Educational Psychology Review (2025) 37:72
https://doi.org/10.1007/s10648-025-10050-2

META-ANALYSIS

Do Pedagogical Agents Enhance Student Motivation? Unraveling the Evidence Through Meta-Analysis

Jessica R. Gladstone[1] · Noah L. Schroeder[2] · Steffi Heidig[3] · Shan Zhang[2] · Chris Palaguachi[1] · Marcin Pitera[1]

Received: 20 September 2024 / Accepted: 27 June 2025 / Published online: 23 July 2025

Abstract

The use of pedagogical agents (PAs) as learning tools within digital learning environments is rising. Previous research shows that PAs can aid learning across various domains and age groups, but their impact on learner motivation is unclear. As PAs become more integrated into learning systems, building an in-depth and theory-driven understanding of how PAs may influence learners' motivation is necessary. We used four prominent motivation theories to guide our examination of the impact of PAs on learner motivation: social cognitive theory, situated expectancy-value theory, interest theory, and self-determination theory. A total of 58 articles met our inclusion criteria. We conducted seven, three-level meta-analyses that included 28 potentially moderating variables and used correlational and hierarchical effects and robust variance estimation. Our results revealed that PAs significantly influenced self-efficacy expectations and interest but did not significantly influence other theory-driven motivational beliefs. We discussed these findings from different theoretical perspectives and provided implications for practice.

Keywords Conversational agents · Virtual characters · Pedagogical agents · Motivation

✉ Jessica R. Gladstone
jglads2@illinois.edu

1 College of Education, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA

2 University of Florida, Gainesville, FL, USA

3 University of Applied Sciences Zittau/Goerlitz, Zittau, Germany

Introduction

Educational technology is rapidly evolving, and one area in particular that has seen increasing attention is the use of virtual characters to enhance the learning experience. Although these characters can play many roles in a learning environment (Clarebout et al., 2002; Schroeder & Gotch, 2015; Siegle et al., 2023), one type that has been widely researched is pedagogical agents (PAs). PAs are specifically designed to facilitate learning and may play various roles, such as acting as a role model and encouraging educational conversations with the learner, providing instruction or feedback, or playing a motivational role while scaffolding learning (Siegle et al., 2023). Previous research has indicated that PAs can aid learning in many situations and for various age groups (Castro-Alonso et al., 2021; Guo & Goh, 2015; Schroeder et al., 2013; Wang et al., 2023), but their impacts on learner motivation are less clear (Heidig & Clarebout, 2011; Schroeder & Adesope, 2014).

As PAs are incorporated into more systems that learners interact with daily, it is necessary to have an in-depth understanding of how PAs may influence learners' motivation. While previous reviews have shown mixed results (Heidig & Clarebout, 2011; Schroeder & Adesope, 2014) or some indication that PAs can aid learners' motivation (Guo & Goh, 2015; Wang et al., 2023), there is still a dire need to operationalize *what motivation is* in these various contexts. Although learning outcomes are often broken down into retention or transfer tasks, motivational constructs are less consistent in the PA literature. Motivation is an established area of research within educational psychology, with well-established theories that have been researched for decades (Wigfield & Koenka, 2020). However, researchers have often not considered the intricacies of motivational theories or measurement in the PA literature, especially the existing research synthesis efforts. Only one meta-analysis we are aware of analyzed a specific motivational construct. Wang et al. (2023) found that PAs can improve learners' intrinsic motivation. Outside of this one meta-analytical finding, we cannot say with any certainty that PAs have a specific impact on learners' confidence-related beliefs, value beliefs, or any other specific measure often associated with various motivational theories. Instead, existing syntheses have primarily focused on motivation as a meta-construct (Guo & Goh, 2015; Heidig & Clarebout, 2011; Schroeder & Adesope, 2014). In this paper, we argue that this approach is theoretically unstable. We seek to change this paradigm and more clearly articulate to what extent PAs may influence specific motivational constructs. We conduct a comprehensive systematic review of the literature across various research fields and use three-level meta-analytic methods with correlated and hierarchical effects and robust variance estimation to understand how the inclusion of a PA influences specific motivational constructs.

Motivation as a Construct in PA Research

Although it is assumed that PAs may improve learners' motivation, the emphasis of research on PAs has been more focused on learning outcomes than motivational measures. For example, a number of meta-analyses of PAs exist (Castro-Alonso et al., 2021; Davis, 2018; Davis et al., 2023; Schroeder et al., 2013), but they only analyzed the effect of PAs on learning, not motivation. Meanwhile, there is some synthesized evidence in regard to motivational measures, and this work shows an increase in researcher interest in more recent times. For example, Heidig and Clarebout's (2011) systematic review (covering the years 2002–2010) reported that only two of 15 studies that included a control group (without a PA) measured motivation as an outcome variable. In more recent reviews (Dai et al., 2022; Schroeder & Adesope, 2014) and a meta-analysis on affective PAs (Guo & Goh, 2015), about one third of the outcome measures examined motivational effects.

Very recently, however, Wang et al. (2023) conducted a meta-analysis investigating the efficacy of affective PAs. They reported 26 effect sizes for intrinsic motivation out of 36 articles. This trend towards increasing numbers of studies within systematically conducted reviews indicates a growing interest in motivation as an outcome measure in PA research. With the exception of Wang et al. (2023), the mentioned reviews and meta-analyses that examined motivation as an outcome did not differentiate between different motivational measures (Dai et al., 2022; Guo and Gho, 2015; Schroeder & Adesope, 2014; Wang et al., 2023). However, important differences may arise when specific motivation constructs are examined. For example, PAs may be particularly effective at boosting learners' confidence-related beliefs but have little impact on learners' interest. This finding would be essential for practitioners to know. The implications across various measures also have different implications for theory. Accordingly, it is necessary to dissect what exactly "motivation" means in the context of PAs.

Relevant Motivation Theories for the Implementation of Pedagogical Agents

Considering specific motivational constructs can provide valuable insights into understanding how and why PAs can facilitate learner motivation. As PA researchers, we can draw from established motivational theories to understand better how PAs can be used as a motivational tool and which specific motivation constructs PAs impact. Research around PAs is commonly grounded in claims that the social cues they can embody and convey are helpful within learning environments (e.g., Mayer et al.'s (Mayer, et al., 2003) social agency theory), and researchers in the field have outlined a number of roles PAs can facilitate within learning spaces (e.g., Baylor & Kim, 2005; Clarebout et al., 2002). As such, in this work, we draw from four prominent motivation theories that focus on the role of socializers (e.g., someone within the learning space) or the social context for understanding learners' motivation. Two of the theories, social cognitive theory and situated expectancy-value theory, have been used to guide other reviews on the use of role models (Gladstone & Cimpian, 2021; Morgenroth et al., 2015), which are conceptually similar to many implementations of PAs in the literature. However, interest theory (Hidi & Renninger, 2006; Renninger & Hidi, 2016, 2019) and self-determination theory (SDT; Ryan & Deci, 2000) are two other prominent motivation theories that emphasize the social context for learners' subsequent motivation and have been used in the PA literature to help guide how PAs can be used as motivational tools (e.g., Moreno et al., 2001; Plant et al., 2009; Zeitlhofer et al., 2023). We briefly summarize these theories and discuss the theoretical implications for the use of PAs next.

Social Cognitive Theory

Social cognitive theory (SCT) is particularly relevant to the use of PAs in digital learning environments. This theory highlights the significance of the social environment and observational learning in strengthening students' confidence-related beliefs, known as self-efficacy (Bandura, 1997; Schunk & DiBenedetto, 2020; Schunk & Usher, 2019). Self-efficacy refers to an individual's belief that they can succeed in a task or domain (Bandura, 1986, 1997), with Bandura (2001) emphasizing that observing others, such as peers or role models, succeed in similar tasks is a key source of self-efficacy. Previous research has found that learners who observe others successfully perform a task may believe they can replicate that success, thus boosting their self-efficacy expectations (Schunk & DiBenedetto, 2020; Schunk & Usher, 2019). In this way, pedagogical agents could play a comparable role to peers or role models, boosting learners' self-efficacy through observational learning.

Situated Expectancy-Value Theory

Situated expectancy-value theory (SEVT; Eccles & Wigfield, 2020) explains that students' situative motivation to engage in achievement tasks depends on their beliefs about their performance or success on a task ("expectancy") and the subjective value and relevance they place on those tasks ("values"). Importantly, these beliefs and subsequent engagement are influenced by social models (Eccles & Wigfield, 2020). Although SEVT provides a useful framework for understanding motivational processes, its constructs align conceptually with other motivational constructs and theories, such as self-efficacy within SCT (Pekrun, 2024). In the context of PA research, PAs may act as social influences that affect students' expectancies of success by modeling competence or providing feedback that bolsters confidence. This connection is reflected in the PA literature, where self-efficacy and expectancies of success are frequently measured as key outcomes (e.g., Baylor & Kim, 2005; Kim et al., 2007; Plant et al., 2009; van der Meij et al., 2015). Accordingly, the present study explores whether PAs can enhance students' self-efficacy and/or expectancies of success, which we categorize collectively as self-efficacy expectations. Learners' value beliefs, or task relevance, may also be impacted by PAs, particularly through the demonstration of reasons for engaging in a task (van der Meij, 2013). Among the components of value beliefs, utility value, defined as the perceived usefulness of a task (Eccles & Wigfield, 2020; Harackiewicz et al., 2014), may be particularly relevant to PA research. Utility value is considered one of the most malleable value beliefs, making it a prime target for experimental and interventions studies (Harackiewicz et al., 2014). For instance, Plant et al. (2009) demonstrated how perceived similarity to a computer-based agent could influence students' utility value perceptions regarding a career path. Thus, PAs designed to embody relatable social model characteristics may help students perceive greater utility in engaging with academic tasks. Therefore, in the present study, we focus on the broader concept of subjective task values (i.e., task relevance as well as the specific construct of utility value beliefs.

Learners' engagement, defined as learners' active and committed participation (Fredricks et al., 2004), is another construct considered within the SEVT model (Gladstone et al., 2022). Unsurprisingly, Heidig and Clarebout (2011) also found evidence that researchers were measuring engagement in the PA literature (i.e., Baylor & Kim, 2005). Although motivation can be considered the driving force of engagement (Eccles & Wang, 2012), these constructs both have important implications for learners' outcomes and can be influenced by socializers (e.g., role models; Eccles & Wang, 2012; Gladstone et al., 2022; Sinatra et al., 2015). Therefore, a PA may also influence learners' engagement in the task.

Interest Theory

Interest theory (Hidi & Renninger, 2006; Renninger & Hidi, 2016, 2019) may also be relevant to the use of PAs in digital learning environments because often, the goal of introducing a student to a PA is to boost interest in a particular topic or field of study (Kim et al., 2007; Moreno et al., (Moreno et al., 2001); Plant et al., 2009). Interest as a motivation construct tends to have two conceptualizations. The first refers to the psychological state as learners engage with a specific topic or material (e.g., an algebra lesson), and the second refers to learners' motivation to work with that topic or material over time (Renninger & Hidi, 2019). According to interest theory, the development of interest can result from the interaction between a student and their social environment (Hidi et al., 2004; Renninger, 2000). In their systematic review of role models, Gladstone and Cimpian (2021) found that researchers often measured learners' interest as an important outcome variable after exposure to a role model in a social environment (e.g., Plant et al., 2009). Previous work on PAs has also examined whether PAs may be an influential affective factor for promoting interest via social cues (Domagk, 2010). Therefore, interest will also be an important outcome to consider when using PAs to boost motivation.

Self-Determination Theory

Given that SCT and SEVT both emphasize that PAs may be used to boost learners' confidence-related beliefs, self-determination theory (SDT) offers an additional framework for understanding how PAs may enhance learner motivation. A central premise of SDT is that social environments should support learners' competence in order to enhance their intrinsic motivation (Ryan & Deci, 2000). In SDT, competence is considered a basic psychological need that is essential for fostering motivation. When learners feel competent, they are more likely to experience confidence in their abilities, which can be intrinsically rewarding and reinforce their internal drive to engage in activities and tasks that they find inherently interesting (Ryan & Deci, 2000). By providing opportunities for success, social environments can support learners' competence, thereby strengthening their intrinsic motivation.

According to SDT, PAs could enhance competence by doing things such as modeling effective problem-solving strategies, providing personalized feedback, or offering encouragement that helps learners overcome challenges. These experiences would thus not only reinforce learners' confidence in their abilities but also satisfy their basic psychological need for competence, which according to SDT is critical for developing intrinsic motivation. Although Wang and colleagues (Wang et al., 2023) did not use SDT as a guiding framework in their meta-analysis, their findings suggest that PAs can indeed increase learners' intrinsic motivation, potentially through their impact on competence. Moreover, intrinsic motivation, as described in SDT, shares similarities with interest, as both involve an internal drive to engage in activities that are perceived as enjoyable or meaningful (Renninger, 2000). Therefore, SDT provides a valuable lens for understanding how PAs can be implemented to foster competence and intrinsic motivation in learning contexts (Zeitlhofer et al., 2023).

Summary of Motivation in PA Research

As noted, with one exception (Wang et al., 2023), previous reviews of the motivational effects of PAs have, often by necessity due to limited data, been broadly scoped, lumping all measures of learners' motivation into one outcome construct for discussion and analysis (Guo & Goh, 2015; Heidig & Clarebout, 2011; Schroeder & Adesope, 2014). Our premise is that this approach needs to be theoretically justified. Thus, we should use prominent motivation theories to guide our work in understanding how PAs may impact learners' specific motivation. Based on the four theories presented above, we consider six discrete types of learner motivation that may be important to consider when implementing PAs as a motivational tool: self-efficacy expectancies (SCT, SEVT), subjective task values or task relevance (SEVT), utility value (SEVT), engagement (SEVT), interest (interest theory), and intrinsic motivation (SDT). Thus, in the present review, we use these theories of learner motivation to guide our examination of the impact of PAs on specific motivation constructs. However, since we found that motivational measures were often not described in sufficient detail to be classified into one of the six categories mentioned above, we added a broad "motivation" category to include studies with unspecified motivation measures in our analyses.

Factors That May Moderate the Effects of PAs on Learners' Motivation

As Heidig and Clarebout (2011) noted, assuming a PA can broadly influence motivation or learning misses important nuances in their implementation that may impact their effectiveness. These nuances have also been noticed in other reviews (Schroeder & Adesope, 2014). Consequently, it is important to understand and examine what factors may moderate the effects of PAs on learners' various motivation outcomes. For this paper, we have categorized these variables as study design factors, system design factors, and character design considerations. Each will be discussed briefly below; however, due to space constraints, we do not elaborate on each individual factor. In short, while the impacts of some of these factors on learning outcomes have been examined in previous PA meta-analyses (Castro-Alonso et al., 2021; Davis et al., 2023; Guo & Goh, 2015; Schroeder et al., 2013), evidence of how they influence motivation outcomes is lacking. However, these factors are nonetheless important for us to disentangle and understand in relation to PAs impact on motivation outcomes and also hold implications for how to design effective PAs. These factors are derived from previous meta-analyses, Heidig and Clarebout's (2011) frameworks for PA design and implementation, and relevant PA research.

Study Design Factors

There are various ways in which the way we design and implement our research studies can influence the effects of the intervention, in this case, the PA. For example, the control condition it is being compared to can influence the overall effectiveness of an intervention, as can the way we assign participants to a condition, the research design, the number of interactions with the system, the duration of the intervention, the domain of the learning materials, the study setting, and the grade level of the learners. All of these are important contextual factors for us to understand. For example, if PAs are significantly better at increasing learners' self-efficacy expectations when compared to instructional videos but are less effective than intelligent tutoring systems, that would have implications for future PA implementation.

System Design Factors

How the PA is implemented within a learning system may also influence its effectiveness in providing learners with motivational benefits (Heidig & Clarebout, 2011). For example, the pacing of the learning system, the type of media the intervention was delivered on, whether the task is collaborative or individual, and the number of PAs in the learning system can all potentially influence the effectiveness of PAs for facilitating learners' motivation, and they provide important insights into the design of educational technologies. For example, if engagement is higher from learning with PAs during collaborative tasks than individual tasks, that provides concrete implications for future PA design.

Character Design Considerations

Although it is important to consider how we design studies and systems, the PA designer must also consider the character's design. This decision space is ample. However, researchers can leverage Heidig and Clarebout's (2011) Pedagogical Agents—Levels of Design (PALD) framework to better understand the specific design considerations. Specifically, in this study, we consider a variety of social cues connected to PA design that may influence how the PA is perceived and interacted with as a social partner or role model, such as whether or not it uses facial expressions, gestures, and gaze. In addition, we examined the PA's physical form and appearance, such as its age and role in the system. By investigating these various factors, we can better understand how PAs can be designed to facilitate specific motivational constructs.

The Present Study

Although a few existing meta-analyses examine PA's influence on learners' motivation, they have some limitations. First, as noted, only one examined a specific motivational construct (Wang et al., 2023) rather than "motivation" more broadly. Second, the studies used conventional meta-analytic techniques. Since conventional meta-analysis relies on each comparison being fully independent (i.e., each participant can only be counted once), researchers analyze only a portion of the data that could otherwise be available. There are now three-level approaches to meta-analysis that allow one to nest comparisons within studies (Assink & Wibbelink, 2016).

Furthermore, we have correlated and hierarchical effects (CHE) models and robust variance estimation (RVE) that allow us to get more precise estimates of effects by making assumptions about the correlations between comparisons within studies and adjusting confidence intervals (Harrer et al., 2021; Pustejovsky & Tipton, 2022). In this study, we combined these approaches by using a three-level meta-analytic (3LMA) model that accounted for CHE and used RVE. This approach allowed us to use more data from each study while providing robust estimates than previous research syntheses have. Not only did we analyze more data compared to previous studies, but we also analyzed motivational constructs specifically, rather than lumping them together into one "motivation" outcome, to provide more precise insights into the effects of PAs.

To summarize, we used 3LMA with CHERVE to explore the following research questions:

RQ1: What is the impact of PAs on learners' "motivation," and what moderates this effect?
RQ2: What is the impact of PAs on learners' self-efficacy expectations, and what moderates this effect?
RQ3: What is the impact of PAs on learners' value beliefs, and what moderates this effect?
RQ4: What is the impact of PAs on learners' utility beliefs, and what moderates this effect?
RQ5: What is the impact of PAs on learners' engagement, and what moderates this effect?
RQ6: What is the impact of PAs on learners' interest, and what moderates this effect?
RQ7: What is the impact of PAs on learners' intrinsic motivation, and what moderates this effect?

Methods

Transparency and Openness

We conducted a systematic review and report it following the PRISMA 2020 guidelines (Page et al., 2021). The review was not pre-registered; however, all data and supporting information are available via the Open Science Framework at the following link: https://osf.io/u8wtg/?view_only=f2d36cc4ed6048a2a4d4243df05eb27b. All analyses were run using Simple Meta-Analysis (Schroeder, 2024).

Literature Search

On September 15, 2023, we conducted a literature search in the following eight databases, which include sources from a variety of professional fields: Medline with Full Text, APA PsycInfo, Academic Search Complete, Education Research Complete, Eric, CINAHL Plus with Full Text, ACM Digital Library. No filters were based on article type or publication year during the search.

Our search string was designed to capture studies investigating the impacts of PAs on students' motivation and consisted of the following: ("virtual human"* OR "embodied agent"* OR "virtual character"* OR "pedagogical agent"* OR "conversational agent"* OR "motivational agent"*) AND (motivat* OR self-efficacy OR self-confidence OR ability belief* OR self-concept OR interest* OR engag* OR value* OR util* OR "sense of belonging" OR belong*). Our search string included a series of constructs that have been used to measure some aspect of students' motivation (Dai et al., 2022) and were driven by SCT, SEVT, interest theory, and SDT. In addition, we added studies that met our criteria that were cited in existing papers around the design of motivational agents (Baylor, 2011), existing systematic reviews (Heidig & Clarebout, 2011; Schroeder & Adesope, 2014), and empirical studies that measured motivational outcomes in existing meta-analyses (Guo & Goh, 2015; Wang et al., 2023).

Our initial search resulted in 2478 studies. After removing duplicates and studies not written in English, 1650 studies remained for the abstract screening. Figure 1 shows the database search results and additional articles from references and reviews.

Inclusion and Exclusion Criteria

To be included in this meta-analysis, studies must have met the following criteria:

1. Must have a visible on screen agent, which must be a virtual character and not a video or image of an actual human.
2. Must compare a PA to a non-PA condition.
3. Must measure some aspect of students' motivation (e.g., motivation, self-efficacy, confidence beliefs, ability beliefs, self-concept, interest, engagement, value, utility, etc.).
4. Must be an empirical study with primary data, with enough information provided to compute an effect size.
5. Must be published in English..
6. Must be publicly accessible via databases or inter-library loan requests.
7. Must be publicly accessible via databases or inter-library loan requests.

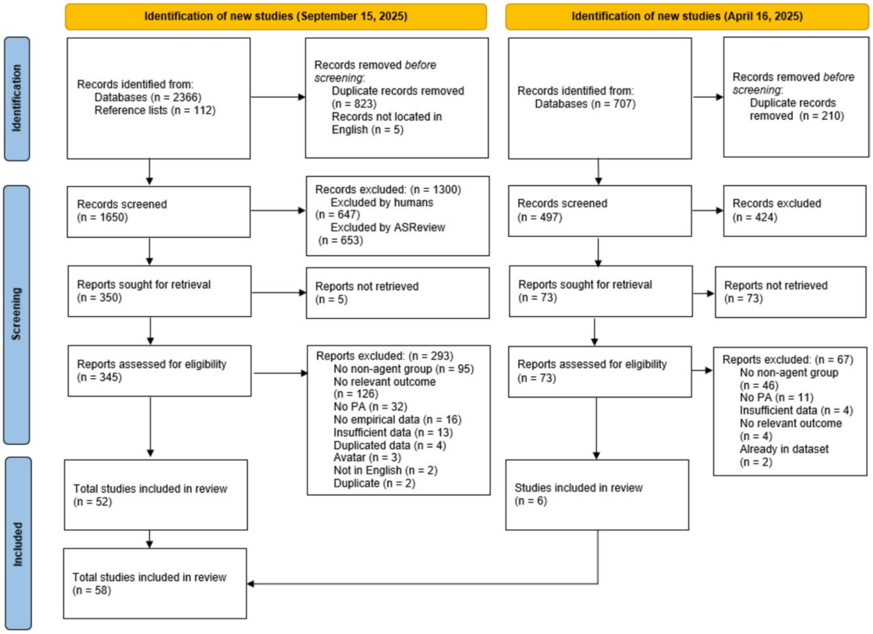

Fig. 1 PRISMA diagram, adapted from Page et al. (2021)

Studies were excluded based on the following criteria:

1. The study examined the use of avatars (representations of self).

Study Screening

The study screening had three phases, as shown in the PRISMA flowchart (Fig. 1).

Phase I Screening

We imported all studies to ASReview (https://asreview.nl/), a screening tool that uses natural language processing to prioritize studies from most to least relevant. We trained the system using five relevant (Baylor & Plant, 2005; Domagk & Niegemann, (2005); Plant et al., 2009; Rosenberg-Kima et al., 2007; Rosenberg-Kima et al., 2010) and five irrelevant studies (Gottsacker et al., 2022; Kolodkin et al., 2012; Koprinkova-Hristova et al., 2013; Sukthankar & Sycara, 2011; Wang et al., 2021). We employed sentence BERT for feature extraction, logistic regression served as the classification method, "maximum" was utilized as the query approach, and dynamic resampling (doubling) was applied as the strategy for maintaining balance.

We used ASReview to help with title and abstract screening based on the results of a recent paper (Campos et al., 2024). Specifically, they found that with a sample of reviews in educational psychology, screening 60% of the abstracts could identify 95% of the included studies when using logistic regression and sentence BERT. As such, we screened 60% of our studies in our database. To confirm if 60% of study screening was sufficient, we investigated the longest sequence of consecutively irrelevant studies and found 126 consecutive irrelevant studies, accounting for 7.58% of our sample. According to Campos et al. (2024), a recall rate of over 95% was associated with data where 7% were coded as consecutively irrelevant. Consequently, we concluded that screening 60% of our database would likely allow us to identify 95% of studies that met our criteria. In the abstract screening phase, 1650 studies were imported to ASReview, 647 studies were excluded by one author, and 653 studies were excluded by ASReview, resulting in 345 studies that were located and included for full-text screening.

Phase II Screening

The second phase was full-text screening. The same author who conducted the abstract screening independently completed the full-text screening ($n=345$). The process resulted in 52 studies that met all of our inclusion criteria.

Updating the Literature Search

On April 16, 2025, we updated our literature search using the same search terms, but constraining the search from 2023 to 2025. We searched most of the same databases, although some were slightly different due to institutional changes and access (see data in OSF repository). After removing duplicates, we screened 497 newly identified studies using the same Phase I screening process as described above, except we reviewed every abstract instead of screening 60%. Additionally, we had an independent screener, someone not an author on the manuscript, screen the first 100 abstracts. The independent screener and first author had 98% agreement with the included and excluded articles. We identified 73 studies for full-text screening. A second reviewer independently examined 15 studies at this phase (~20% of the sample) and inter-rater agreement was 100%. In sum, updating the search added six new studies to the analyses.

Data Extraction

Data extraction occurred in three phases, as outlined below.

Selecting Relevant Comparisons

3LMA offers an advantage compared to conventional meta-analysis in that it can account for dependencies within the data, allowing for multiple effect sizes to be extracted from each set of participants. However, as meta-analysts, we feel it is critically important to reduce confounding variables where possible to facilitate the interpretation of the results. Accordingly, we emphasized selecting comparisons that reduced, or preferably eliminated, confounding variables between the conditions being compared. For example, Park (2005) explored using various PAs and seductive graphics, resulting in eight experimental conditions. We specifically compared groups with similar seductive graphics conditions to isolate the variable of interest, in this case, the presence or absence of a PA, as the only difference between conditions.

Another difference worth noting between conventional and 3LMA models regarding their structural differences in the data is in studies with only one control group but multiple experimental groups. In these cases, the multiple experimental groups were all compared to the one control group on each variable. For example, Jung et al. (2022) explored the impacts of five PAs and one control condition (i.e., no visual PA) on various outcomes. As such, we extracted five comparisons for each outcome variable for this study: each experimental group compared to the same control group.

Variables Extracted

Researchers have explored many aspects of PA design that could influence learners' motivation. Moreover, previous meta-analyses and reviews in the field have explored various study designs and methodological factors that could influence the results. Consequently, we coded several variables in order to build a comprehensive understanding of how PA implementation and design may influence learners' motivation. We categorized the variables similarly to Schroeder et al. (2025) in that we categorized variables as study descriptives (year, publication type), study design factors (control condition, randomization strategy, research design, duration of intervention, number of interactions, domain, setting, grade level, location, pre-test presence), system design factors (task type, pacing, media type, number of agents), character design considerations (voice type, gaze, facial expressions, gestures, form, age, agent role), and effect size data. We also examined study quality (same learning materials, attrition rate, assignment strategy, outcome measure understandability, outcome measure reliability, overall quality indicators). The categorization scheme for all variables is described in detail in Appendix A on OSF: https://osf.io/u8wtg/?view_only=f2d36cc4ed6048a2a4d4243df05eb27b.

Inter-rater Agreement

Two authors independently coded 11 studies (18.96% of the sample). Inter-rater agreement was 94.62%. Disagreements were reconciled through discussion and re-examination of the primary studies.

Data Analysis

We used simple meta-analysis (Schroeder, 2024) to run a 3LMA with CHERVE. We assumed a correlation of $r = 0.60$ within studies. We made inferences using restricted maximum likelihood estimation and used t distributions (Viechtbauer, 2022). Effect sizes were interpreted using Hattie's (2015) criteria. Specifically, in educational settings, Hattie specified $g=0.20$ as a small effect, $g=0.40$ as a moderate effect, and $g=0.60$ as a large effect. Since we used CHERVE, during moderator analyses, we removed levels of each moderator that only contained one comparison, as the presence of only a single comparison within a level can prevent the analysis from computing a *p*-value. When this occurred, we noted the level removed from each analysis in the relevant table.

Outliers and Influential Cases

To identify outlying cases, simple meta-analysis uses Van Lissa's (n.d.) method of looking for comparisons with non-overlapping confidence intervals with the overall meta-analytic effect size. Simple meta-analysis also checked if these, or any other, cases were influential using Cook's distance, DFBETAS,[1] and hatvalues (Viechtbauer, n.d.). Simple meta-analysis uses the following metrics as indicators of significant influence: Cook's distance over 0.50 (Viechtbauer & Cheung, 2010), DFBETAS over 1 (Viechtbauer & Cheung, 2010), and hatvalues over $3 \times$(*number of model coefficients* ÷ *number of studies*) (Viechtbauer, n.d.b).

Forest Plots

We used simple meta-analysis to create comparison-level and study-level forest plots.

Publication Bias

Conventional approaches to quantifying publication bias do not function well in 3LMA models (Rodgers & Pustejovsky, 2021). As such, we used simple meta-analysis to create a comparison-level funnel plot to examine for asymmetry.

Results

Our sample included 58 studies that included 304 comparisons. We ran seven 3LMAs using CHERVE (Table 1) and examined the impact of 28 potentially moderating variables across five categories. In sum, we report the results of 84 moderator analyses, as some meta-analyses did not have sufficient studies to warrant moderator analysis. Due to space constraints and the considerable number of analyses run, only significant findings are reported here. A complete reporting of each set of analyses, including discussions of outlier and influence decisions, forest plots, funnel plots,

[1] We note that DFBETAS was based on the CHE model (as opposed to the CHERVE model) as the function does not work with robust objects in R. The same is the case for evaluating the variance of the meta-analytic models using I^2.

Table 1 The overall results of each 3LMA with CHERVE

Outcome	$k_{comparisons}$	$k_{studies}$	g	p	Q	τ^2_{within}	$\tau^2_{between}$	Quality indicator median	Quality indicator range
"Motivation"	84	22	**0.20**	0.03	294.90*	0.05	0.08	4	1–5
Self-efficacy expectations	57	29	**0.29**	.001	475.51*	0.29	0.00	4	1–5
Value beliefs	25	9	0.09	0.40	58.91*	0.04	0.02	4	4–5
Utility beliefs	8	5	0.34	0.06	36.49*	0.13	0.00	4	4–5
Engagement	57	8	−0.02	0.73	217.68*	0.09	0.00	3	2–5
Interest	50	12	**0.30**	0.01	130.76*	0.08	0.01	4	3–5
Intrinsic motivation	23	3	0.00	0.99	56.97*	0.05	0.01	3	3

Bolded numbers indicate significant findings. Quality indicator scores range from 0 to 5. * indicates $p < .001$

and the raw data, is available in the project OSF repository (https://osf.io/u8wtg/?view_only=f2d36cc4ed6048a2a4d4243df05eb27b).

Figure 2 shows the trend of publications over time. Specifically, it shows the number of comparisons in the analysis by year of study publication. Researchers have been interested in the use of PA compared to non-PA conditions since the early 2000 s; however, there seems to be a surge of comparisons roughly every 5 years. Notably, we found 72 comparisons in 2021, showing that researchers are still quite interested in the impacts of PA on learners' motivation.

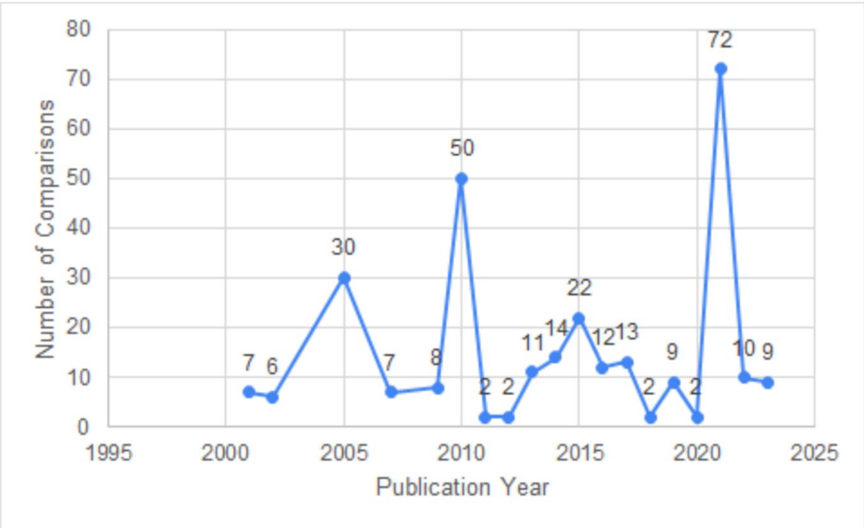

Fig. 2 Number of comparisons by publication year

Of the 304 comparisons, 172 were extracted from journal articles, 83 were from conference proceedings, and 49 were from dissertations. As shown in Fig. 3, most studies that reported where they took place were in the USA ($k_{\text{studies}} = 19$), Netherlands ($k_{\text{studies}} = 4$), and China ($k_{\text{studies}} = 7$).

In regard to study quality indicators, most comparisons had a reasonable number of quality indicators. However, as expected, there was a range of scores present. The maximum possible number was 5 quality indicators, with a minimum score of 0. The results for each analysis are in Table 1.

RQ1: What Is the Impact of PAs on Learners' "Motivation" and What Moderates This Effect?

Our first research question addresses those outcome measures which were classified as "motivation" measures but could not be directly tied to a relevant theoretical framework. The random effects 3LMA with CHERVE examining the impact of PAs on learners' "motivation" showed that PAs have a small but statistically significant effect ($g = 0.20$, $p = 0.03$, $k_{\text{comparisons}} = 84$, $k_{\text{studies}} = 22$). There was significant heterogeneity within the sample, $Q(83) = 294.90$, $p < 0.0001$, $\tau^2_{\text{within}} = 0.05$, $\tau^2_{\text{between}} = 0.08$. In addition, the model[1] explains 73.36% of the variance ($I^2_{\text{total}} = 73.36\%$), with some attributable within studies ($I^2_{\text{within}} = 28.11\%$) and more attributable to between studies ($I^2_{\text{between}} = 45.25\%$). Sensitivity analyses indicated that there were not noteworthy differences between setting ρ = 0.60 versus ρ = 0.40 or ρ = 0.80. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 4.

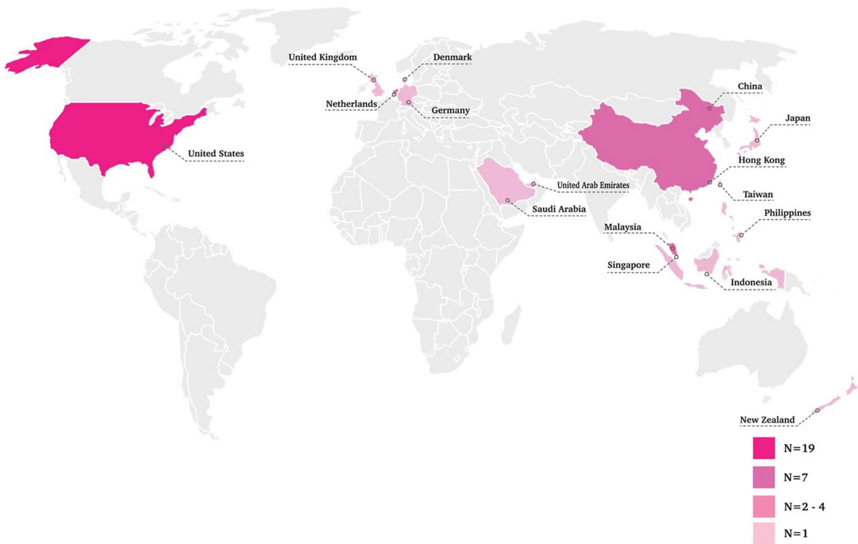

Fig. 3 A figure depicting the number of studies distributed globally that met the inclusion criteria. Note that studies which did not explicitly report where data were collected are not represented in this figure

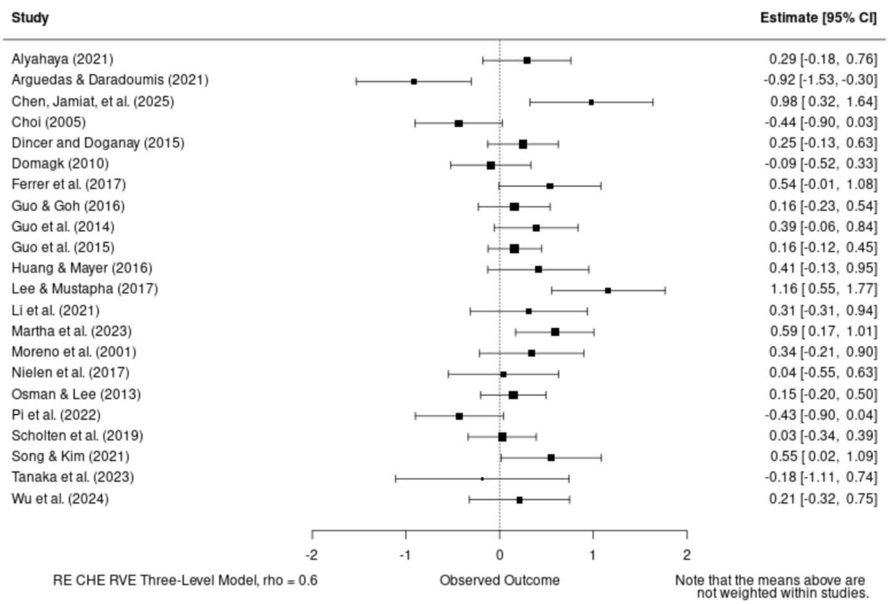

Fig. 4 Study-level forest plot of “motivation” outcomes

Moderator Analyses

We examined the effects of 28 potential moderating variables across five categories. A table in the OSF repository presents the results of the moderator analyses, including the omnibus test of moderators for each analysis. Only significant results are presented here.

We found that whether or not PAs used facial expressions significantly moderated the impact of PAs on learners’ “motivation” ($F(2.00, 4.52)=32.06$, $p<0.01$). Specifically, PAs with facial expressions improved learners’ “motivation” ($g=0.33$, $p=0.03$), whereas those without facial expressions had no significant effect.

The role of the PA also significantly moderated the impact of PAs on learners’ “motivation” ($F(4.00, 2.27)=19.164$, $p=0.04$). PAs that played multiple roles had a small positive effect ($g=0.97$, $p=0.04$) on learners’ motivation, but this data represents findings from only two comparisons within one study. As such, it should be interpreted with caution, especially given the size of the effect size being notably larger than the overall effect of PAs on learners’ “motivation.”

Publication Bias

Examination of the funnel plot indicates that effect sizes are reasonably symmetrically distributed. As such, publication bias was not viewed as a significant issue with this analysis.

RQ2: What Is the Impact of PAs on Learners' Self-Efficacy Expectations and What Moderates This Effect?

A random effects 3LMA with CHERVE found that the overall impact of PAs on learners' self-efficacy expectations was $g=0.29$, $p<0.001$ ($k_{\text{comparisons}}=57$, $k_{\text{studies}}=29$). There was significant heterogeneity within the sample, Q(56) = 475.51, $p<0.001$, $\tau^2_{\text{within}}=0.29$, $\tau^2_{\text{between}}=0.00$. Overall, the model[1] explains 99.17% of the variance (I^2), with 99.17% within studies and 0.00% between studies. Sensitivity analyses indicated that there were not noteworthy differences between setting $\rho=0.60$ versus $\rho=0.40$ or $\rho=0.80$. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 5.

Moderator Analyses

We examined the effects of 28 potential moderating variables across five categories. A table with the full moderator analysis details is available on OSF. Only significant results are presented here.

We found that the grade level of the participants significantly moderated the effects of PAs on learners' self-efficacy expectations (F(4.00, 6.83) = 11.42, $p<0.01$). Specifically, the strongest effect was found for students in grades 7–9 ($g=0.77$, $p<0.001$); however, there were only two comparisons extracted from one study (Plant et al., 2009). We also found a moderate, significant impact on adult learners outside of post-secondary education ($g=0.44$, $p<0.01$). While there were eight studies in this category containing 17 comparisons, we note that one of these studies had a very large sample size, which could have influenced the overall effect size found. Primary, upper secondary, and post-secondary learners did not find significant benefits or detriments for their self-efficacy expectations when compared to non-PA groups.

Our results also showed that the presence or absence of facial expressions significantly impacted learners' self-efficacy expectations compared to non-PA systems (F(2.00, 9.41) = 6.46, $p=0.02$). Specifically, PAs that used facial expressions had a moderate impact on learners' self-efficacy expectations ($g=0.46$, $p<0.01$), while those without facial expressions had no significant impact compared to non-PA systems. Noteworthy is that this effect size is from 12 studies that included 24 comparisons.

Finally, whether or not the reliability was reported for the outcome measure significantly moderated the effects of PAs on learners' self-efficacy expectations (F(1.00, 20.55) = 6.94, $p=0.02$). Studies that reported reliability of the measure with their sample were associated with a higher effect size ($g=0.44$, $p<0.001$) than those that did not ($g=0.07$, $p=0.56$).

Publication Bias

Examination of the comparison-level funnel plot shows that the effect sizes are reasonably symmetrically distributed. As such, publication bias is not viewed as a significant concern in this analysis.

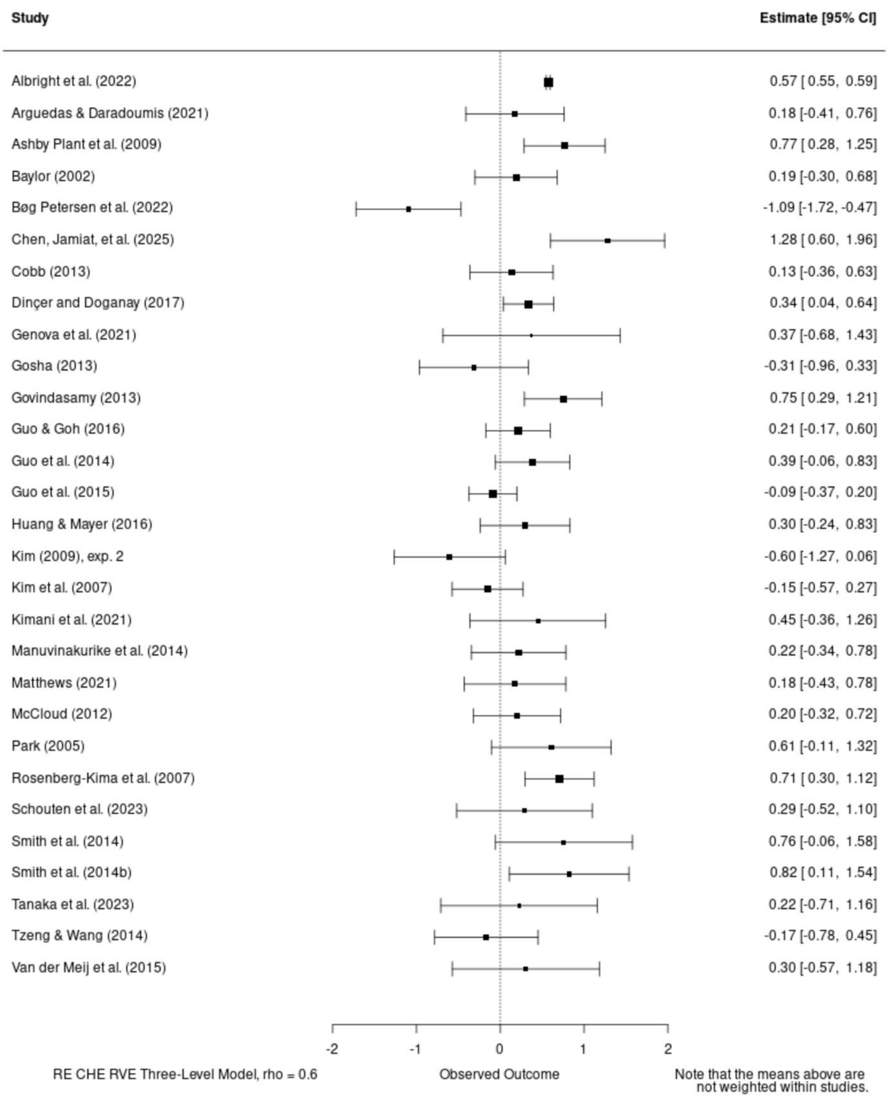

Fig. 5 Study-level forest plot of self-efficacy expectations

RQ3: What Is the Impact of PAs on Learners' Value Beliefs and What Moderates This Effect?

The random effects 3LMA with CHERVE exploring the impact of PAs on learners' value beliefs showed that PAs have a small, non-significant effect ($g=0.09$, $p=0.40$, $k_{\text{comparisons}}=25$, $k_{\text{studies}}=9$). There was significant heterogeneity within the sample, $Q(24)=58.91$, $p<0.001$, $\tau^2_{\text{within}}=0.04$, $\tau^2_{\text{between}}=0.02$. In addition, the model[1] explained 53.51% of the variance ($I^2_{\text{total}}=53.51\%$), with variation due to within studies ($I^2_{\text{within}} = 36.76\%$) and between studies ($I^2_{\text{between}} = 16.75\%$). Sensitivity analyses indicated that there were not noteworthy differences between setting $\rho = 0.60$ versus $\rho = 0.40$ or $\rho = 0.80$. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 6.

Moderator Analyses

Due to the relatively small number of studies contributing effect sizes ($k_{\text{studies}} = 9$), we did not conduct moderator analyses.

Publication Bias

Examination of the funnel plot indicates that effect sizes are reasonably symmetrically distributed. However, we note that the results of this analysis are based on only nine studies.

RQ4: What Is the Impact of PAs on Learners' Utility Beliefs and What Moderates This Effect?

Our random effects 3LMA with CHERVE of the impact of PAs on learners' utility beliefs found that PAs have a small to moderate, non-significant effect ($g = 0.34$, $p = 0.05$, $k_{\text{comparisons}} = 8$, $k_{\text{studies}} = 5$). There was significant heterogeneity within the sample, $Q(7) = 36.49$, $p < 0.001$, $\tau^2_{\text{within}} = 0.13$, $\tau^2_{\text{between}} = 0.00$. In addition, the model[1] explained 71.13% of the variance in the sample ($I^2_{\text{total}} = 71.13\%$), with

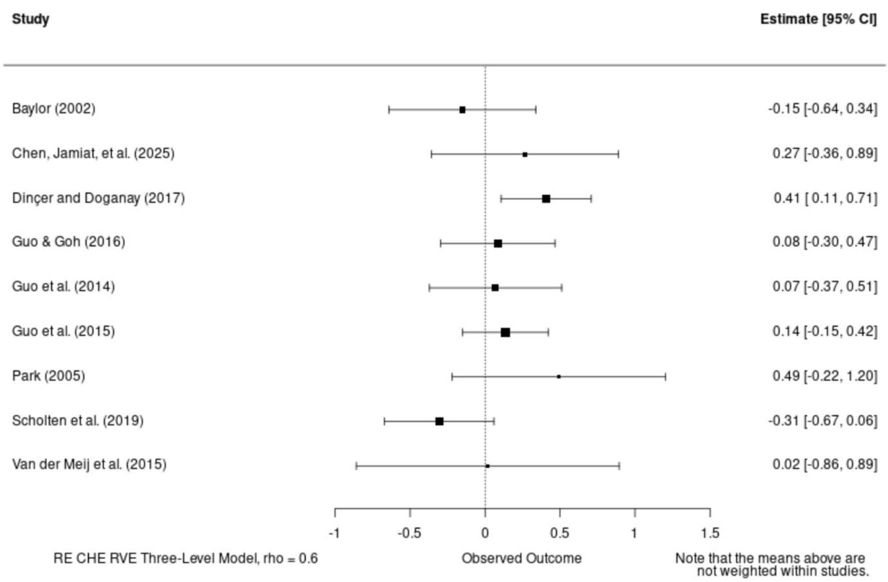

Fig. 6 Study-level forest plot of value belief outcomes

71.13% attributable to within study variance ($I^2_{\text{within}} = 71.13\%$), and none attributable to between study variation ($I^2_{\text{between}} = 0.00\%$). No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 7.

Moderator Analyses

Due to the small number of studies and comparisons in this analysis, we did not run moderator analyses.

Publication Bias

Examination of the funnel plot indicates that effect sizes are reasonably symmetrically distributed. However, we note that the results of this analysis are based on only eight comparisons from five studies.

RQ5: What Is the Impact of PAs on Learners' Engagement and What Moderates This Effect?

The result of a random effects 3LMA with CHERVE meta-analysis found that the impact of PAs on learners' engagement was small and non-significant ($g = -0.02$, $p = 0.73$, $k_{\text{comparisons}} = 57$, $k_{\text{studies}} = 8$). There was significant heterogeneity within the sample ($Q(56) = 217.68$, $p < 0.001$, $\tau^2_{\text{within}} = 0.09$, $\tau^2_{\text{between}} = 0.00$), and $I^2_{\text{total}} = 56.24\%$, indicating that the model[1] explained 56.24% of the variance, with

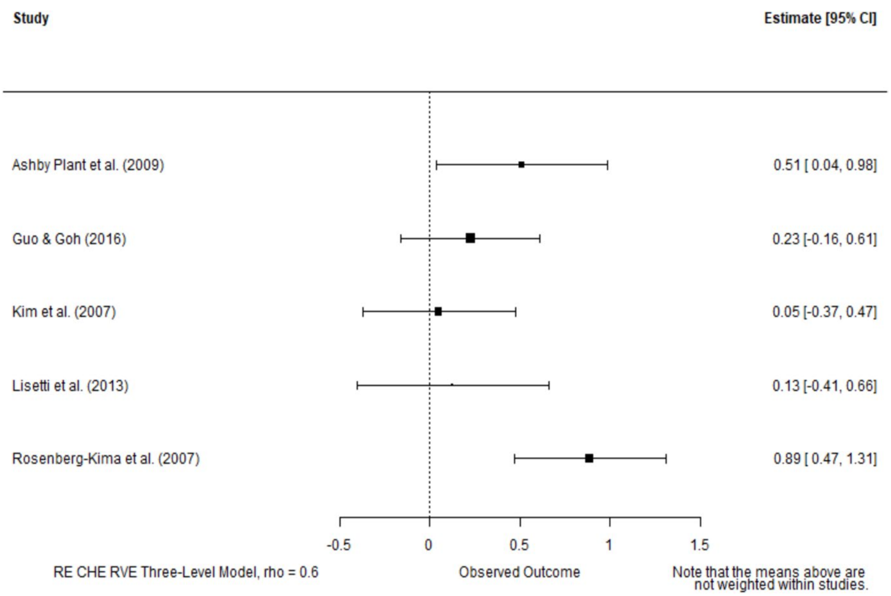

Fig. 7 Study-level forest plot of utility value belief outcomes

within study variation (I^2_{within} = 56.24%) explaining the variance rather than between study variation ($I^2_{between}$ = 0.00%). Sensitivity analyses indicated that there were not notable differences between setting ρ = 0.60 versus ρ = 0.40 or ρ = 0.80. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 8.

Moderator Analyses

Due to the small number of studies that contributed effect sizes to this analysis ($k_{studies}$ = 8) we did not conduct moderator analyses.

Publication Bias

Examination of the comparison-level funnel plot showed a reasonably symmetrical distribution of effect sizes. However, we note one study (Jung et al., 2022) accounted for 40 of the 57 effects in this analysis, so the results may not be as generalizable as the number of comparisons in the analysis may imply.

RQ6: What Is the Impact of PAs on Learners' Interest and What Moderates This Effect?

A random effects 3LMA with CHERVE exploring the impact of PAs on learners' interest found that PAs have a small to moderate, statistically significant effect (g = 0.30, p = 0.01, $k_{comparisons}$ = 50, $k_{studies}$ = 12). We found significant heterogeneity

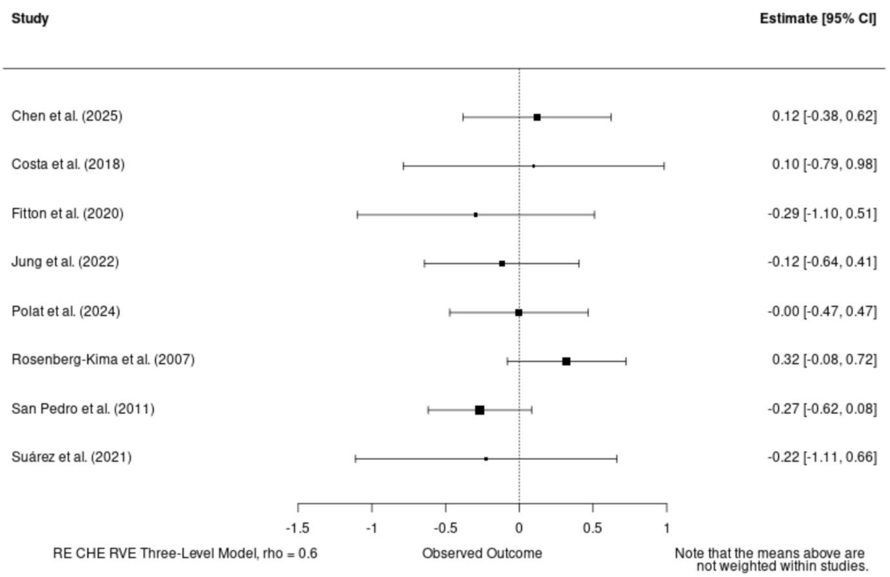

Fig. 8 Study-level forest plot of engagement outcomes

within the sample, $Q(49)=130.76$, $p<0.001$, $\tau^2_{\text{within}}=0.08$, $\tau^2_{\text{between}}=0.01$. The model[1] explained 50.87% of the variance ($I^2_{\text{total}}=50.87\%$), with some variance attributed to within studies ($I^2_{\text{within}}=43.03\%$) and additional variance attributed to between studies ($I^2_{\text{between}}=7.84\%$). Sensitivity analyses indicated that there were not notable differences between setting $\rho=0.60$ versus $\rho=0.40$ or $\rho=0.80$. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 9.

Moderator Analyses

We examined the effects of 28 potential moderating variables across five categories. A table with the full moderator analysis details is available on OSF. Only significant results are presented here.

The type of control condition significantly impacted the influence of PAs on students' interest ($F(3.00, 3.06)=400.67$, $p<0.001$). Specifically, PAs were significantly more effective than video-based interventions ($g=0.54$, $p<0.01$) and no intervention comparisons ($g=0.33$, $p=0.02$). However, they were not significantly more effective than software programs without PAs ($g=0.26$, $p=0.16$), and they were significantly less effective than traditional training ($g=-0.31$, $p=0.02$), although there was only one study with two comparisons examining the use of traditional training to the use of PAs.

The PAs form and age were marginally significant moderators ($F(1.00, 4.22)=7.91$, $p=0.045$), and since the statistics are identical for both moderators, we report both together (see moderator table on OSF). Non-human PAs, where age

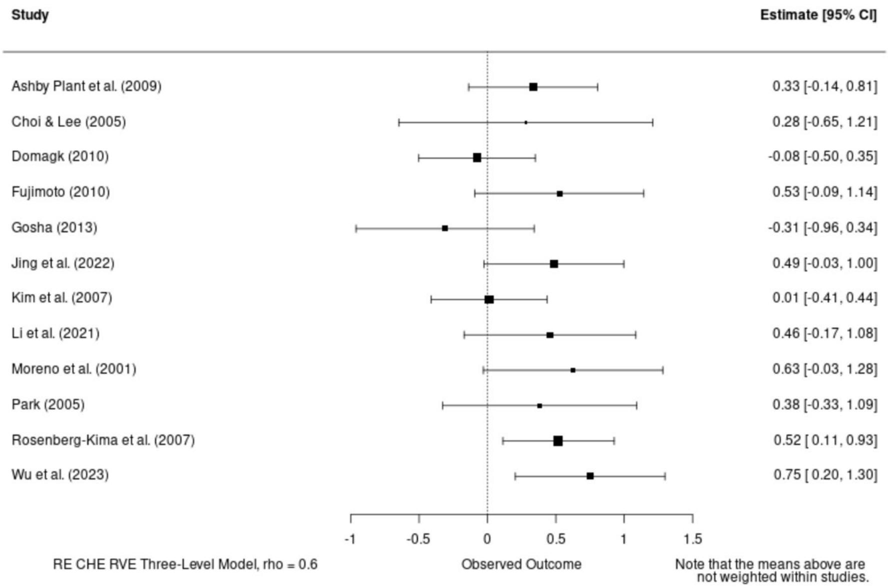

Fig. 9 Study-level forest plot of interest outcomes

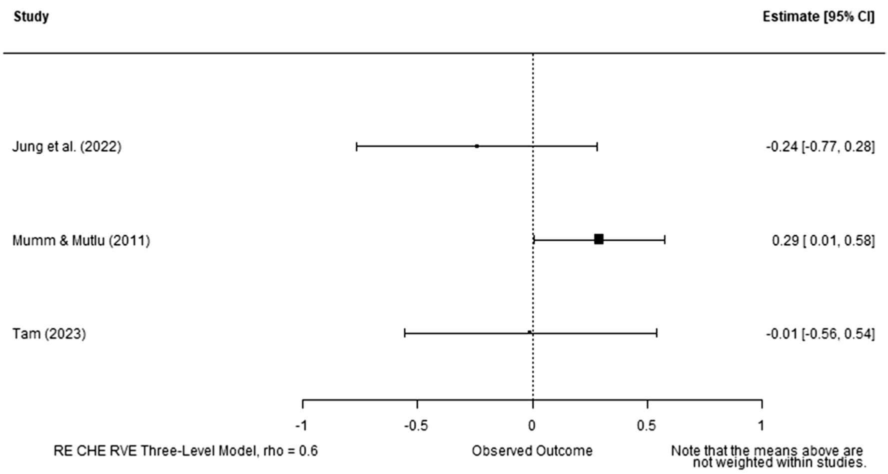

Fig. 10 Study-level forest plot of intrinsic motivation outcomes

was not relevant, were associated with a statistically significant effect ($g = 0.58$, $p = 0.01$); however, there are only four studies with eight comparisons that used non-human PAs, and the effect size was marginally significant, so this should be interpreted with caution.

Publication Bias

Examination of the comparison-level funnel plot shows that the effect sizes are quite symmetrically distributed. As such, publication bias is not viewed as a significant concern in this analysis.

RQ7: What Is the Impact of PAs on Learners' Intrinsic Motivation and What Moderates This Effect?

The random effects 3LMA with CHERVE examining the effects of PAs on learners' intrinsic motivation showed that PAs have essentially no influence ($g = 0.00$, $p = 0.99$, $k_{\text{comparisons}} = 23$, $k_{\text{studies}} = 3$). There was significant heterogeneity in the sample $Q(22) = 56.97$, $p < 0.001$, $\tau^2_{\text{within}} = 0.05$, $\tau^2_{\text{between}} = 0.01$. The model[1] explained 46.21% of the variation within the sample ($I^2_{\text{total}} = 46.21\%$), with variation within studies ($I^2_{\text{within}} = 39.44\%$) and less attributable to between studies ($I^2_{\text{between}} = 6.76\%$). Sensitivity analyses indicated that there were not notable differences between setting $\rho = 0.60$ versus $\rho = 0.40$ or $\rho = 0.80$. No outliers or influential studies were removed from the dataset. A forest plot is provided in Fig. 10.

Moderator Analyses

We did not examine potentially moderating variables because there were only three studies contributing comparisons.

Publication Bias

The comparison-level funnel plot showed that more studies appeared on the negative side of the overall effect size than the positive, indicating asymmetry. Since there were only three studies contributing effect sizes to this analysis, the study-level funnel plot is difficult to interpret for symmetry.

Discussion

Although investigating the use of PAs to motivate learners has increased over the last several years, it has remained to be seen how PAs influence specific motivational constructs (Heidig & Clarebout, 2011; Schroeder & Adesope, 2014). Therefore, in the present meta-analysis, we examined the impact of PAs on specific theory-driven motivation constructs and what may moderate the influence of PAs on these constructs. Alongside examining broad motivation, the specific motivation constructs included self-efficacy expectations, value beliefs (including utility value), engagement, interest, and intrinsic motivation. Overall, the meta-analyses revealed that using PAs influenced students' self-efficacy expectations, interest, and "motivation." However, the lack of significant findings for the other motivation constructs is similarly important to consider, as are the few significantly moderating variables. In the following sections, we discuss the findings and their implications for theory and practice.

Implications for Theory

This meta-analysis has implications for integrating two fields of study, PAs and learners' motivation, which are often siloed into their respective "research homes." One issue Heidig and Clarebout (2011) identified more than a decade ago was that of researchers in the field of PAs claiming that PAs could be used as motivational tools. However, those same studies often did not measure motivation compared to a non-PA control group. Previous reviews and meta-analyses around PAs have since found support for the idea that PAs can be used as motivational tools, at least in specific circumstances, if not more broadly (Guo & Gho, 2015; Heidig & Clarebout, 2011; Schroeder & Adesope, 2014; Wang et al., 2023). Yet, previous meta-analyses investigating the motivational effects of PAs only considered intrinsic motivation (Wang et al., 2023) or examined motivation as a broad construct and averaged multiple measures of motivation (Guo & Gho, 2015). In the present study, we did find more research examining specific motivation constructs. However, despite the increasing theoretical clarity in the field of motivation, more than 25% of the studies in our sample that examined "motivation" broadly rather than specific constructs were published after 2020. This shows a dire need for PA researchers to ensure there is conceptual clarity and theoretical grounding for their studies based on the well-established motivation literature. We discuss our findings in the context of the motivational theories that guided this work. However, before doing so, we briefly discuss the broad "motivation" measures and what the findings mean for the field.

Broad "Motivation" Implications

Alongside our specific motivation constructs of interest, we also examined "motivation" more broadly when authors did not measure or mention a specific motivational construct. We did find a small, significant effect of PAs on "motivation" ($g = 0.20$, $p = 0.03$). While we may have expected to find a significant effect because previous reviews have found that PAs significantly influence "motivation" (Guo & Gho, 2015; Schroeder et al., 2025), we do not feel as though this finding advances our understanding of motivational theory in relation to PAs due to different theoretical constructs being measured between studies. Rather, our finding simply leads to a call for more theoretically driven research and measurement in the field.

Consequently, it is of the utmost importance for researchers focused on using PAs as a motivational tool to situate their work within extant motivational theory to better ensure accuracy and consistency in the measurement of motivation constructs. While there is a broad consensus in the PA literature for differentiating between retention, comprehension, and transfer when assessing learning outcomes, there is not a consensus for assessing learners' motivation. The motivation constructs examined in this paper (self-efficacy expectations, value and utility beliefs, and engagement as tied back to SCT and SEVT, and learners' interest (interest theory) and intrinsic motivation (SDT)) are possible categories for theory-driven constructs that researchers can use in future research.

SCT and SEVT

We found that the presence of PAs compared to no PAs had a small, significant effect on learners' self-efficacy expectations. This finding is in line with SCT and SEVT, which posit that learners are likely to have a boost in their self-efficacy expectations when they observe or are told that someone else succeeded in a similar task (Bandura, 2001; Eccles & Wigfield, 2020; Schunk & DiBenedetto, 2020; Schunk & Usher, 2019). PAs may play this role in the learning environment since many PAs provide information, coaching, or scaffolding (Authors, in press; Schroeder & Gotch, 2015). This finding also replicates previous reviews, such as by Heidig and Clarebout (2011), that found an effect of PAs on self-efficacy (although only four studies in their sample measured self-efficacy). Therefore, it is promising that this finding remained with the inclusion of additional studies published since Heidig and Clarebout's review.

In the context of SEVT, we initially anticipated that PAs would influence learners' overall value beliefs (i.e., task relevance), utility value, and engagement. However, we found no significant effects of PAs on these constructs. Our synthesized results did not reveal a significant effect of PAs on learners' value beliefs or utility value. This is surprising, given the malleability of these constructs. The lack of significant effects may be due, in part, to the PAs and how they were implemented in the studies we examined. Gladstone and Cimpian (2021) argued that a role model would likely impact a learner's value beliefs if the learner perceived the role model to be meaningfully like them. This may not be the case in the studies we examined, as few considered any similarity between the learner and the PA. Thus, if a PA researcher is interested in impacting learners' value beliefs, careful consideration must be made regarding the agent's role[2] and whether the learner will perceive some degree of similarity (e.g., demographic and/or psychological) between themself and the PA. We speculate that designing PAs in such a way may significantly impact learners' value beliefs.

The lack of significant findings regarding engagement could be because engagement as a construct suffers from similar theoretical framing issues as measures of motivation. That is, engagement is a broad construct that is often conceptualized and operationalized differently across research groups (Sinatra et al., 2015). This is why incorporating engagement into a theoretical framework, such as SEVT (Gladstone et al., 2022), is important. Without a guiding theoretical framework and clear operationalizations, replicating or extending any of the previous PA research around the construct of engagement will be difficult.

Interest Theory

We found that PAs had a small to moderate significant effect on learners' interest. While this implies that, in a broad sense, PAs facilitate interest, it is important to note that the control condition moderated this effect. In particular, PAs boosted learners' interest significantly more than when learners just watched a non-interactive video. However, PAs did not significantly boost interest compared to an interactive software program that did not include a PA. These findings align with interest theory in that learners are likely to develop an interest in a topic when they are actively interacting with their social environment rather than passively receiving information (Hidi et al., 2004; Renninger, 2000).

Further, although we had 50 comparisons in the interest meta-analysis, they were extracted from only 12 studies, and therefore, the results should be interpreted in this context. For example, despite the overall meta-analysis showing a statistically significant positive result, we argue that the results of the control group moderator analysis demonstrate that one should not believe that the mere presence of a PA is enough to boost interest in all situations. Rather, the results indicate that the PA needs to have a role in the system that the learner is interacting with, but more

[2] There are a variety of perspectives on the role PAs can play in the learning environment (Baylor & Kim, 2005; Clarebout et al., 2002; Schroeder & Gotch, 2015). We do not argue that one approach is "better" or "worse" than others, as that was not the intention nor possible to parse from our analysis (i.e., our analyses do not compare PAs to other PA conditions, but rather a PA to a non-PA condition). Rather, we simply argue that it is a consideration designers must take into account.

research is needed. For example, research indicates that students prioritize which elements of a PA-enhanced learning environment they allocate attention to (Davis et al., 2024), so the role the agent plays in the learning environment could be a central consideration. Future work around the use of PAs to influence interest can focus on the role of the PA within the system to examine how to design PAs in this context most effectively.

SDT and Intrinsic Motivation

Based on a previous meta-analysis by Wang and colleagues (Wang et al., 2023) on the effects of affective PAs on learners' intrinsic motivation, we expected also to find an effect of PAs on intrinsic motivation. However, our results revealed a non-significant relationship between the use of PAs and learners' intrinsic motivation. One reason for this discrepancy could be that we used SDT to help guide our inclusion criteria and took a more conservative approach by only including studies that explicitly measured intrinsic motivation. Wang and colleagues classified various constructs into the broader intrinsic motivation category, whereas we used the specific terms authors of papers used to classify constructs. For example, both the current review and Wang and colleagues included the study by Guo et al. (2014). However, we did not include this study in our intrinsic motivation analysis, whereas Wang and colleagues did. Guo et al. do not specifically mention that they measured intrinsic motivation; rather, the authors measured attention, relevance, confidence, satisfaction, affective enjoyment, cognitive enjoyment, and behavioral enjoyment. While it is common practice for meta-analysts to lump measures into more 'meta' measures for the sake of their analysis (e.g., retention and transfer scores may both be considered 'learning'), it is difficult to know with certainty which of these constructs were classified as intrinsic motivation by Wang and colleagues. Having experienced these types of decision points in our prior meta-analytical work has driven our pursuit of transparency, which is one reason why we chose to not only classify constructs based on how the authors of the papers classified them but also pursue three-level meta-analytic models that allow each data point to be included rather than having to select one, or average "like" measures as a two-level meta-analytic model would require.

To summarize our work around intrinsic motivation and the implications of our findings, we believe that future work on PAs would greatly benefit from using SDT as a theoretical framework when researchers are interested in measuring intrinsic motivation. Otherwise, intrinsic motivation will likely continue to be used as an umbrella term encompassing many other motivation constructs, making it difficult to understand how PAs can impact specific aspects of learners' motivation.

Summary

Overall, we found that PAs can improve learners' self-efficacy expectations and interest. These findings suggest that PAs may be effective social models in learning contexts, which is consistent with predictions from SCT, and in some aspects SEVT (Bandura, 2001; Eccles & Wigfield, 2020; Gladstone & Cimpian, 2021; Schunk & DiBenedetto, 2020; Schunk & Usher, 2019), and interest theory (Hidi & Renninger, 2006; Renninger & Hidi, 2016, 2019). However, PAs were not found to be an effective tool for boosting learners' intrinsic motivation. This seems plausible given that a PA may not easily influence one's internal drive to learn something in a short time frame, and studies in the field are generally of relatively short duration. Therefore, focusing on intrinsic motivation in a short-term intervention may not reap the benefits that may be seen had the focus been on other motivation constructs. In general, these results indicated that PAs can be used as a motivational tool for specific purposes, such that focusing on learners' self-efficacy expectations and interest seems to be a promising approach.

Although we did not find significant effects for value beliefs, utility value, engagement, or intrinsic motivation, these findings are important to discuss. The lack of significant findings for these constructs is indicative of a larger, long-standing problem in the field, which is that there needs to be more cross-disciplinary collaborations to ensure better measurement and operationalization of these motivation constructs. As has been repeatedly noted in earlier reviews on PAs over the last 20 years (Clark & Choi, 2005; Dai et al., 2022; Schroeder & Adesope, 2014; Schroeder & Gotch, 2015), more coherent measures that are theoretically and psychometrically strong are needed in studies around PAs if we wish to have truly meaningful findings in the field.

Implications for Practice

The current review has several implications for practice. Interestingly, we found that the grade level of the participants significantly moderated the effects of PAs on learners' self-efficacy expectations. Specifically, the strongest effect was found for students in grades 7–9. This finding is particularly important when considering that students' self-efficacy expectations have been found to decline by the middle school years (Jacobs et al., 2002; Schunk & Miller, 2002). Therefore, if implemented appropriately, PAs could be used to help buffer the typical declines in confidence-related beliefs that adolescent learners experience. However, it should be noted that this finding is primarily driven by results from one study (Plant et al., 2009), and thus, future work should continue to examine this. We also found that PAs significantly impacted adult learners outside of postsecondary education. Although we did not find a significant moderator effect of domain, this finding has practical implications for the use of PAs in postsecondary domains that learners from traditionally marginalized groups report experiencing harmful stereotypes that lower their confidence-related beliefs in, such as STEM domains (Johnson et al., 2019; Stout et al., 2011). It is promising that PAs can boost postsecondary learners' confidence-related beliefs; however, research in the future needs to continue examining how PAs may be leveraged in specific domains for specific students at risk of declining confidence-related beliefs.

Our results, particularly those regarding interest, also highlight the importance of PA design. The results revealed that in order to use PAs as a motivational tool, it is not enough to simply include a character, but rather to think carefully about its function in the learning environment, its purpose in the learning process, and its design (Craig & Schroeder, 2018; Heidig & Clarebout, 2011). Heidig and Clarebout's (2011) framework can be applied to guide decisions on the design of the PA (PALD, pedagogical agents—levels of design) in accordance with the conditions of its use, such as the characteristics of the learning environment, the learner, and the function of the PA (PACU, pedagogical agents—conditions of use). Identifying specific motivational constructs rather than aiming for "broad motivation" may also be helpful for the practice of designing PAs. Clarity about the desired motivational effects of the PA can help define the function and purpose of the PA: Should it promote self-efficacy expectations, learner interest, or utility value?

Limitations and Directions for Future Research

Although the present study has several strengths, several limitations are worth noting. The first is that we were not able to examine the long-term impacts of PAs on learners' motivation. Some studies allow for longer periods (e.g., more than five hours) of interaction between the learner and the PA (Nielen et al., 2018; Osman & Lee, 2014), but studies that include delayed post-tests to detect long-term effects on motivational outcomes are scarce. Especially when considering motivation, effects may develop over time and not necessarily be detectable in immediate post-tests. If self-efficacy expectations and interest are positively affected by the presentation of PAs, this may encourage learners to continue working with the learning material, contributing to learning gains in the domain. Additionally, the situational interest created when interacting with the PA may, in the long run, turn into individual interest as a relatively stable affective orientation towards the subject matter presented (Schiefele, 2009). However, studies with delayed post-tests are needed to evaluate these effects.

We also could not examine the potential moderating roles of PA gender and ethnicity or learners' gender and ethnicity. Relatively few studies included in our review examined whether demographic characteristics of the PA or learners could moderate their effectiveness for certain students (e.g., Plant et al., 2009). It will be important for future research to examine these PA and learner characteristics more closely to understand better whether perceived demographic similarity matters in boosting learners' motivation. There has been a limited amount of research on this topic in the past in regard to learning; however, the results were mixed, and generally speaking, improving specific motivational constructs was not the focus of the work (Huang et al., 2022; Makransky et al., 2019; Moreno & Flowerday, 2006; Ozogul et al., 2013).

Finally, it is important to note methodological limitations. Of note is that some of the constructs, such as intrinsic motivation, included relatively few studies. Thus, caution should be taken when interpreting these results, and we have explicitly reported the number of studies and comparisons for each analysis to try and make this information transparent. In addition, we noted that the field lacks publication bias tests for three-level meta-analytic models. We view this as a minor hindrance rather than something that strongly influences the interpretation of our findings, as even with conventional meta-analytic models, to the best of our knowledge, no one publication bias test has been broadly agreed upon as the "best" indicator of "publication bias." Rather, publication bias can be thought of as adding contextual information that aids in interpreting the model. While lacking this contextual information is less than optimal, we believe that the benefits of 3LMA with CHERVE, specifically utilizing all available data, accounting for correlations between dependent data, and creating robust estimates, outweigh the data loss that would occur when using conventional meta-analytic models that have various publication bias tests available.

Conclusion

Pedagogical agents are often claimed to facilitate learner motivation. However, motivation is a complex construct that has not often been closely examined for theoretical alignment in this field. Consequently, we dissected "motivation" into theory-driven constructs and meta-analyzed the primary studies in the field to better understand to what extent and how PAs can influence various theoretical perspectives of motivation. We conducted seven 3LMA with CHERVE and found that PAs can significantly influence learners' self-efficacy expectations, interest, and measures of "motivation." Meanwhile, they did not significantly influence intrinsic motivation, value beliefs, utility beliefs, or engagement. Consistent with prior research, our results show the importance of intentionality in PA design and implementation and highlight that even though it has been noted in the literature for nearly 20 years, researchers in the field should use theoretically and psychometrically strong measures in their studies.

Funding This research was supported by the National Science Foundation and the Institute of Education Sciences under Grant #2229612.

Declarations

Conflict of interest The authors declare no competing interests.

References

*indicates study included in the meta-analysis

*Albright, G., Fazel, M., Khalid, N., McMillan, J., Hilty, D., Shockley, K., & Joshi, S. (2022). High school educator training by simulation to address emotional and behavioral concerns in school settings: A randomized study. *Journal of Technology in Behavioral Science, 7*(3), 277–289. https://doi.org/10.1007/s41347-022-00243-9

*Alyahya, S. M. (2021). *Social cues in animated pedagogical agents for second language learners: The application of the embodiment principle in video design* [Doctoral dissertation, University of South Florida].

*Arguedas, M., & Daradoumis, T. (2021). Analysing the role of a pedagogical agent in psychological and cognitive preparatory activities. *Journal of Computer Assisted Learning, 37*(4), 1167–1180. https://doi.org/10.1111/jcal.12556

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

*Baylor, A. L. (2002). Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents. *Educational Technology Research and Development, 50*(2), 5–22.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development, 59*, 291–300. https://doi.org/10.1007/s11423-011-9196-3

Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15*(2), 95–115.

*Bøg Petersen, G., Klingenberg, S., & Makransky, G. (2022). Pipetting in virtual reality can predict real-life pipetting performance. *Technology, Mind, and Behavior*, *3*(3). https://doi.org/10.1037/tmb0000076

Campos, D. G., Fütterer, T., Gfrörer, T., Lavelle-Hill, R., Murayama, K., König, L., Hecht, M., Zitzmann, S., & Scherer, R. (2024). Screening smarter, not harder: A comparative analysis of machine learning screening algorithms and heuristic stopping criteria for systematic reviews in educational research. *Educational Psychology Review, 36*(1), 19. https://doi.org/10.1007/s10648-024-09862-5

Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories: A meta-analysis. *Educational Psychology Review, 33*, 989–1015. https://doi.org/10.1007/s10648-020-09587-1

*Chen, C., Jamiat, N., Rabu, S. N. A., & Du, S. (2025). Effects of augmented reality with pedagogical agent on EFL primary students' vocabulary acquisition, motivation, and technology perceptions. *International Journal of Human–Computer Interaction,* 1–19. https://doi.org/10.1080/10447318.2024.2443534

*Chen, J., Mokmin, N. A. M., Shen, Q., & Su, H. (2025). Leveraging AI in design education: Exploring virtual instructors and conversational techniques in flipped classroom models. *Education and Information Technologies*, 1–21. https://doi.org/10.1007/s10639-025-13458-z

*Choi, S. (2005). *Cognitive efficiency of animated pedagogical agents for learning English as a second language* [Doctoral Dissertation, University Of Southern California].

*Choi, S., & Lee, H. (2005). Cognitive and motivational effects of animated pedagogical agent for learning English as a second language. *Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology*, 771–773.

Clarebout, G., Elen, J., Johnson, W. L., & Shaw, E. (2002). Animated pedagogical agents: An opportunity to be grasped? *Journal of Educational Multimedia and Hypermedia,11*(3), 267–286.

Clark, R. E., & Choi, S. (2005). Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research, 32*(3), 209–225.

*Cobb, C. (2013). *The use of an animated pedagogical agent as a mnemonic device to promote learning and motivation in online education* [Doctoral Dissertation, Walden University].

*Costa, S., Brunete, A., Bae, B., & Mavridis, N. (2018). Emotional storytelling using virtual and robotic agents. *International Journal of Humanoid Robotics*. https://doi.org/10.1142/S0219843618500068

Craig, S. D., & Schroeder, N. L. (2018). *Design principles for virtual humans in educational technology environments*. In Deep Comprehension (pp. 128–139). Routledge.
Dai, L., Jung, M. M., Postma, M., & Louwerse, M. M. (2022). A systematic review of pedagogical agent research: Similarities, differences and unexplored aspects. *Computers & Education*. https://doi.org/10.1016/j.compedu.2022.104607
Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review, 24*, 193–209. https://doi.org/10.1016/j.edurev.2018.05.002
Davis, R. O., Lee, Y. J., Vincent, J., & Wan, L. (2024). Exploring gesture frequencies and images in multimedia environments with pedagogical agents. *Journal of Computer Assisted Learning,40*(6), 3055–3071.
Davis, R. O., Park, T., & Vincent, J. (2023). A meta-analytic review on embodied pedagogical agent design and testing formats. *Journal of Educational Computing Research, 61*(1), 30–67. https://doi.org/10.1177/07356331221100556
*Dincer, S., & Doganay, A. (2015). The impact of pedagogical agent on learners' motivation and academic success. *Practice and Theory in Systems of Education, 10*(4), 329–348.
*Dinçer, S., & Doğanay, A. (2017). The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load. *Computers & Education, 111*, 74–100. https://doi.org/10.1016/j.compedu.2017.04.005
Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes? The role of the appeal of agent's appearance and voice. *Journal of Media Psychology: Theories, Methods, and Applications,22*(2), 84–97. https://doi.org/10.1027/1864-1105/a000011
Domagk, S., & Niegemann, H. M. (2005). Pedagogical agents in multimedia learning environments: Do they facilitate or hinder learning?. In *Towards sustainable and scalable educational innovations informed by the learning sciences* (pp. 654–657). IOS Press.
Eccles, J. S., Wang, M. T., (2012). Part 1 commentary: So what is student engagement anyway? Christenson, S. L., Reschly, A. L., Wylie, C. (Eds.), *Handbook of research on student engagement Springer*, 133-145.
Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*. https://doi.org/10.1016/j.cedpsych.2020.101859
*Ferrer, V., Perdomo, A., Ali, H., Fies, C., Quarles, J., & IEEE. (2017). Virtual humans for temperature visualization in a tangible augmented reality educational game. In *2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR)* (pp. 1–6).
*Fitton, I. S., Finnegan, D. J., & Proulx, M. J. (2020). Immersive virtual environments and embodied agents for e-learning applications. *Peerj Computer Science, 6*, e315.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059
*Fujimoto, T. (2010). *Story-based pedagogical agents: A scaffolding design approach for the process of historical inquiry in a web-based self-learning environment* [Doctoral Dissertation, The Pennsylvania State University].
*Genova, H., Lancaster, K., Morecraft, J., Haas, M., Edwards, A., DiBenedetto, M., Krch, D., DeLuca, J., & Smith, M. (2021). A pilot RCT of virtual reality job interview training in transition-age youth on the autism spectrum. *Research in Autism Spectrum Disorders,89*, Article 101878. https://doi.org/10.1016/j.rasd.2021.101878
Gladstone, J. R., & Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM. *International Journal of STEM Education, 8*, 1–20. https://doi.org/10.1186/s40594-021-00315-x
Gladstone, J. R., Wigfield, A., spsampsps Eccles, J. S. (2022). Situated expectancy-value theory, dimensions of engagement, and academic outcomes. In S. L. Christenson spsampsps A. L. Reschly (Eds.), *Handbook of research on student engagement* (pp. 57–76). Springer International Publishing.
Gosha, K. (2013). *The application of embodied conversational agents for mentoring African American stem doctoral students* [Doctoral Dissertation, Clemson University].
Gottsacker, M., Norouzi, N., Schubert, R., Guido-Sanz, F., Bruder, G., & Welch, G. (2022). Effects of environmental noise levels on patient handoff communication in a mixed reality simulation.

Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology. https://doi.org/10.1145/3562939.3565627

*Govindasamy, M. K. (2013). Embodied agent in tutor role: Effects on field dependent and independent low achiever's retention and perceived science self-efficacy beliefs. *Journal of Educational Multimedia and Hypermedia, 22*(3), 273–297.

Guo, Y. R., & Goh, D.H.-L. (2015). Affect in embodied pedagogical agents: Meta-analytic review. *Journal of Educational Computing Research,53*(1), 124–149. https://doi.org/10.1177/0735633115588774

*Guo, Y. R., & Goh, D.H.-L. (2016). Evaluation of affective embodied agents in an information literacy game. *Computers & Education, 103*, 59–75. https://doi.org/10.1016/j.compedu.2016.09.013

Guo, Y. R., Goh, D. H. L., & Luyt, B. (2014). Using affective embodied agents in information literacy education. *IEEE/ACM Joint Conference on Digital Libraries*, 389–398. https://doi.org/10.1109/JCDL.2014.6970195

Guo, Y. R., Goh, D.H.-L., Luyt, B., Sin, S.-C.J., & Ang, R. P. (2015). The effectiveness and acceptance of an affective information literacy tutorial. *Computers & Education,87*, 368–384. https://doi.org/10.1016/j.compedu.2015.07.015

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Systematic Reviews*. https://doi.org/10.1002/cl2.1230

Harackiewicz, J. M., Tibbetts, Y., Canning, E., & Hyde, J. S. (2014). Harnessing values to promote motivation in education. In *Motivational interventions* (pp. 71–105). Emerald Group Publishing Limited.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). Doing meta-analysis with R: A hands-on guide Chapmann spsampsps Hall/CRC Press 978-0-367-61007-4.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning DY Dai RJ Sternberg Eds Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development. Lawrence Erlbaum Associates Inc, 89-115.

Huang, X., Mathews, J. L., & Hsiao, E. (2022). Choice of pedagogical agents as virtual math tutors: Perspectives from children and college students. *Journal of Educators Online, 19*(2).

*Huang, X., & Mayer, R. E. (2016). Benefits of adding anxiety-reducing features to a computer-based multimedia lesson on statistics. *Computers in Human Behavior, 63*, 293–303. https://doi.org/10.1016/j.chb.2016.05.034

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development,73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

*Jing, B., Liu, J., Gong, X., Zhang, Y., Wang, H., & Wu, C. (2022). Pedagogical agents in learning videos: Which one is best for children? *Interactive Learning Environments*, 1–17. https://doi.org/10.1080/10494820.2022.2141787

Johnson, I. R., Pietri, E. S., Fullilove, F., & Mowrer, S. (2019). Exploring identity-safety cues and allyship among Black women students in STEM environments. *Psychology of Women Quarterly, 43*(2), 131–150. https://doi.org/10.1177/0361684319830926

*Jung, J.-Y., Qiu, S., Bozzon, A., & Gadiraju, U. (2022). Great chain of agents: The role of metaphorical representation of agents in conversational crowdsourcing. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3491102.3517653

*Kim, C., Keller, J. M., & Baylor, A. L. (2007). effects of motivational and volitional messages on attitudes toward engineering: Comparing text messages with animated messages delivered by a pedagogical agent. *Proceedings of the IADIS International Conference on Cognition & Exploratory Learning in Digital Age*, 317–320.

*Kim, Y. (2009). The role of learner attributes and affect determining the impact of agent presence. *International Journal of Learning Technology, 4*(3/4), 2–2. https://doi.org/10.1504/IJLT.2009.028808

*Kimani, E., Parmar, D., Murali, P., & Bickmore, T. (2021). Sharing the load online: Virtual presentations with virtual co-presenter agents. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411763.3451670

Kolodkin, A., Boogerd, F. C., Plant, N., Bruggeman, F. J., Goncharuk, V., Lunshof, J., Moreno-Sanchez, R., Yilmaz, N., Bakker, B. M., Snoep, J. L., Balling, R., & Westerhoff, H. V. (2012). Emergence of the silicon human and network targeting drugs. *European Journal of Pharmaceutical Sciences, 46*(4), 190–197.

Koprinkova-Hristova, P., Oubbati, M., & Palm, G. (2013). Heuristic dynamic programming using echo state network as online trainable adaptive critic. *International Journal of Adaptive Control and Signal ProcessiNg,27*(10), 902–914.

*Lee, T. T., & Mustapha, N. H. (2017). Who is more efficient: Teacher or pedagogical agents? *AIP Conference Proceedings, 1847*(1). https://doi.org/10.1063/1.4983903

*Li, W., Wang, F., Mayer, R. E., & Liu, T. (2022). Animated pedagogical agents enhance learning outcomes and brain activity during learning. *Journal of Computer Assisted Learning, 38*(3), 621–637. https://doi.org/10.1111/jcal.12634

*Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I can help you change! An empathic virtual agent delivers behavior change health interventions. *Acm Transactions On Management Information Systems*. https://doi.org/10.1145/2544103

Makransky, G., Wismer, P., & Mayer, R. E. (2019). A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *Journal of Computer Assisted Learning, 35*(3), 349–358.

*Manuvinakurike, R., Velicer, W. F., & Bickmore, T. W. (2014). Automated indexing of internet stories for health behavior change: Weight loss attitude pilot study. *Journal of Medical Internet Research, 16*(12), e285–e285. https://doi.org/10.2196/jmir.3702

*Martha, A. S. D., Santoso, H. B., Junus, K., & Suhartanto, H. (2023). The effect of the integration of metacognitive and motivation scaffolding through a pedagogical agent on self-and co-regulation learning. *IEEE Transactions on Learning Technologies, 16*(4), 573–584.

*Mathews, J. L. (2021). *The effect of pedagogical agent persona on performance, self-efficacy, and attitudes in adult learners in an online environment* [Doctoral Dissertation, Western Kentucky University].

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419–425. https://doi.org/10.1037/0022-0663.95.2.419

*McCloud, A. M. (2012). *Pedagogical agents: The effects of agent visual fidelity on perceived self-efficacy* [Doctoral Dissertation, Northern Illinois University].

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31*(2), 186–207.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*(2), 177–213. https://doi.org/10.1207/S1532690XCI1902_02

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology, 19*, 465–483. https://doi.org/10.1037/gpr0000059

*Mumm, J., & Mutlu, B. (2011). Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback. *Computers in Human Behavior, 27*(5), 1643–1650. https://doi.org/10.1016/j.chb.2011.02.002

*National Center for Education Statistics. (n.d.). *Education indicators: An international perspective/indicator 1 side bar. ISCED Levels of Education.* https://nces.ed.gov/pubs/eiip/eiip1s01.asp

Nielen, T. M. J., Smith, G. G., Sikkema-de Jong, M. T., Drobisz, J., van Horne, B., & Bus, A. G. (2018). Digital guidance for susceptible readers: Effects on fifth graders' reading motivation and incidental vocabulary learning. *Journal of Educational Computing Research, 56*(1), 48–73. https://doi.org/10.1177/0735633117708283

Osman, K., & Lee, T. T. (2014). Impact of interactive multimedia module with pedagogical agents on students' understanding and motivation in the learning of electrochemistry. *International Journal of Science and Mathematics Education, 12*(2), 395–421. https://doi.org/10.1007/s10763-013-9407-y

Ozogul, G., Johnson, A. M., Atkinson, R. K., & Reisslein, M. (2013). Investigating the impact of pedagogical agent gender matching and learner choice on learning outcomes and perceptions. *Computers & Education,67*, 36–50.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71
*Park, S. (2005). *The effects of seductive augmentation and agent role on learning interest, achievement, and attitude* [Doctoral Dissertation, Florida State University].
Pekrun, R. (2024). Overcoming fragmentation in motivation science: Why, when, and how should we integrate theories? *Educational Psychology Review, 36*(1), 27. https://doi.org/10.1007/s10648-024-09846-5
*Pi, Z., Deng, L., Wang, X., Guo, P., Xu, T., & Zhou, Y. (2022). The influences of a virtual instructor's voice and appearance on learning from video lectures. *Journal of Computer Assisted Learning, 38*(6), 1703–1713. https://doi.org/10.1111/jcal.12704
Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education,53*(2), 209–215. https://doi.org/10.1016/j.compedu.2009.01.013
*Polat, H., Taş, N., Kaban, A., Kayaduman, H., & Battal, A. (2024). Human or humanoid animated pedagogical avatars in video lectures: The impact of the knowledge type on learning outcomes. *International Journal of Human–Computer Interaction*, 1–16. https://doi.org/10.1080/10447318.2024.2415762
Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science, 23*(3), 425–438.
Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone, J. M. Harackiewicz, C. Sansone, & J. M. Harackiewicz (Eds.), *Intrinsic motivation: Controversies and new directions* (pp. 373–404). Academic Press. https://doi.org/10.1016/B978-012619070-0/50035-0
Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and engagement*. Routledge.
Renninger, K. A., & Hidi, S. E. (2019). Interest and internal motivation. *The Cambridge Handbook of Motivation and Learning* (pp. 263–394). Cambridge University Press.
Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods, 26*(2), 141–160. https://doi.org/10.1037/met0000300
*Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2007). The importance of interface agent visual presence: Voice alone is less effective in impacting young women's attitudes toward engineering. In Y. De Kort, W. IJsselsteijn, C. Midden, B. Eggen, spsampsps B. J. Fogg (Eds.), *Persuasive Technology* (Vol. 4744, pp. 214–222). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77006-0_27
Rosenberg-Kima, R. B., Plant, E. A., Doerr, C. E., & Baylor, A. L. (2010). The influence of computer-based model's race and gender on female students' attitudes and beliefs towards engineering. *Journal of Engineering Education,99*(1), 35–44. https://doi.org/10.1002/j.2168-9830.2010.tb01040.x
Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68. https://doi.org/10.1037/0003-066X.55.1.68
*San Pedro, M., Rodrigo, M., & Baker, R. (2011). The relationship between carelessness and affect in a cognitive tutor. In S. D'Mello, A. Graesser, B. Schuller, & J. Martin (Eds.), *Affective Computing and Intelligent Interaction Proceedings Part I, 4* (pp. 306–315). Springer.
Schiefele, U. (2009). Situational and individual interest. In K. R. Wenzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197–222). Routledge/Taylor & Francis Group.
*Scholten, M. R., Kelders, S. M., & Van Gemert-Pijnen, J. E. W. C. (2019). An empirical study of a pedagogical agent as an adjunct to an ehealth self-management intervention: What modalities does it need to successfully support and motivate users? *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2019.01063
*Schouten, D., Deneka, A., Theune, M., Neerincx, M., & Cremers, A. (2022). An embodied conversational agent coach to support societal participation learning by low-literate users. *Universal Access in the Information Society, 22*, 1215–1241. https://doi.org/10.1007/s10209-021-00865-5

Schroeder, N. L. (2024). Simple meta-analysis. Available from https://github.com/noah-schroeder/simple_meta-analysis/

Schroeder, N. L., & Adesope, O. O. (2014). A systematic review of pedagogical agents' persona, motivation, and cognitive load implications for learners. *Journal of Research on Technology in Education, 46*(3), 229–251.

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research, 49*(1), 1–39. https://doi.org/10.2190/EC.49.1.a

Schroeder, N. L., & Gotch, C. M. (2015). Persisting issues in pedagogical agent research. *Journal of Educational Computing Research, 53*(2), 183–204.

Schroeder, N. L., Zhang, S., Jaldi, C. D., Gladstone, J. R., Lopez, A. A., & Dorley, E. (2025). *Pedagogical agents can help kids learn: A series of three-level meta-analyses*. Denver, Colorado: American Educational Research Association.

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*. https://doi.org/10.1016/j.cedpsych.2019.101832

Schunk, D.H., & Miller, S.D. (2002). Self-efficacy and adolescents' motivation. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents*. Greenwich, CT: Information Age

Schunk, D. H., & Usher, E. L. (2019). Social cognitive theory and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (2nd ed., pp. 11–26). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190666453.013.2

Siegle, R. F., Schroeder, N. L., Lane, H. C., & Craig, S. D. (2023). Twenty-five years of learning with pedagogical agents: History, barriers, and opportunities. *Techtrends,67*(5), 851–864.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

*Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., Olsen, D. E., Bell, M. D., & Fleming, M. F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(10), 2450–2463. https://doi.org/10.1007/s10803-014-2113-y

*Smith, M. J., Ginger, E. J., Wright, M., Wright, K., Boteler Humm, L., Olsen, D., Bell, M. D., & Fleming, M. F. (2014). Virtual reality job interview training for individuals with psychiatric disabilities. *Journal of Nervous & Mental Disease,202*(9), 659–667. https://doi.org/10.1097/NMD.0000000000000187

*Song, D., & Kim, D. (2021). Effects of self-regulation scaffolding on online participation and learning outcomes. *Journal of Research on Technology in Education, 53*(3), 249–263. https://doi.org/10.1080/15391523.2020.1767525

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology,100*, 255–270. https://doi.org/10.1037/a0021385

*Suarez, G., Jung, S., & Lindeman, R. (2021). Evaluating virtual human role-players for the practice and development of leadership skills. *Frontiers in Virtual Reality*. https://doi.org/10.3389/frvir.2021.658561

Sukthankar, G., & Sycara, K. (2011). Activity recognition for dynamic multi-agent teams. *ACM Transactions on Intelligent Systems and Technology*. https://doi.org/10.1145/2036264.2036282

*Tam, F. L. (2023). *The impact of game design elements on adolescents' motivation and executive function skills in a cognitive skills training game* [Doctoral Dissertation, New York University].

*Tanaka, H., Saga, T., Iwauchi, K., Honda, M., Morimoto, T., Matsuda, Y., Uratani, M., Okazaki, K., & Nakamura, S. (2023). The validation of automated social skills training in members of the general population over 4 weeks: Comparative study. *JMIR Formative Research, 7*, e44857. https://doi.org/10.2196/44857

*Tzeng, S.-C., & Wang, P.-T. (2014). Identification features and pedagogical agents in a mathematical game. *2014 Computer Games: AI, Animation, Mobile, Multimedia, Educational and Serious Games*, 1–5. https://doi.org/10.1109/CGames.2014.6934153

van der Meij, H. (2013). Motivating agents in software tutorials. *Computers in Human Behavior, 29*(3), 845–857. https://doi.org/10.1016/j.chb.2012.10.018

van der Meij, H., van der Meij, J., & Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment.

Educational Technology Research and DevelopmeNt, 63(3), 381–403. https://doi.org/10.1007/s11423-015-9378-5

*van Lissa, C. (n.d.). *Doing meta-analysis in R and exploring heterogeneity using metaforest*. https://cjvanlissa.github.io/ng-Meta-Analysis-in-R/index.html

*Viechtbauer, W. (n.d.). *Model diagnostics for "rma.mv" objects—Influence.rma.mv*. https://wviechtb.github.io/metafor/reference/influence.rma.mv.html

*Viechtbauer, W. (n.d.b). *Model diagnostics for "rma.uni" objects—Influence.rma.mv*. https://wviechtb.github.io/metafor/reference/influence.rma.uni.html

Viechtbauer, W. (2022). *Package "metafor."* https://cran.r-project.org/web/packages/metafor/metafor.pdf

Viechtbauer, W., & Cheung, M.W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods,1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Wang, J., Li, X., Pan, L., & Zhang, C. (2021). Parametric 3d modeling of young women's lower bodies based on shape classification. *International Journal of Industrial Ergonomics*. https://doi.org/10.1016/j.ergon.2021.103142

Wang, Y., Gong, S., Cao, Y., Lang, Y., & Xu, X. (2023). The effects of affective pedagogical agent in multimedia learning environments: A meta-analysis. *Educational Research Review*. https://doi.org/10.1016/j.edurev.2022.100506

Wigfield, A., & Koenka, A. C. (2020). Where do we go from here in academic motivation theory and research? Some reflections and recommendations for future work. *Contemporary Educational Psychology*. https://doi.org/10.1016/j.cedpsych.2020.101872

*Wu, C., Jing, B., Gong, X., & Ma, X. (2023). The zoomorphic effect: A contribution to the study of images of pedagogical agents for children's learning in instructional videos. *Journal of Computer Assisted Learning, 39*(5), 1620–1635.

*Wu, C., Li, Y., Yang, H., Wang, X., Li, X., & Jing, B. (2024). The greater the interaction, the stronger the learning performance? Examining pedagogical agents' interactive presence in instructional videos. *Applied Cognitive Psychology,38*(5), Article e4242.

Zeitlhofer, I., Zumbach, J., & Aigner, V. (2023). Effects of pedagogical agents on learners' knowledge acquisition and motivation in digital learning environments. *Knowledge, 3*(1), 53–67.