new conditions for educators' professional signification and evidence-informed practice.

Review limitations

Whilst various steps were taken to limit bias and ensure the rigour of the review process, there are some limitations associated with this study. First, only studies published in the English language were eligible for inclusion in this review. This may mean that relevant research findings published in other languages have not been taken into account. Moreover, as most studies employed qualitative designs and many quantitative studies did not employ probability sampling methods, the participants in most studies may not be representative of the wider teaching populations. Furthermore, while 19 studies were eligible for inclusion in this review, it is important to note that most included studies only reported a few paragraphs or sentences concerning teachers' use of generative AI tools to find and incorporate research evidence into their practice. These studies tended to focus

on teachers' use of generative AI tools in a more general sense than their research use in particular. We therefore recommend for more dedicated empirical research to be carried out on teachers' use of generative AI tools to support their RIEP.

Conclusion

This study reviewed evidence on the benefits and challenges presented by teachers' use of generative AI tools to find and integrate research findings into their practice. Several implications for RIEP theory were also uncovered. Most notably, evidence suggests generative AI tools can mutually reinforce teachers' ideation processes and reconstruction of learner-centred plans. Yet, unreliable outputs can also disrupt these processes and create challenging conditions for teachers' professional identity. These findings raise several important issues for educational policy and practice. To support teachers in using generative AI tools in their RIEP, policymakers and school leaders can provide ongoing professional development opportunities to aid teachers in crafting prompts that extract contextually relevant insights and feed into a meaningful theory of action. Educational institutions can play a role in nurturing teachers' AI confidence and positive professional identity by establishing communities of practice in which teachers not only share best practices concerning generative AI use but also guidance on how to synthesise and weigh up different sources of evidence based on informed human-led judgement.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Stephen Sowa works as a Teaching Fellow at the University of Southampton, Southampton Education School. His research interests include topics such as the artistry of teaching and the implications of technological advances for teaching.

Chris Brown is Professor of Education at the University of Southampton, Head of the Southampton Education School, and Distinguished Visiting Professor, University of Tübingen. His published works span the broad sphere of research, evidence, and ideas-use in education and society more broadly.

Tae-Hee Choi is an Associate Professor and the Director of Internationalisation at Southampton Education School, University of Southampton. Her research interests include education policy and change in the global context, and teacher development therein.

Rachele Newman is a Principal Teaching Fellow and Director of Initial Teacher Education at Southampton Education School, University of Southampton. Rachele is passionate about supporting and developing teachers at all stages of their careers.

ORCID

Stephen Sowa  <http://orcid.org/0000-0002-7095-1843>

Chris Brown  <http://orcid.org/0000-0002-9759-9624>

Tae-Hee Choi  <http://orcid.org/0000-0001-8840-4082>

Data availability statement

The datasets collected and analysed during the study are available in the 'Generative AI tools and teachers' research-informed practice' repository, <https://doi.org/10.5258/SOTON/D3185>.

References

- Al-Mughairi, H., and P. Bhaskar. 2024. "Exploring the Factors Affecting the Adoption AI Techniques in Higher Education: Insights from Teachers' Perspectives on chatgpt." *Journal of Research in Innovative Teaching & Learning*. <https://doi.org/10.1108/JRIT-09-2023-0129>.
- Alammari, A. 2024. "Evaluating Generative AI Integration in Saudi Arabian Education: A mixed-Methods Study." *PeerJ Computer Science* 10:e1879. <https://doi.org/10.7717/peerj-cs.1879>.
- Baudrillard, J. 1968. *The System of Objects*. London: Verso.
- Bhaskar, P., and S. Rana. 2024. "The chatgpt Dilemma: Unravelling Teachers' Perspectives on Inhibiting and Motivating Factors for Adoption of chatgpt." *Journal of information, Communication and Ethics in Society* 22 (2): 219–239. <https://doi.org/10.1108/JICES-11-2023-0139>.
- Biesta, G. 2007. "Why 'What Works' Won't Work: Evidence-Based Practice and the Democratic Deficit in Educational Research." *Educational Theory* 57 (1): 1–22. <https://doi.org/10.1111/j.1741-5446.2006.00241.x>.
- Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Brown, C. 2020. *The Networked School Leader: How to Improve Teaching and Student Outcomes Using Learning Networks*. London: Emerald.
- Brown, C., S. MacGregor, J. Flood, and J. Malin. 2022. "Facilitating Research-Informed Educational Practice for Inclusion. Survey Findings from 147 Teachers and School Leaders in England." *Frontiers in Education* 7:1–17. <https://doi.org/10.3389/feduc.2022.890832>.
- Cain, T. 2016. "Research Utilisation and the Struggle for the Teacher's Soul: A Narrative Review." *European Journal of Teacher Education* 39 (5): 616–629. <https://doi.org/10.1080/02619768.2016.1252912>.
- Cain, T., S. Brindley, C. Brown, G. Jones, and F. Riga. 2019. "Bounded decision-Making, Teachers' Reflection and Organisational Learning: How Research Can Inform Teachers and Teaching." *British Educational Research Journal* 45 (5): 1072–1087. <https://doi.org/10.1002/berj.3551>.
- Celik, I., M. Dindar, H. Muukkonen, and S. Järvelä. 2022. "The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research." *Technology Trends* 66 (4): 616–630. <https://doi.org/10.1007/s11528-022-00715-y>.
- Chiu, T. K. F. 2023. "The Impact of Generative AI (genai) on Practices, Policies and Research Direction in Education: A Case of Chatgpt and Midjourney." *Interactive Learning Environments* 32 (10): 6187–6203. <https://doi.org/10.1080/10494820.2023.2253861>.
- Chounta, I.-A., E. Bardone, A. Raudsep, and M. Pedaste. 2022. "Exploring Teachers' Perceptions of Artificial Intelligence as a Tool to Support Their Practice in Estonian k-12 Education." *International Journal of Artificial Intelligence in Education* 32 (3): 725–755. <https://doi.org/10.1007/s40593-021-00243-5>.
- Cordingley, P. 2015. "The Contribution of Research to Teachers' Professional Learning and Development." *Oxford Review of Education* 41 (2): 234–252. <https://doi.org/10.1080/03054985.2015.1020105>.
- Davis, R. O., and Y. Jik Lee. 2023. "Prompt: ChatGPT, Create My Course, Please!" *Education Science* 14 (1): 24. <https://doi.org/10.3390/educsci14010024>.
- Department for Education. 2024. *Generative AI in Education: Educator and Expert Views*. London: UK Government Department for Education. <https://www.gov.uk/government/publications/generative-ai-in-education-educator-and-expert-views>.
- Derakhshan, A., and F. Ghiasvand. 2024. "Is chatgpt an Evil or an Angel for Second Language Education and Research? A Phenomenographic Study of research-Active EFL Teachers'

- Perceptions." *International Journal of Applied Linguistics* 34 (4): 1246–1264. <https://doi.org/10.1111/ijal.12561>.
- ElSayary, A. 2024. "An Investigation of Teachers' Perceptions of Using chatgpt as a Supporting Tool for Teaching and Learning in the Digital Era." *Journal of Computer Assisted Learning* 40 (3): 931–945. <https://doi.org/10.1111/jcal.12926>.
- Galindo-Domínguez, H., N. Delgado, D. Losada, and J.-M. Etxabe. 2024. "An Analysis of the Use of Artificial Intelligence in Education in Spain: The in-Service Teacher's Perspective." *Journal of Digital Learning in Teacher Education* 40 (1): 41–56. <https://doi.org/10.1080/21532974.2023.2284726>.
- Godfrey, D. 2016. "Leadership of Schools as research-Led Organisations in the English Educational Environment: Cultivating a research-Engaged School Culture." *Educational Management Administration and Leadership* 44 (2): 301–321. <https://doi.org/10.1177/1741143213508294>.
- Govindarajan, R., and G. Christuraj. 2024. "Opportunities and Challenges of Using ChatGPT in the ELT Scenario of Utas, Nizwa, Oman." *Journal of Teaching English for Specific and Academic Purposes* 11 (3): 593–605. <https://doi.org/10.22190/jtesap230529046g>.
- Grant, M. J., and A. Booth. 2009. "A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies." *Health Information & Libraries Journal* 26 (2): 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>.
- Gustilo, L., E. Ong, and M. Rose Lapinid. 2024. "Algorithmically-Driven Writing and Academic Integrity: Exploring Educators' Practices, Perceptions, and Policies in AI Era." *International Journal for Educational Integrity* 20 (1). <https://doi.org/10.1007/s40979-024-00153-8>.
- Hasanein, A. M., and A. E. E. Sobaih. 2023. "Drivers and Consequences of chatgpt Use in Higher Education: Key Stakeholder Perspectives." *European Journal of Investigation in Health, Psychology and Education* 13 (11): 2599–2614. <https://doi.org/10.3390/ejihpe13110181>.
- Hong, Q. N., P. Pluyea, S. Fàbreguesb, G. Bartletta, F. Boardmanc, M. Cargod, P. Dagenaise, et al. 2018. *Mixed Methods Appraisal Tool (MMAT) Version 2018: User Guide*. Montreal, Québec: McGill Department of Family Medicine. http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf.
- Kartal, G. 2024. "The Influence of chatgpt on Thinking Skills and Creativity of EFL Student Teachers: A Narrative Inquiry." *Journal of Education for Teaching* 50 (4): 627–642. <https://doi.org/10.1080/02607476.2024.2326502>.
- Laak, K.-J., and J. Aru. 2024. "Generative AI in k-12: Opportunities for Learning and Utility for Teachers." *Communications in Computer and Information Science*: 2150. https://doi.org/10.1007/978-3-031-64315-6_49.
- Moorhouse, B. L., and L. Kohnke. 2024. "The Effects of Generative AI on Initial Language Teacher Education: The Perceptions of Teacher Educators." *System* 122:103290. <https://doi.org/10.1016/j.system.2024.103290>.
- Nguyen Thi Thu, H. 2023. "EFL Teachers' Perspectives Toward the Use of chatgpt in Writing Classes: A Case Study at Van Lang University." *International Journal of Language Instruction* 2 (3): 1–47. <https://doi.org/10.54855/ijli.23231>.
- Prakasha, G. S., R. Sanskriti, and B. Ishani. 2024. "User Experiences of chatgpt Among Engineering Students, Teachers, and Working Professionals in India." *Journal of Educators Online* 21 (2): 1–17. <https://doi.org/10.9743/JEO.2024.21.2.12>.
- Tapan-BROUTIN, M. S. 2024. "Exploring Mathematics Teacher Candidates' Instrumentation Process of Generative Artificial Intelligence for Developing Lesson Plans." *Yükseköğretim Dergisi* 14 (1): 165–176. <https://doi.org/10.53478/yuksekogretim.1347061>.
- Ulla, M. B., W. F. Perales, and S. O. Busbus. 2023. "To Generate or Stop Generating Response?: Exploring EFL Teachers' Perspectives on chatgpt in English Language Teaching in Thailand." *Learning: Research and Practice* 9 (2): 168–182. <https://doi.org/10.1080/23735082.2023.2257252>.
- Uribe, S. E., I. Maldupa, A. Kavadella, M. El Tantawi, A. Chaurasia, M. Fontana, R. Marino, N. Innes, and F. Schwendicke. 2024. "Artificial Intelligence Chatbots and Large Language Models in Dental Education: Worldwide Survey of Educators." *European Journal of Dental Education* 28 (4): 865–876. <https://doi.org/10.1111/eje.13009>.

- Weiss, C. 1979. "The Many Meanings of Research Utilisation." *Public Administration Review* 39 (5): 426–431. <https://doi.org/10.2307/3109916>.
- Wrigley, T. 2018. "The Power of 'Evidence': Reliable Science or a Set of Blunt Tools?" *British Educational Research Journal* 44 (3): 359–376. <https://doi.org/10.1002/berj.3338>.