Learning and Individual Differences 122 (2025) 102727

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Bringing learners into focus: A systematic review of learner characteristics in AR-supported STEM education☆

Zoya Kozlova *, Katharina M. Bach, Peter A. Edelsbrunner, Sarah I. Hofer

Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-University, 80802 Munich, Germany

A R T I C L E I N F O

Keywords:
Augmented reality
Individual differences
Learning outcome
STEM education
ATI perspective

A B S T R A C T

To understand how AR effects on STEM learning depend on individual differences, it is essential to follow the ATI (Aptitude-Treatment Interaction) perspective and investigate interactions between individual differences and AR- vs. non-AR conditions. This systematic review explored the extent to which individual characteristics are examined in AR research as predictors to further review if and how AR research in STEM education follows an ATI approach. Our findings reveal that from 2013 to 2022, $k = 38$ studies investigated the role of individual variables as predictors with only $k = 5$ studies considering how individual differences interact with AR vs. non-AR conditions. Spatial ability emerged as the most frequently studied learner characteristic in ATI-AR research, yet its impact on learning outcomes remains inconclusive. We discuss possible reasons for this gap and propose solutions, offering a study design framework to conduct AR studies considering the ATI perspective.

The integration of cutting-edge technologies like artificial intelligence (AI), Virtual Reality (VR), and Augmented Reality (AR) into education provides considerable potential for personalized learning. Through interactive guidance and real-time adaptation to students responses and behaviors (Bhutoria, 2022; Marougkas et al., 2023), these technological advancements allow for prioritizing the learner and addressing their individual interests and needs, thus realizing the fundamental aspect of personalized learning (Bernacki et al., 2021).

In other words, in personalized learning, learners characteristics shape the constituents of the learning path and the ways the path is paved, namely, instruction, such as formats and methods used in teaching. This reflects the key postulate of the aptitude-treatment interaction (ATI), formulated back in the 1970s. According to Snow (1977), an individual's abilities or characteristics (aptitudes) respond differently to specific educational interventions (treatment).

In a traditional classroom, any instructional intervention is a mere element of the educational environment the learner is already familiar with. The use of collaborative learning in a classroom is an instance of such an intervention. Collaborative learning is an instructional format that presupposes the students to work together. Even though this sort of group cooperation might be novel to students, the other learning constituents remain familiar, for example, the classroom settings, group/interpersonal interaction, per se. Watching a video is another instance of a format variation in learning and instruction. Bringing some novelty to the instructional format, interaction with the video screening still takes place in the learning environment (the classroom) the learners have experience dealing with.

In the case of immersive technology, such as VR and AR, the learner gets exposed to a novel surrounding (or environment), for example, being inside the human body in a VR learning environment or seeing the solar system in AR glasses in the classroom. In the case of VR, the new environment is fully immersive, i.e., the learner is meant to completely surrender themselves to the environment to start feeling physically present in a different place or setting. For AR, which this study is focused on, the immersion is usually manifested through some non-regular features displayed in the still familiar surroundings (e.g., seeing the planets of the Solar system in the familiar classroom).

Beyond novelty effects, such environmental experiences in AR also have a variety of characteristics of their own that may affect, and, crucially, put new demands on the learner. For example, they present information in three dimensions instead of two dimensions such as in written information, and they provide potential for immersion.

These environmental characteristics, which we label here *AR affordances*, should not be seen to automatically provide additional learning potential. Whether and to which extent AR supports - or even hinders - learning might depend on the instructional design used and the

☆ This article is part of a Special issue entitled: ‘Metaanalyses & Systematic Reviews’ published in Learning and Individual Differences.

* Corresponding author at: Leopoldstr. 13, 80802 München, Germany.
E-mail address: zoya.kozlova@psy.lmu.de (Z. Kozlova).

https://doi.org/10.1016/j.lindif.2025.102727
Received 10 May 2024; Received in revised form 4 May 2025; Accepted 12 May 2025
Available online 13 June 2025

characteristics that learners require to realize the full potential of these AR affordances. For some learners, AR might be supportive because they possess the spatial abilities required to process augmented information (Krüger et al., 2022). For other learners, AR affordances might be unnecessary, or overwhelming because they might induce additional cognitive load that prevents learners from integrating the new information with their prior knowledge (Ling et al., 2021).

Yet, to the best of our knowledge, there is no review of research on the technological affordances of AR and their interaction with learner characteristics available. Without such an overview, it is difficult to gauge which questions in this regard have been worked on and to draw conclusions regarding the specific AR affordances × learner characteristics interaction. Given that AR affordances, such as visualizing the invisible and showing the object in a 3D perspective, are particularly fitting for learning activities in STEM subjects, we conducted a systematic literature review to investigate the extent to which individual characteristics and their interaction with the AR environment have been investigated in STEM fields. Examining this question, we aim to provide a comprehensive overview of the current landscape within individual differences research related to learning with AR. We first identify studies that treat individual differences as predictors for the learning outcome. In a second step, we examine which of these studies followed the ATI perspective. That means that our systematic review encompasses all studies that examine the individual variables as predictors of learning outcomes in AR learning settings, and then we focus on studies that look at the interaction of the effect that an AR condition has in comparison to another condition with learners' individual characteristics. Only the latter study design allows for determining which demands on learners' individual characteristics are specific to AR and not more generally required in STEM learning. Based on this perspective, we illuminate exemplary methodologies and map out a prospective research agenda.

1. Theoretical background

1.1. Personalized learning

In the past years, there have been multiple attempts to coin (or recycle) the universal definition of personalized learning. One of the most common definitions, developed by Spector (2018) characterizes personalized learning as the learning environment adapted to the learners individual knowledge and interest. Technological affordances come crucial in the implementation of such an environment in order to adjust the pace and content of instruction based on individual learner performance (OECD, 2006; Shemshack et al., 2021; Spector, 2014).

This implies the necessity to integrate the learners' needs and specificities into the learning environment afforded by technology. For instance, Benhamdi et al. (2017) integrated personalization in the learning environment by accounting for students' preferences, interests, background knowledge, and working memory capacity. Narciss et al. (2014) considered the learners' motivation, prior knowledge, and gender differences in adapting the learning material to the learners' needs.

The realization of such adaptivity presupposes a clear-cut idea of how exactly different learners employ or/and react to different learning environments enabled by technology. The idea of tailoring instruction to accommodate students' individual differences is not new. Locating the most effective instructional strategies to meet the varying needs of students was articulated back in the 70s at the dawn of ATI and is currently undergoing a revival. However, today, there is still little evidence on how individual differences align with different instructional approaches (Tetzlaff et al., 2023). In the next section, we are going to delve into what exactly ATI means and explore the state-of-the-art research in learning with technology, particularly in STEM education.

1.2. Aptitude-treatment interaction research in STEM

The key premise of the ATI concept is that the same instructional setting may affect different learners differently and that this difference is conditioned by inherent learner characteristics (Snow, 1980). Learner characteristics - unique traits of the learners - have been classified in various ways (Murphy, 2012; Sackett et al., 2017). In this review, we utilize the individual differences classification developed by Vandewaetere et al. (2011) in relation to the technology use in personalized learning. According to this classification, individual differences expand across the three main domains cognition, affect, and behavior. The first domain encompasses cognition-related traits such as intelligence, prior knowledge, working memory capacity, and reasoning ability. The affect domain includes a vast variety of affective traits (note that these are also based on cognition, but in contrast to the cognition-related traits, they exclude for example skills and abilities; Kell, 2018), such as motivation, self-efficacy, interest, engagement, or creativity. The behavior domain contains the so-called interaction parameters (not to be confused with statistical interaction), that is, the individual characteristics occurring when interacting with the learning environment (e.g., self-regulation and help-seeking). These characteristics are also closely connected to cognitive and/or affective domains.

Normally, aptitude is treated as a moderator, a variable influencing the strength or direction of the relationship between the outcome variable (e.g., knowledge of electromagnetism) and the treatment (e.g., learning with a traditional experimental set-up or with AR; Preacher & Sterba, 2019). It is also necessary to check the presence or absence of the interaction between a specific learner characteristic (e.g., prior knowledge) and the treatment (e.g., introduction to electromagnetism through magnetic field demonstrations with iron filings) and its magnitude - i.e., how differently learners with various levels of aptitude will react to different treatments. For this reason, a control group is required to provide a baseline for comparison for these different treatments. A control group allows to isolate and examine the specific effects of moderator variables on the relationships between predictor and outcome variables, namely the individual characteristic and the treatment, to properly assess the moderator effects across various contexts.

Since the present research is concerned with individual characteristics that are required for learning successfully from the affordances of AR and the affordances-informed instructional design, a control condition in our case is any comparison condition that does not use AR or any other technology (e.g., a traditional STEM learning condition using pen & paper or ball & stick models).

Current research findings have identified specific individual variables that have an impact on knowledge acquisition in STEM (Alexander, 2017 - for relational reasoning; Berkowitz et al., 2022 - for working memory capacity; Sorby et al., 2018 - for spatial ability; Edelsbrunner et al., 2023 - for representational competence). Simultaneously, other challenges in learning STEM are associated with the instructional approaches and methods utilized. For instance, Quinn et al. (2020) emphasizes the lack of integration of scientific inquiry practices into real-world learning, stating that the scientific phenomena lose its meaning to many students, as the way science is presented seems to be disconnected from their real-world experiences. Another challenge related to STEM learning is that some students experience great difficulty with theoretical representations of scientific concepts (Sahin & Yilmaz, 2020).

However, STEM education has emerged as a growing priority in global policy agendas (Tytler, 2020). The importance of STEM education has been explicitly highlighted by recent research findings, emphasizing that the number of STEM graduates is positively associated with GDP, employment, and labor productivity (Bacovic et al., 2022; Ray, 2015). Given the importance of boosting STEM education, it is crucial to tackle the challenges it poses, i.e., integrating the scientific concepts into real-life settings and supporting learners with the visual representation of the scientific phenomena. One way to do this is by using AR technology. In the next section, we will map out possible ways to approach it.

1.3. Affordances of augmented reality for STEM learning

One widely adopted definition of AR, developed by Azuma (1997), describes AR as the technology which simultaneously represents real and virtual objects in a 3D perspective within an authentic environment in real-time. Another characteristic is a proper alignment of real and virtual objects that achieves a seamless integration of virtual objects into the real world (Buchner et al., 2022).

Aligned with these defining characteristics, AR offers various affordances. Particularly in STEM education, AR enables multiple affordances that may support learning. Wu et al. (2013) mapped out AR affordances in their literature review. According to their research, AR environments present the following properties: depicting learning content in 3D perspectives, offering ubiquitous, collaborative and situated learning, providing learners with senses of presence, immediacy, and immersion, visualizing the invisible, and connecting formal and informal learning. Even though these affordances are not specific to AR environments alone (Krüger et al., 2019) and can also be applied to other immersive (e.g., VR) and traditional environments (e.g., classroom settings), we selected this classification of affordances because it captures the overall complexity and variability of what AR environments can offer.

The first affordance AR provides is a 3D perspective of the study object. In other words, the three-dimensional virtual object gets displayed in the user's real-world environment, granting the learner the opportunity to examine it from various angles. This affordance is particularly important when it comes to learning three-dimensional geometric shapes (Bhagat et al., 2021). The study revealed that the 3D depiction of these concepts leads to an increased learner's satisfaction and engagement.

The second affordance is the provision of situated, collaborative learning. Namely, the use of AR incentivizes the learners to work together (collaboratively and ubiquitously) and acquire knowledge in authentic and contextually relevant settings (situated learning). For instance, Tarng et al. (2015) developed a virtual butterfly ecological system by combining the campus host plants with the butterfly breeding activities displayed virtually by means of AR. Students could use smartphones or tablet PCs to "breed" virtual butterflies on the plants to further gain a deeper understanding of their life cycles and stages of growth.

The third affordance, according to Wu et al. (2013), is granting learners a sense of presence, immediacy, and immersion. Sense of presence experienced in AR stands for the capacity to interact effortlessly and naturally with all the genuine and mediated elements of the environment (Benyon, 2012). Immediacy refers to the quality of providing real-time and direct interaction or feedback between the virtual/augmented and real elements within the user's environment, thus affording instant and responsive interactions to contribute to a more immersive and engaging AR experience. Immersion is defined as vividness provided by a system, representing the system's capacity to exclude external factors (Cummings & Bailenson, 2016). All three aspects (the sense of presence, immediacy, and immersion) can positively affect the motivation and engagement students experience while learning (Cai et al., 2021).

The fourth affordance, which might be particularly relevant for STEM learning, is making the invisible visible. Revealing the unseen, AR enables learners to develop a deeper understanding of abstract concepts by effectively creating visual representations, such as diagrams or graphs, that can accurately shape the learner's understanding of elaborate science concepts. Yoon and Wang (2014) allowed the students to manipulate bar magnets in real-time and observe the visualized magnetic fields dynamically depicted on a computer screen. This visualization of the fields increased the students' interaction with the magnets and improved learners' engagement as well as collaboration.

Finally, connecting formal and informal learning, that is, extending traditional boundaries of learning from the school settings to everyday learning contexts, AR enables learning in various environments and shapes. An example of such a use of AR is the study of the effect of visualizing multiple representations to help convey basic concepts of current and resistance in a Science museum (Beheshti et al., 2017). In this study, parents and the kids who took part were offered to see the circuit visualized by means of AR. Not only did this intervention lead to the kids' improved performance in the post-test, but also stimulated an increased parental engagement in helping the kids understand the essence of current when examining the topic with AR. This change occurred due to the different nature of exploratory questions kids posed when "seeing" the current flow.

All the above-described affordances, that is 3D perspective, sense of presence, immediacy and immersion, collaborative and situated learning, making the invisible visible and bridging the formal and informal learning, emphasize the potential impact that AR can have on STEM learning. Previous research highlighted the advantages AR brings about in terms of the learning outcome in the following fields: physics (Akçayır et al., 2016; Tarng et al., 2022), chemistry (Chao et al., 2016), anatomy (Ferrer-Torregrosa et al., 2015), and medicine (Aebersold et al., 2018). Another encouraging finding of AR in STEM educational contexts is that most studies report positive effects on affective student characteristics, such as motivation and attitudes towards STEM subjects (Cao & Yu, 2023; Khan et al., 2019).

At the same time, studies more generally investigating differential effects of technology use also highlight the effect of certain learner characteristics on the learning outcome (Hofer & Reinhold, submitted). For instance, students with higher working memory capacity benefited more from computer-assisted instruction compared to those with lower working memory capacity (Chevalère et al., 2021). This might be attributed to the complexity of the computer-assisted environment, which can overload the learners with lower working memory capacity. When it comes to the use of VR, compared to high spatial ability students, low spatial ability students seem to profit more from learning with VR (Lee & Wong, 2014). Another recent study also revealed that students with lower prior knowledge gained a better learning outcome when being instructed with the help of VR in comparison to students with higher prior knowledge. However, this was true when a signaling principle (making the important information more prominent by means of highlighting the text, for instance) was used in the VR learning environment (Han et al., 2023). In other words, for learners with higher levels of prior knowledge, VR technology had no impact.

Given the affordances AR offers and thus, the novel learning environment this entails, it is crucial to systematically examine learner characteristics × AR affordances interactions explored in STEM education research.

1.4. Treatment in AR studies

In the classical ATI research, treatment is traditionally defined as an instructional method or environment applied to explore its interaction with learners' characteristics and its effects on the learning outcomes (Snow, 1977). In technology-enhanced learning, however, defining treatment becomes more complex. On the one hand, the use of technology, when compared to a control group (typically a traditional classroom setting), predefines the essence of intervention group treatment. On the other hand, technology alone, despite its unique affordances, cannot substitute for the absence or inadequacy of learning strategies in place. In other words, we cannot look at the treatment in the technology-supported condition as a mere use of technological affordances in the instruction.

Media comparison studies, i.e., studies that compare the effectiveness of one medium (e.g., AR, VR, videos) with another (e.g., traditional classroom methods, print materials) to determine which is better for learning outcomes, have faced criticism for comparing fundamentally different treatments, often overlooking the broader instructional design and pedagogical ground underlying the use of technology (Buchner & Kerres, 2023; Feldon et al., 2021). In their systematic review, Buchner and Kerres (2023) argue that 80 % of the existing AR research focuses on media comparison neglecting that solely technology does not drive learning. It is the alignment of the instructional methods with the learning objectives, learning environment and learning tools (e.g., traditional paper-based or technology-supported ones) that impact the "how" and the "why" of learning.

For this reason, our definition of treatment in AR research encompasses the multiple facets of learning with technology. We define treatment as an instructional intervention in a learning environment where specific instructional design that accounts for the environmental affordances and learning outcomes is used.

1.5. Learning outcomes

One of the most widely-used learning outcomes taxonomies developed by Bloom looks into the learning processes which best manifest the cognitive process the learner is involved in while studying for a specific learning outcome (Bloom, 1956; Forehand, 2010). Considering the complexity of the aims pursued in STEM education, we looked for the means to also integrate the type of knowledge aimed at, as well as the extent of knowledge complexity expected from the learner to master a new topic in our exploration of the learning outcomes.

To this end, we relied on the pragmatic instructional alignment taxonomy (PIAT) as a taxonomy to categorize learning outcomes in AR research (Hofer & Schalk, in preparation). The PIAT is a taxonomy that encompasses three characteristics of the learning outcome: The cognitive process, type of knowledge, and knowledge complexity. The first characteristic, the cognitive process learners are undertaking, describes whether to achieve a learning goal, learners have to reproduce what they have learned, whether and to which extent they have to transfer the learning content to apply it in new contexts or over time (Barnett & Ceci, 2002), and whether they have to produce something new based on what they have learned. In the context of the STEM disciplines, knowledge reproduction can be exemplified by memorization and reproduction of scientific facts, formulas, or equations. For knowledge transfer – application of the law of physics to explain the motion, and for knowledge production students apply theoretical concepts of Newton's laws to the practical experience in a pendulum experiment.

In terms of the type of knowledge, the PIAT taxonomy differentiates between declarative (factual and conceptual knowledge, i.e., facts and their relationship) and procedural knowledge (practical understanding of how to perform tasks and activities, involving the sequential steps or actions required for their execution). Knowledge complexity is represented by a high or low number of knowledge elements learners need to possess to achieve the learning outcome (e.g., clear indication of the necessity to have a certain extent of prior knowledge before the intervention).

Different cognitive processes, types of knowledge and complexity of knowledge manifested in different learning outcomes require different abilities of individuals to achieve these learning outcomes (Hofer & Schalk, in preparation). We therefore considered it crucial to embed the diversity of learning outcomes in the scope of our research considering individual differences in the use of AR in STEM.

1.6. Learning in technology-afforded environments triad: a study design framework

Given the multitude of affordances inherent in technology, along with the complexity of learning outcomes and individual characteristics, the interaction among these three elements - technological affordances, individual variables, and learning outcomes - is crucial for understanding intervention effectiveness in technology-afforded environments. This interaction emphasizes the specific affordances inherent to each technology, rather than the variation of specific instructional properties within the technology itself (e.g., using worked examples versus problem-solving approaches in virtual reality learning environments). We developed a study design framework called Learning in Technology-Afforded Environments Triad (L-Tech Triad) to underline the necessity of incorporating all the aspects in designing studies embracing technological affordances (Fig. 1).

It also pinpoints the need to consider how individual variables may vary depending on the learning content/outcome and technology under study. Therefore, careful consideration of which individual variables to analyze and specifying the corresponding learning outcomes addressed in each instance is crucial. The framework highlights the complexity of the "treatment" notion, i.e., the combination of instructional strategies, principles and the technological affordances, when it comes to learning in the technology-supported environments. In other words, it elaborates on the key facets of the treatment that should be considered in the study design.

1.7. Present study

In the recent years, there have been multiple systematic reviews and meta-analyses in the field of AR and STEM (see Chang et al., 2022 for a meta-analysis, Hidayat & Wardat, 2023 for AR in STEM; Ibáñez & Delgado-Kloss, 2018 on AR in Science Education; Xu et al., 2022 for meta-analysis on moderators of AR in science learning), which extensively cover the state of the art in research in AR. However, an examination of existing ATI studies in AR research, thus identification of interactions between learner characteristics and the treatment in the AR environment, has yet to be performed. At the same time, when it comes to learning success in STEM, there is strong evidence to suggest that learning strongly depends on individual differences of the learners (e.g., Alexander, 2017 – for relational reasoning; Berkowitz et al., 2022 – for working memory capacity; Sorby et al., 2018 - for spatial ability). However, these findings were derived from studies carried out in traditional classroom settings without AR usage.

In this systematic review, we attempt to build on and extend existing research syntheses to determine the extent to which individual differences have been explored in AR research and to clarify what is currently known (and not) about the role of learner characteristics in AR-supported learning in STEM education. Concentrating on the STEM fields allows for a more targeted synthesis of how individual learner differences affect learning outcomes in domains where AR affordances are particularly relevant.

We do this we first investigate the current state of the art regarding individual differences as predictors in AR research, and then narrow the list to those studies that examine the learner characteristics from an ATI perspective. We look into the following research questions:

RQ1: To what extent are individual characteristics investigated as predictors in AR research within STEM education?
RQ2: What is the state of the art in AR research in terms of considering AR affordances, learning outcomes, and individual differences from an ATI perspective?

Importantly, we focused on studies that included the control group (non-AR settings) as a baseline for a moderation analysis.

RQ3: What are the key design characteristics of these studies?

We synthesize our findings and the results of the selected studies on the role of individual differences in the learning effectiveness of AR in STEM subjects and the statistical tests used to further highlight the challenges associated with examining learner characteristics in AR-supported environments. From this synthesis, we derive theoretical implications and propose a research agenda for the future, anchored in the Learning in Technology-Afforded Environments Triad. This

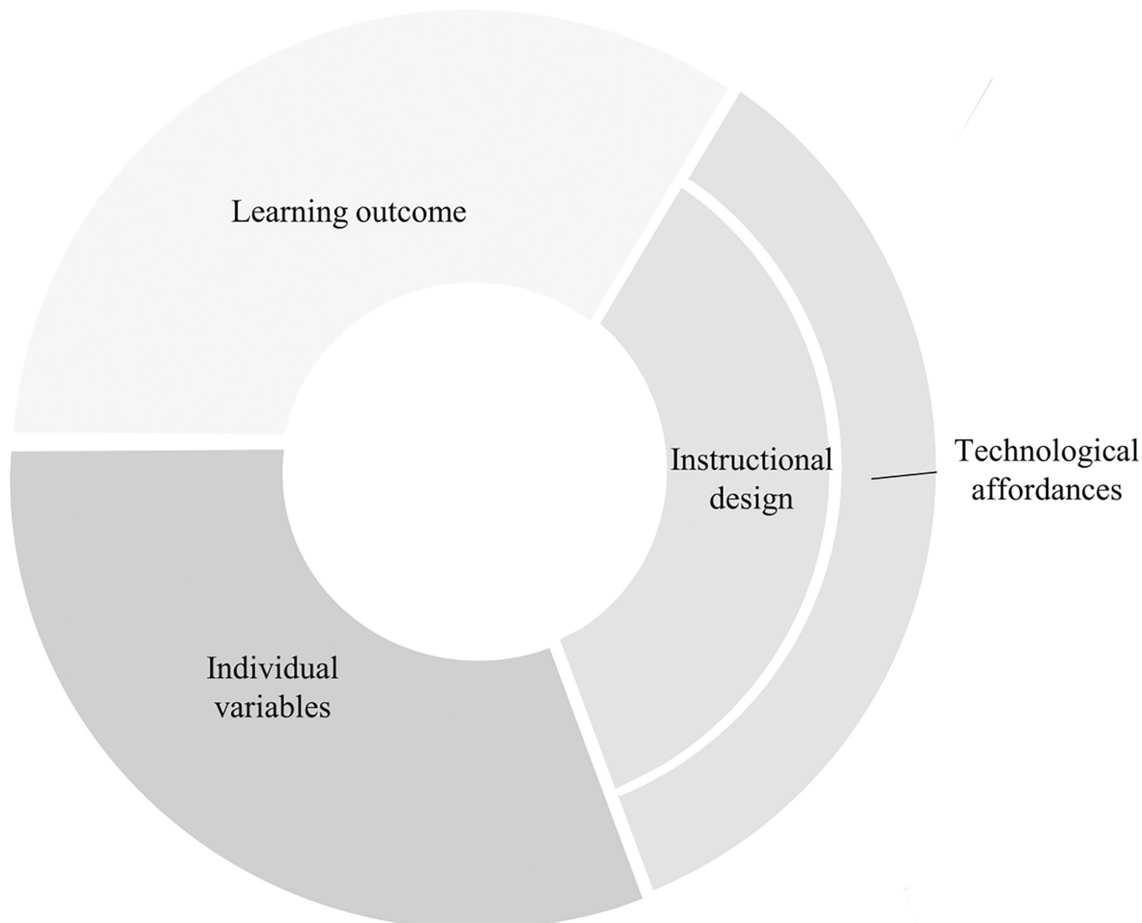

Fig. 1. Learning in Technology-Afforded Environments Triad (L-Tech Triad).
Note. This figure shows the three components of study design with the use of technology for STEM education.

framework is designed to guide researchers in developing AR-based studies, addressing the individual characteristics, instructional design guided by AR affordances and learning outcomes.

2. Methods

2.1. Search method

To obtain an in-depth overview of the current research, considering the latest technological advances, we focused on the last decade, including studies dating 2013–2022 in the sample. We considered this particular decade, as this review, in many ways, is an extension to the existing literature reviews and meta-analyses, which also focus on this time span, as the advances in the AR research had yet to be synthesized (Chang et al., 2022).

For our study selection, we utilized the search parameters and the findings of a meta-analysis by Chang et al. (2022) as in January 2023, when the first phase of the literature search was performed, it was the most recent and broad-ranged work reviewing AR research. In addition, Chang et al. (2022) also investigated various learning outcomes addressed with the affordances of AR alongside examining its application in a wide range of subject areas, including STEM. From this meta-analysis, we retrieved the studies published from 01.01.2013 to 31.12.2021 ($k = 134$). For the studies from 01.01.2022 to 31.12.2022, in line with the search string used by Chang et al. (2022), we utilized the same basic selection criteria and the study exclusion criteria in the initial phase. That is, we searched for the following terms “augmented reality”, “augmenting reality”, and “mixed reality”, combined with “learning”, “education”, “training”, “teaching”, and “instruction” in the two databases Web of Science and Scopus, which were the two databases utilized by Chang et al. (2022). The search was performed in January 2023 and yielded $k = 756$ results for the Web of Science and $k = 318$ for Scopus. After removing duplicates and conference proceedings, the remaining 498 titles were examined. These studies, together with the 134 studies derived from Chang et al. (2022) sample and three more studies found in other sources by means of snowballing, went through the two-phased selection procedure.

In the literature search and selection, we were guided by the procedure of the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA; Page et al., 2021). Fig. 2 provides more detail on how the selection procedure was implemented.

2.2. Selection procedure

In *Phase 1*, we imported the search results from the database search and other sources into Citavi v.6 (Swiss Academic Software GmbH, 2018), where the studies went through title and abstract screening. We then excluded 409 records based on the following exclusion criteria: written not in English; conference proceeding; a non-empirical study; pre-school students or teachers as a sample (while we focused on STEM education at schools); non-learning outcome addressed, i.e., motivation, enjoyment, self-efficacy etc. (since we aimed at inspecting, the effect AR has on the learning achievement in STEM); studies done in special education; vocational education; non-STEM subjects (e.g., art or English as a foreign language); conceptual paper (e.g., AR environment design is described or a possible use of AR in teaching, in general).

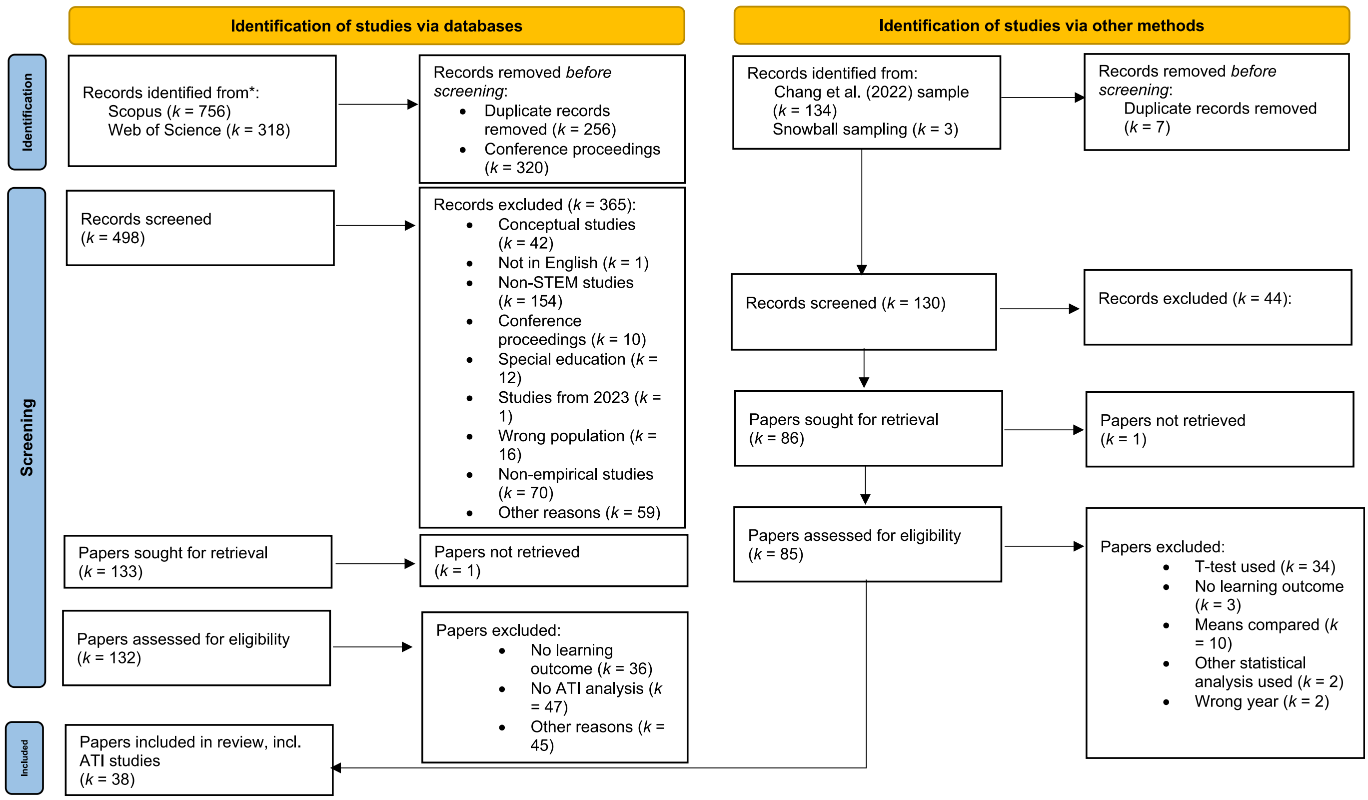

Fig. 2. Diagrammatic representation of the PRISMA review process.
Note. The PRISMA chart describes the process for the studies selection.

Coder 1 initially established the coding criteria. Subsequently, coder 2 conducted pilot testing with a subset of studies ($k = 25$) using this coding scheme. Then an assessment of interrater reliability followed. After performing consistency check, discussing the inconsistencies and reviewing discrepancies in the coding, this procedure was iterated three additional times with randomly-selected subsets of studies. After four rounds of pilot testing, an interrater reliability coefficient of Cohen's kappa 0.8 was achieved. Any inconsistencies were discussed until 100 % agreement was reached. At this point, coder 1 proceeded with the coding process. The same procedure was repeated for phases two and three of the studies' coding process.

In *Phase 2*, the remaining $k = 217$ documents were imported into the open-source CADIMA software for the full-text inspection (Kohl et al., 2018). We particularly looked into the research questions, hypotheses, and the statistical analysis used for the hypotheses testing.

In our sample at this phase, we included studies that examined individual characteristics as predictors. If the individual variable was used as a moderator, mediator, or as a covariate in the statistical methods used for the data analysis, the study was included in the sample. We also checked if the measurements used in the experiments indicated reliability and validity. The presence of a control group was another criterion for the study to be included. The total number of the studies at this phase (RQ1) amounted to $k = 38$ studies.

In the final selection phase (RQ2), we extracted only the studies which examined learners' individual characteristics as moderators according to the traditional ATI approach. The final sample for RQ2 and RQ 3 amounted to $k = 5$ studies.

2.3. Coding strategy

We coded the sample of $k = 38$ studies (RQ 1) based on the STEM subject in which the AR intervention was used, the specific individual characteristic analyzed, the test used to measure these characteristics, and the statistical analysis applied. For the coding of the sample for RQ 2 and 3 ($k = 5$), we additionally focused on the sample size, learning outcomes, and AR affordances. Table 1 provides a thorough overview of the key variables we coded based on existing frameworks.

2.3.1. Learner characteristics

Learner characteristics encompass a wide range of unique traits and individual characteristics, such as cognitive (e.g., working memory capacity, spatial ability, relational reasoning), affective (motivation, achievement emotions) and behavioral (self-regulation, metacognition). Initially, we aimed at narrowing down our examination to solely cognitive traits. However, given a scarce number of studies identified, we widened the scope to include all the three groups.

2.3.2. Statistical approach

It has been pointed out that in order to test the differential effectiveness of an educational approach such as AR-supported learning across individual characteristics, an appropriate statistical approach has to be used. Such analyses have to include an interaction term between the aptitude and the condition (e.g., 2×2 ANOVA or an interaction term between condition and aptitude in a regression analysis). If only one group of learners with AR is used and learning gains within this group are predicted by an individual characteristic of the learners, then this analysis does not allow examining whether the effect of the individual characteristic is caused by specific requirements of the AR setting, or whether it is just a more general predictor of learning (Tetzlaff et al., 2023). Therefore, a control condition has to be used in which AR or a specific implementation of it are not realized. The statistical analysis has to implement a moderation effect that shows whether, under the AR condition, the effect of the individual characteristic differs in comparison to its effect under the control condition. Such a moderation effect would indicate that an individual characteristic clearly depends on the AR condition and is not just a mere general learning prerequisite that is

Table 1
Coding scheme.

Code	Represented meaning	Research focus
Learner characteristics (Vandewaetere et al., 2011)	Learner characteristics are unique abilities and differences in cognitive, affective and behavioral domains that influence how the individual interacts with the learning environment	Research question 1 Research question 2 Research question 3
Statistical analyses (Preacher & Sterba, 2019)	Aptitude (learner characteristic) is treated as a predictor of the intervention effect	Research question 1 Research question 2 (Aptitude is treated as a moderator) Research question 3 (Aptitude is treated as a moderator)
Instructional function of AR	AR is used as a supplement or a substitute to the traditional learning material	Research question 2 Research question 3
Instructional design	Instructional methods, strategies and principles used to achieve the learning goals set	Research question 2 Research question 3
AR affordances (Wu et al., 2013)		Research question 2
Three-dimensional (3D) perspective	Enabling the learner to see the object from various angles	
Immediacy	Direct interaction with the environment with a feedback potential	
Situated learning	Authentic contexts or environments relevant to the content taught	
Making invisible visible	Visualizing normally invisible objects	
Connecting formal and informal learning	Integrating educational activities in everyday settings	
Learning outcomes (Hofer and Schalk, in preparation)		Research question 2
Cognitive processes Type of knowledge Knowledge complexity	Reproduce, transfer, produce Declarative or procedural	

not specific to AR (Tetzlaff et al., 2023). Consequently, we systematically coded the statistical analyses employed across the studies to ascertain whether aptitude was treated as a moderator (Preacher & Sterba, 2019). We also noted whether and in how many cases individual characteristics have been assessed but not been used in this manner (e.g., because they were treated as covariates but not as moderators).

2.3.3. *AR affordances*

For AR affordances in line with Wu et al. (2013), we identified the following subcategories: three-dimensional perspective, immediacy, making the invisible visible, situated learning, and connection formal and informal learning. Originally, Wu et al. (2013) proposed treating immersion and sense of presence alongside immediacy. However, after examining our sample, we employed only the "immediacy" affordance, as the immersion and sense of presence did not particularly apply to the context of the studies examined.

2.3.4. *Instructional design*

To avoid the "media comparison trap", in this review, we did not contrast technology- supported conditions with traditional settings per se, but rather highlighted how AR affordances are embedded in the instructional design. That is, we first examined which AR affordances were used in the technology-supported group, and how these affordances informed the instructional design, and the other way around. We deliberately avoided employing any particular framework for the instructional principles or methods, as we sought to explore the full capacity of how design is currently reflected, without restricting our investigation to any formal parameters.

2.3.5. *Instructional function of AR*

We also looked at the role of AR in instructional design. We wanted to find out whether AR was used as a supplement or as a replacement for the traditional learning material (e.g., a printed handout). In other words, we investigated whether AR complemented the learning material and was used alongside the book (e.g., AR provided the visualization of normally invisible magnetic fields in the horseshoe magnet in addition to the visualizations provided by the book) or replaced the traditional learning material and became the central medium of content delivery.

2.3.6. *Learning outcomes*

The PIAT taxonomy was selected as it properly corresponds to the real-life demands of instruction Hofer & Schalk, in preparation). Offering three dimensions to the goal-setting, this taxonomy examines the learning outcomes considering (1) cognitive processes (reproduce, transfer, and produce), (2) types of knowledge (procedural and declarative), and (3) complexity (number of knowledge elements, including the prior knowledge students need to engage with to achieve the learning outcome). This holistic approach to the learning outcome considers the practice-based nature of the studies. In other words, running a study, these are the exact three dimensions the learning outcome is centered around. Proximity to the real-world classroom makes the taxonomy practical for distinguishing the learning outcomes from the research, which is also practice-based. To code the cognitive processes, type of knowledge, and knowledge complexity, we looked into the intervention and the post-test used to check which type of knowledge and cognitive process were in focus of the experiment.

3. Results

3.1. *Research question 1*

Our first research question was to identify studies that addressed individual differences in AR research in STEM education. The key requirement for selecting the study was that the individual characteristic was considered as a predictor in the data analysis. Table 2 provides an insight into the studies identified at this stage.

Out of 38 studies identified, 29 studies were published in the last five years (from 2019 to 2024). The studies were performed in the following subject areas: physics, including astronomy ($k = 10$), biology, including anatomy ($k = 7$), chemistry ($k = 5$), mathematics, including geometry ($k = 3$), general science, including natural science ($k = 6$), geography ($k = 2$), programming ($k = 2$), archeology ($k = 1$), geology ($k = 1$) and medical science ($k = 1$).

In terms of the statistics employed, prior knowledge as a covariate in the statistical analysis (ANCOVA and MANOVA) was included in $k = 28$ studies. Three further studies also included individual characteristics as covariates, denoting them explicitly as predictors (which statistically is the same procedure as including covariates; Bhagat et al., 2021; Hu et al., 2021; Jackson et al., 2019). Other analyses used included fuzzy set qualitative comparative analysis (Ling et al., 2021) and a cluster analysis followed by a two-way ANCOVA (Yu et al., 2022).

3.1.1. *Prior knowledge as the most examined individual variable*

The prevailing individual characteristic tackled was prior knowledge ($k = 29$) measured by means of a self-designed pre-test ($k = 20$). Prior knowledge was used as a covariate in the statistical analyses ($k = 28$). Although the overall findings suggest that prior knowledge has an impact on learning outcomes, the direction of the effect of prior knowledge, the magnitude of this relationship and the effect of condition (control group vs. AR group) cannot be determined due to the lack of interaction analysis in the studies. With regard to the types of prior knowledge studied, Zumbach et al. (2022) differentiated between two types of prior knowledge examined - factual knowledge and

Table 2
General characteristics of the selected studies.

Study	Year	Subject	Individual variable	Tests	Statistics
Barmaki et al. (2019)	2019	Anatomy	Prior knowledge	Pre-test	ANCOVA 2 × 2 with pre-test score as covariate
Bhagat et al. (2019)	2019	Biology	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Bhagat et al. (2021)	2021	Geometry	Motivation, visual attention	Instructional Materials Motivation Survey (Keller, 2010); eye-tracking for visual attention including fixation count, total fixation duration, and total viewing duration	Regression analysis was used to explore the learning achievement predictors
Bogomolova et al. (2020)	2020	Anatomy	Visual-spatial ability	MRT (Peters et al., 1995), Paper Folding Test (PFT) (Ekstrom, 1976)	Linear regression analysis
Chao and Chang (2018)	2018	Mathematics	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Chao et al. (2016)	2016	Chemistry	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Chen (2020)	2020	Natural science	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Chen et al. (2022)	2022	Physics	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Chu et al. (2019)	2019	Architecture	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Elford et al. (2022)	2022	Chemistry	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Fidan and Tuncel (2019)	2019	Physics	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Furió Ferri et al. (2015)	2015	Science	Grade and gender	Self-report	ANCOVA with grade and gender as factors
Habig (2020)	2020	Chemistry	Gender, spatial ability	Self-report, Purdue Visualization of Rotation Test (Guay, 1976)	ANOVA with gender as a single factor and ANOVA with spatial ability as a covariate
Hsiao et al. (2013)	2013	Natural and Life Science and Technology	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Hu et al. (2021)	2021	Physics	Physics background, AR experience	Self-report	Multiple linear regression with AR experience and physics background as predictors
Huang et al. (2022)	2022	Natural science course	Gender	Self-report	Two-way ANOVA to examine the effects of learning method and gender, as well as their interaction, on posttest scores
Jackson et al. (2019)	2019	Geology	Gender, engagement, major, confidence, ethnicity, challenge	Self-designed questionnaire	Multiple regression analysis
Jones et al. (2022)	2022	Anatomy	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Kao and Ruan (2022)	2022	Programming	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Ke and Hsu (2015)	2015	Science-related TPACK knowledge	Survey of Pre-service Teachers' Knowledge of Teaching and Technology (Schmidt et al., 2009)	Pre-test	ANCOVA with pre-test score used as covariate
Lai et al. (2019)	2019	Geography	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Ling et al. (2021)	2021	Chemistry	Prior knowledge, attitude, SA	ROT (Purdue Spatial Visualisation revised); Unified Chemistry Final Exam (UCFE)	Fuzzy set qualitative comparative analysis
McNeal et al. (2020)	2020	Geography	Spatial ability (mental rotation)	PSVT:R-Revised Purdue Spatial Visualization Test (Guay, 1976; Maeda et al., 2013)	Linear regression analysis
Nagayo et al. Nagayo et al. (2022)	2022	Medical science (surgery)	Prior knowledge	Perform the suture twice in four minutes without referring to the AR system or video	ANCOVA with pre-test score used as covariate
Sun and Chen (2019)	2019	Mathematics	Prior knowledge	Midterm exam scores were used as the pre-test to be individuals' initial capability	ANCOVA with pre-test score used as covariate
Tarng et al. (2016)	2016	Physics	Prior knowledge	Lunar phase achievement test	ANCOVA with pre-test score used as covariate
Tarng et al. (2018)	2018	Physics	Prior knowledge	Sun pass achievement test	ANCOVA with pre-test score used as covariate
Tarng et al. (2022)	2022	Chemistry	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Thees et al. (2020)	2020	Physics	Prior knowledge	A selection of 13 single-choice items from the Heat and Temperature Concept Evaluation (HTCE) (Thornton & Sokoloff, 1998)	ANCOVA with pre-test score used as covariate
Thees et al. (2022)	2022	Physics	Prior knowledge	Ten items from a power test used by Altmeyer et al. (2020)	ANCOVA with pre-test score used as covariate

(*continued on next page*)

Table 2 (*continued*)

Study	Year	Subject	Individual variable	Tests	Statistics
Tsai and Lai (2022)	2022	Programming	Prior knowledge	Learning Effectiveness Questionnaire	ANCOVA with pre-test score used as covariate
Weng et al. (2019)	2019	Physics (astronomy)	Spatial ability (SA)	Visualization of Rotations (Revised PSVT: R) (Guay, 1980)	Two-way ANOVA followed by simple main effect tests.
Weng et al. (2020)	2020	Biology	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Wu et al. (2018)	2018	Natural science course	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Yoon et al. (2017)	2017	Physics (science museum visit)	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Yu et al. (2022)	2022	Physics	Learning physics (PA) anxiety, prior knowledge	Physics Anxiety Rating Scale (PARS) (Sahin et al., 2015), pre-test	Cluster analysis to dichotomize students into high PA and low PA followed by two- way ANCOVA with prior knowledge as a covariate
Zhang et al. (2020)	2020	Biology	Prior knowledge	Pre-test	ANCOVA with pre-test score used as covariate
Zumbach et al. (2022)	2022	Biology	Factual knowledge and mental representations knowledge pre-tests (self-designed), intrinsic and extrinsic motivation, metacognitive learning strategies as well as general and biology-related ability self-concept.	Pre-test, cognitive load (Klepsch et al., 2017), ability self-concept (Schöne et al., 2012), Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991), learning strategies (L IST questionnaire) (Wild & Schiefele, 1994)	MANCOVA with both knowledge pre-tests, intrinsic and extrinsic motivation, metacognitive learning strategies as well as general and biology-related ability self-concept used as covariates

representational competence - using two different instruments to measure those. Nagayo et al. (2022), who conducted the only study in medical science, assessed learners' prior knowledge by asking them to perform a suture (a stitch or row of stitches holding together the edges of a wound or surgical incision) twice in four minutes without reference to any supporting material. Ke and Hsu (2015) investigated the science-related technological pedagogical content knowledge (TPACK) of pre-service teachers to see whether it affected the learning outcome, i.e., improved the TPACK of teacher education students. With regard to the other studies, the type of prior knowledge was not specifically explored.

As for the other individual characteristics identified, spatial ability and gender were the second most commonly addressed ($k = 4$ for each characteristic). In terms of the facets of spatial ability, spatial visualization was predominantly investigated using the Mental Rotation Test (Peters et al., 1995) and the Purdue Visualization of Rotation Test (Guay, 1976). Other examined individual variables were cognitive load ($k = 1$) and motivation ($k = 2$).

3.1.2. Excluded studies

Having analyzed the final set and identified only a few studies that address the effects of the learners' characteristics on the learning outcomes in AR environments, we decided to gain a deeper understanding if the individual differences have been handled in any way in the remaining studies. After reexamining the 217 studies, we found out that 57 studies used *t*-tests to check the group differences in the learning outcome of the control and the intervention groups. An interesting example is the study by Habig (2020), where moderation analysis was conducted (gender as a moderator), but it did not encompass the crucial interaction term. In other words, although moderation was examined, the study did not investigate whether the effect of the independent variable on the dependent variable varied significantly at different levels of the moderator. Instead, the authors subjectively compared the effect of the learner characteristic on the learning outcome in the AR-condition to that in a control condition. Based on a significant effect in the AR condition but not in the control condition, the authors concluded that the role of the learner characteristic differed between the two conditions. However, this is a case of the "difference in significances"- fallacy (Edelsbrunner & Thurn, 2023): If an effect is significant in one group but not in the other, this does not yet indicate whether the effect reliably differed between the groups or conditions. To find this out, an interaction term in the sense of an ATI-analysis has to added and if this term is significant, it would support the assumption that the effect indeed differs between conditions. Among other analyses used were error pattern analysis (Cen et al., 2019), single factor covariate analysis (Chang et al., 2019) and principal component analysis (Mendez-Lopez et al., 2022). This look into statistical approaches in the excluded studies indicates that individual characteristics have been included in quite a few further studies but they are frequently not examined as moderator variables, which would be the appropriate approach to gauge their AR-specific effects.

3.2. Research questions 2 and 3

With regard to the second research question, out of 38 studies we sought to identify those that addressed individual characteristics from an ATI perspective, only five studies ($k = 5$) (<1 % of the initial hits) could be included. These were the studies that examined the effect of individual differences on learning outcomes in an AR environment. Although Huang et al. (2022) also examined the interaction between the individual variable (gender) and the learning condition, we deliberately excluded the study, since gender is a proxy for many individual characteristics that commonly vary across genders (e.g., cognitive abilities, interests and motivation, self-efficacy) and as such it may not be clearly categorized within our three kinds of variables.

The five studies from the final sample were published in the last five years, and they deal with physics ($k = 2$), anatomy ($k = 1$), geography ($k = 1$) and chemistry ($k = 1$) subjects. In terms of the intervention, AR was predominantly used to highlight the important features of the scientific concept or the object depicted in the AR environment, by making some parts visible or more prominent (e.g., Bogomolova et al., 2020; Yu et al., 2022) and allowing the students to inspect the object from various angles (e.g., McNeal et al., 2020; Weng et al., 2019). The key characteristics of the studies are summarized in Table 3.

As visible from Table 4, spatial ability is the most examined individual difference in AR research, being investigated in four studies (Bogomolova et al., 2020; Ling et al., 2021; McNeal et al., 2020; Weng et al., 2019). As learners' individual characteristics can be either compensated for or hyperbolized for a better learning outcome, for instance, when the level of learning material difficulty or the format is adapted to the learner's aptitude (Kühl et al., 2022), we found it crucial to examine the effect the spatial ability has on the learning outcome in AR. Interestingly, results from the reviewed studies indicate that spatial ability has a controversial impact on the learning outcome in STEM in AR. On the one hand, there is evidence that the negative effect of lacking spatial ability on the learning outcome in STEM is less pronounced in AR conditions in comparison to the traditional classroom (Weng et al.,

Table 3
General characteristics of the selected studies in ATI paradigm.

Study	Year	Country	Subject	Topic	Intervention
Bogomolova et al.	2020	Netherlands	Anatomy	Lower limb anatomy	3D virtual model of a lower leg with an object-centered view to enable exploration, visual and audio feedback on the structures, size adjustment.
Ling et al.	2021	Taiwan	Chemistry	Molecular structure and properties of organic compounds, mainly including alkanes (e.g., methane), alkenes (e.g., ethylene), alkynes (e.g., acetylene), and aromatic hydrocarbons (e.g., benzene).	To a two-dimensional picture the corresponding 3D structure would be displayed on the tablet, and the 3D model could rotate freely with the movement of the two-dimensional picture. Then compare different molecule models with each other.
McNeal et al.	2020	USA	Geography	Topographic Maps Assessment	Manipulation of sand within a physical sandbox, while seeing real-time visual projections, such as contour lines and color-coded elevation representations, augmented to generate a dynamic 3D topographic map.
Weng et al.	2019	Indonesia	Astronomy	Solar and lunar eclipse	A book with 3D pictures of the eclipse.
Yu et al.	2022	China	Physics	Magnetism	Students scan real magnets with a mobile device, superimposing virtual magnet models and explore magnetic attraction and magnetic field patterns.

2019). However, the findings by McNeal et al. (2020) revealed that compared to the control group, low spatial performers did not do as well on the post-test after using the AR Sandbox. The study by Bogomolova et al. (2020) also revealed only a marginal effect of the intervention × spatial ability interaction. In other words, it is yet to be examined whether students with lower spatial ability really benefit from the use of AR in instruction more strongly than those with higher spatial ability.

In terms of the learning outcomes pursued, all five of the studies aimed at developing declarative knowledge, considering the high complexity of the knowledge chunks under discussion (Ling et al., 2021; Yu et al., 2022). As for the cognitive processes, the five studies sought to get the students to reproduce and transfer the knowledge gained, for instance, by enabling them to demonstrate comprehension in a post-knowledge test. Weng et al. (2019) also used knowledge transfer by inviting the students to create an AR book on solar and lunar eclipse after examining which, the students could answer a more analytical question of how solar and lunar eclipses happen. None of the studies sought to achieve the “produce” learning goal though. Three studies explicitly described the learning outcomes pursued in the experiment (Bogomolova et al., 2020; Weng et al., 2019; Yu et al., 2022).

3.2.1. Instructional design

In terms of instructional design, most of the studies provide a detailed overview of how exactly and why the AR technology was used in the instructional design to deliver the specific learning content. Namely, the authors explained the rationale for using AR for the specific content, the learning methods used, and the difference between the control and experimental groups in terms of instructional design was clearly stated. In all five studies, the learning methods used for the control group were identical to those used for the AR group. For instance, Yu et al. (2022) used active experimentation in both the control and experimental groups. However, in the latter, objects generated by AR were used for experimentation. In three studies (Ling et al., 2021; McNeal et al., 2020; Weng et al., 2019), AR was used as a supplement to the learning materials used. That is, AR served as an additional medium of instruction due to its unique affordances. For example, in Ling et al. (2021) when learning about the molecule structure, in the AR group students the physical model of a molecule, used in the control group an additional more sophisticated AR model. Two studies (Bogomolova et al., 2020; Yu et al., 2022) used AR as a substitute for the traditional learning tool. For instance, Bogomolova et al. (2020) used AR to help students memorize the names of bones and muscles, the location and organization of their structures, and the function of muscles based on their origin. A traditional 2D atlas and a 3D desktop environment were replaced by AR, which allowed students to move around the 3D model of a lower leg to explore it from different angles and perspectives.

3.2.2. Individual differences

Spatial ability is the most extensively researched learner characteristic in ATI - AR studies (Bogomolova et al., 2020; Ling et al., 2021; McNeal et al., 2020; Weng et al., 2019). For instance, Bogomolova et al. (2020) found that learners with high mental rotation ability (one of the facets of spatial ability) performed equally well across all conditions, suggesting that AR does not provide a significant advantage for individuals with strong visual-spatial skills. In contrast, learners with lower mental rotation ability benefited significantly from AR, outperforming their counterparts in the monoscopic 3D desktop condition and achieving similar scores to those in the 2D anatomical atlas group. Similarly, Weng et al. (2019) reported that students with higher spatial ability generally demonstrated strong learning outcomes regardless of instructional medium, whereas those with lower spatial ability benefited more from AR. The AR tool might have helped bridge the performance gap by compensating for weaker spatial visualization skills, ultimately reducing disparities between the students with high and low spatial ability. However, McNeal et al. (2020) found that students with high mental rotation ability consistently outperformed their peers across all conditions, with structured AR activities yielding the highest scores. In contrast, students with low mental rotation ability struggled in AR Sandbox activities, particularly in the semi-structured condition, where their performance was even lower than in the control group. Although structured AR activities offered the best results among AR treatments for these students, the AR group as a whole did not significantly outperform the control group.

Ling et al. (2021) identified distinct learner profiles in relation to various learning outcomes - immediate and knowledge retention in AR-supported environments. The learner profiles emerged from the qualitative comparative analysis rather than being predefined. The study categorized learners and applied necessity and sufficiency analyses to identify the individual variables combinations that lead to good or poor learning outcomes. The study suggests that there is no single trait universally essential for learning with AR. The necessity and sufficiency analyses revealed four distinct profiles associated with both positive and negative learning outcomes for immediate and lasting learning outcomes. Ling et al. (2021) identified four distinct learner profiles based on their characteristics and how they interacted with AR:

- Strong foundational learning and high spatial ability: These learners generally achieved good immediate learning outcomes and required AR support to maintain long-term retention.
- Weak foundational learning and high spatial ability: Without AR, these learners had poor learning outcomes. However, when AR was used, they could achieve good immediate learning outcomes, but only if they had a passionate attitude towards AR could they sustain good long-lasting outcome.

Table 4
AR affordances, learning outcomes and individual differences from an ATI perspective.

Study	Learning goal			AR affordance	Instructional design	Instructional function of AR	Individual difference	Tests	Role of individual difference
	Cognitive process	Type of knowledge	Complexity						
Bogomolova et al., 2020	Reproduce	Declarative	Low	Making invisible visible, 3D perspective, immediacy	Constructive alignment theory was applied to ensure alignment between learning goals, activities, and assessments in all the three groups: AR group, Monoscopic 3D Desktop Group and 2D Anatomical Atlas Group. In AR group, participants interacted with DynamicAnatomy, an AR application for HoloLens, featuring a fully interactive 3D model of the lower leg. Monoscopic 3D Desktop Group: A Windows desktop application with the same anatomical model as DynamicAnatomy but displayed on a 2D screen was used. Interaction was limited to rotating the model along the Y-axis using a mouse. Other features, such as auditory feedback and scaling, were included. 2D Anatomical Atlas Group: Used selected handouts from anatomy atlases and textbooks with 2D images of bones, muscles, and ankle movements. Handouts included an index for navigation but only listed anatomical names, with no additional descriptive content.	Walking around the 3D model to explore the structure of the limb from various perspectives with AR used as a substitute for the traditional learning material.	Visual-spatial ability	MRT (Peters et al., 1995), Paper Folding Test (PFT) (Ekstrom, 1976)	MRT-High Group: Performed equally well across all three groups, indicating no significant advantage of AR for participants with high visual-spatial ability. MRT-Low Group: AR Group significantly outperformed the monoscopic 3D desktop group and achieved similar scores to the 2D anatomical atlas group. Monoscopic 3D Desktop Group scored significantly lower compared to other groups.2D Anatomical Atlas Group achieved better results than the desktop group, with a moderate effect size.
Ling et al., 2021	Reproduce, transfer	Declarative	High	3D perspective, making invisible visible, immediacy	Active learning principles, e.g., comparing several molecule models and collaborative activities were used in both groups. Students observed either only a physical model of a molecule (control group) or the physical model supplemented by an additional more sophisticated AR model (in AR	Depicting the molecule in 3D from various angles. AR is used as a supplement to the traditional learning material (ball-and-stick model of a molecule).	Prior knowledge, attitude, spatial ability	ROT (Purdue Spatial Visualisation revised); Unified Chemistry Final Exam (UCFE)	Learning success: Solid foundation for learning (prior knowledge): Students with a solid foundational knowledge (SFL) benefited from AR because it enhanced their ability to visualize and understand molecular structures,

(continued on next page)

Table 4 (*continued*)

Study	Learning goal			AR affordance	Instructional design	Instructional function of AR	Individual difference	Tests	Role of individual difference
	Cognitive process	Type of knowledge	Complexity						
					group). The observation was followed by a group discussion and independent examination of the molecule models to examine their structure.				improving both immediate (GILO) and long-term learning outcomes (GLLO).AR provided complementary advantages by linking abstract concepts to concrete visualizations, especially for those with higher prior knowledge. High spatial ability (HSA) enabled students to effectively perceive and understand 3D spatial structures using AR. AR further amplified their natural abilities by providing manipulable models, resulting in improved immediate and lasting outcomes. Even students with low foundational knowledge (SFL) but high HSA could achieve GILO and GLLO when supported by AR. Learning failure: Poor Foundational Knowledge (Prior Knowledge): Students with weak foundational knowledge (~SFL) struggled with AR, as they lacked the prior knowledge necessary to understand and connect the visualizations to concepts. AR alone could not compensate for the lack of basic understanding, leading to poor immediate (GILO) and lasting outcomes (GLLO).
McNeal et al., 2020	Reproduce, transfer	Declarative	Low	3D perspective,	In the AR group the integration of	AR-based models of common molecules	Spatial ability	PSVT: R-Revised	Students with high mental

(continued on next page)

Table 4 (*continued*)

Study	Learning goal			AR affordance	Instructional design	Instructional function of AR	Individual difference	Tests	Role of individual difference
	Cognitive process	Type of knowledge	Complexity						
				making invisible visible, immediacy	structured, semi-structured, and unstructured learning activities was used. In the unstructured activities students independently explored the AR Sandbox with minimal guidance. In semi-structured, students received moderate guidance, such as prompts for contour-line rules and relationships between water flow and elevation. In the structured AR group, detailed activities included topographic mapping, slope analysis, and creating landscape profiles under the instructor's direction. In the control group students followed a more linear instructional approach using traditional tools and lab materials. Activities were guided primarily by worksheets without the interactive and visual aids provided by the AR Sandbox. Assignments without AR integration.	were used, which allowed for simultaneous comparison for several molecules at the same time. AR used as a supplement to the traditional learning material (traditional topographic map).	(mental rotation)	Purdue Spatial Visualization Test (Guay, 1976; Maeda et al., 2013)	rotation (MR) ability showed better topographic map assessment score across all conditions, with structured AR activities yielding the highest results. Students with low MR ability struggled in AR Sandbox activities, particularly in the semi-structured condition, which resulted in lower scores than the control group. Structured AR activities provided the highest scores among AR treatments for these students but the AR group did not significantly outperform the control group.
Weng et al., 2019	Reproduce, transfer	Declarative	Low	Situated learning, 3D perspective, making invisible visible	In the AR/MR group active learning principles, e.g., manipulating 3D models and visualizations of celestial movements in the MR group were used. In the control group the students learnt more passively from the book only.	Depicting the eclipse from a 3D perspective. MR is used as a supplement integrated in the traditional learning materials (a printed book).	Spatial ability (SA)	Visualization of Rotations (Revised PSVT: R) (Guay, 1980)	Higher spatial ability (HSA) students showed strong learning outcomes, but the impact of AR was less pronounced compared to lower spatial ability (LSA). LSA students benefited more from the AR tool, as the 3D visualizations and MR features compensated for their lower spatial visualization skills. The gap in learning outcomes between HSA and LSA students was smaller in the AR group

(continued on next page)

Table 4 (*continued*)

Study	Learning goal			AR affordance	Instructional design	Instructional function of AR	Individual difference	Tests	Role of individual difference
	Cognitive process	Type of knowledge	Complexity						
									compared to the control group, demonstrating AR's role in reducing disparities.
Yu et al., 2022	Reproduce, transfer	Declarative	High	Making invisible visible	Active experimentation with either the real physical objects or the augmented models. In the AR group, hands-on experiments were guided through the MagAR modules during the three lessons: Introduction to the Magnetic World: Interaction with virtual objects to explore magnetic attraction and repulsion. Magnetic field inquiry: visualization of iron filings' distribution and magnetic induction lines using AR. Knowledge extension & recall: interaction sub-modules with AR markers were used to explore geomagnetic fields, conclude knowledge, and practice magnetic field drawing. In the control group, students learned using traditional experimental materials during the three lessons. The experiments included observing magnetic field lines by spreading iron filings on paper over magnets, interacting with physical magnets, small magnetic needles, and objects.	Visualizing magnetic fields with AR used a substitute for the traditional experimentational materials.	Learning physics anxiety	Physics Anxiety Rating Scale (PARS) (Sahin et al., 2015)	Students with higher physics anxiety performed significantly better in the AR group compared to the non-AR group. Students with lower physics anxiety also performed better in the AR group compared to the non-AR group.

- Strong foundational learning and low spatial ability: Learners in this profile could achieve good immediate leaning outcomes but struggled to maintain good knowledge retention unless they had passionate attitude for AR. If they had a negative attitude towards AR, their learning outcomes also deteriorated.
- Weak foundational learning and low spatial ability: These learners consistently struggled to achieve poor immediate and long-term outcomes unless they had a passionate attitude towards AR. However, their engagement was often driven by novelty rather than meaningful learning, that made AR less effective for them.

As for the final study, Yu et al. (2022) found that both high and low physics anxiety students benefited from AR-supported learning, with significantly better performance in the AR group compared to the non-AR group. However, the positive impact of AR was particularly notable for students with higher physics anxiety. This reveals the potential of AR to support the learners with higher physics anxiety who may otherwise struggle with anxiety-related issues in learning.

3.2.3. *AR affordances*

As for the AR affordances, all five studies stipulated making the invisible visible as the key property of the AR technology for their research purposes. In addition, 3D perspective was also extracted as another environmental feature suitable for the selected studies (Bogomolova et al., 2020; Ling et al., 2021; Weng et al., 2019). In their study, Ling et al. (2021) also utilized the immediate feedback the AR affords (we coded it as "immediacy"). That is, the students could get the 3D structured model of the molecule and its components by pointing the tablet camera to the two-dimensional picture. By timely seeing the molecule components, the students could analyze the structure more thoroughly and connect their new knowledge chunks to the existing ones.

3.2.4. *Key design characteristics of the reviewed studies*

The third research question focused on the key design characteristics of these studies. The key findings are condensed in Table 5.

Table 5
Key study design characteristics.

Study	Groups	Experimental design	Sample size	Statistics
Bogomolova et al., 2020	AR group, Monoscopic 3D Desktop Group, 2D Anatomical Atlas Group	Double-center randomized controlled trial	$N = 58$	Linear regression analysis, with anatomy knowledge test score as dependent variable, intervention group as a fixed factor, visual-spatial abilities test score as a covariate, and "visual-spatial abilities test score" × "intervention group" as in interaction term
Ling et al., 2021	AR group vs. control group	Quasi-experimental study, between-subject design	$N = 97$	Fuzzy set qualitative comparative analysis
McNeal et al., 2020	AR group (structured learning activities) vs. AR group (unstructured learning activities) vs. AR group (semi-structured learning activities) vs. control group	Quasi-experimental study, between-subject design	$N = 545$	Linear regression analysis
Weng et al., 2019	AR group vs. control group	True between-subjects experimental research design	$N = 80$	Two-way ANOVA followed by simple main effect tests
Yu et al., 2022	AR group vs. control group	A 2 × 2 factorial quasi-experimental design	$N = 96$	Cluster analysis to dichotomize students into high PA and low PA followed by two- way ANCOVA

3.2.5. *Study designs, groups and samples*

In accordance with the ATI-approach, all the five studies used between subject designs, where control group performance was compared to the experimental group (AR vs. non-AR condition). In the reviewed studies, three employed a traditional two-group experimental design, comparing learning outcomes between a technology-supported group and a control group using paper-and-pencil learning material (Ling et al., 2021; Weng et al., 2019; Yu et al., 2022). Bogomolova et al. (2020) utilized a three-group experimental design, where the AR group was compared to both a 3D monoscopic version of the AR application implemented on a desktop and a 2D atlas condition, allowing for an assessment of added value of AR beyond other technological and traditional learning materials. McNeal et al. (2020) implemented a four-group experimental design, in which three AR groups with varying instructional designs - structured, semi-structured, and unstructured - were compared against a control group to examine the impact of different instructional approaches within AR environments on learning outcomes.

In terms of the sample size, the studies were conducted with rather small sample sizes, from 58 to 97 students, apart from McNeal et al. (2020), where 545 participants took part in the study, distributed between four conditions.

3.2.6. *Statistical analyses*

As to the statistical analyses employed for the learner differences and AR environment interaction, one study used ANOVA, one - ANCOVA and two studies used regression analysis. Ling et al. (2021) employed fuzzy set qualitative comparative analysis, aimed at handling situations where variables are not strictly dichotomous (either present or absent) but can exist in degrees of membership. This can provide a more in-depth understanding of the relationship between the independent and dependent variables beyond strict dichotomies that, for example, result from median splits (Kumar et al., 2022).

4. Discussion

The aim of the present systematic literature review was to examine the state of the art of AR research in terms of considering the individual differences in general as predictors and specifically from an ATI perspective together with the learning outcomes, AR affordances and instructional design. This paper adds to the existing body of research (meta-analyses and literature syntheses) that deals with the effect of AR on learning effectiveness (Chang et al., 2022; Hidayat & Wardat, 2023; Ibáñez & Delgado-Kloos, 2018). Out of the 217 studies selected <20 % look into individual characteristics (38 studies). Prior knowledge is examined in 29 studies. Even though over 70 % of the identified studies consider the effect of the prior knowledge when learning with AR, their findings are rather uninformative. The statistical analyses implemented treat prior knowledge as a covariate. That means that the direction of the relationship (e.g., whether higher prior knowledge enhances learning outcomes) or the magnitude of prior knowledge impact (e.g., how strongly prior knowledge influences learning outcomes) in the AR settings are not examined. As for the types of prior knowledge, only three studies describe the exact types of prior knowledge addressed in the test. This scarcity of findings makes it impossible to draw any definite conclusion on the role of the prior knowledge type in learning with AR.

Our findings also revealed that limited attention is currently given to individual differences from an ATI perspective in AR research. This was evidenced by only five included studies out of 217. Four of the five studies looked into the effect of spatial ability on the learning outcome in an AR environment (Bogomolova et al., 2020; Ling et al., 2021; McNeal et al., 2020; Weng et al., 2019). The other individual variables examined were prior knowledge and physics anxiety (Ling et al., 2021; Yu et al., 2022). Declarative knowledge reproduction and transfer were the most common learning outcomes addressed. Making the invisible visible (e.g., Bogomolova et al., 2020; McNeal et al., 2020: Yu et al., 2022) and 3D perspective (Bogomolova et al., 2020; Ling et al., 2021; McNeal et al., 2020; Weng et al., 2019) were the most utilized AR affordances, which goes in line with our expectations. Overall, the number of studies examining learner characteristics is small, and among those studies spatial ability has been the most frequently examined characteristic. On the one hand, the findings align in suggesting that AR is particularly valuable for learners with lower spatial ability (Bogomolova et al., 2020; Weng et al., 2019). However, sharp differences in study design make it difficult to draw final conclusions. For instance, Weng et al. (2019) conducted a study with fifth-grade students in astronomy, while Bogomolova et al. (2020) performed their study with first- and second-year undergraduate students. Moreover, Bogomolova et al. (2020), even though examining the same facet of spatial ability, i.e., mental rotation, employed a different instrument to measure it. At the same time, McNeal et al. (2020), who examined the effect of AR on understanding topographic maps in the population of college students, revealed that in their settings the students with low mental rotation ability struggled in AR Sandbox activities. To sum up, the differences in learning content, participant demographics (e.g., educational level) and different facets of spatial ability impede generalization and a deeper understanding of the role of spatial ability when learning with AR. The complexity of the results reported by Ling et al. (2021) further highlights the necessity of a more nuanced approach to investigating the role of individual differences in learning with AR. Namely, the study underscores the importance of considering multiple cognitive characteristics simultaneously rather than focusing on isolated variables. According to Ling et al. (2021) learning outcomes the treatment aims to achieve should also be considered.

In the next section we will attempt to map out the possible implications of the findings and how those could be properly exploited in AR research in the future. As we believe that the inclusion of learner characteristics in AR research from an ATI perspective is the most robust way to examine the impact of AR on individual learners, in the next section we will focus on the ways in which the ATI perspective can be integrated into AR studies.

4.1. Embracing individual differences in AR research: possible solutions

We believe there are several reasons that could account for a relatively low number of studies considering individual differences in AR studies. Firstly, this might be attributed to the methodological challenges that the traditional ATI analyses pose. The inclusion of interaction parameters in the analysis is a key characteristic of the traditional ATI approach. This necessitates larger sample sizes compared to the analysis of the main effect (Blake & Gangestad, 2020; Perugini et al., 2018; Tetzlaff et al., 2023). Relatively small sample sizes ($N < 100$) used in AR studies impede the utilization of such analyses. At the same time, the use of AR, being a novel technology, is associated with technological limitations. These limitations include the absence of required infrastructure such as high-speed internet at school, connectivity issues affecting the integration of AR technology into real-world experimental setups used in STEM studies, and the high costs of AR technology along with the technical skills needed for its proper implementation. Consequently, AR technology has yet to achieve mass adoption to be utilized at a grander scale in educational research. For this reason, the limited sample sizes will likely persist as an enduring issue for individual differences research in AR studies. We will attempt to examine potential ways to overcome these methodological challenges further in the paper.

Another factor that could explain a relatively small number of AR studies looking into the individual characteristics might be aligned with the keen focus on the effectiveness of AR technology in teaching STEM per se. That means scrutinizing whether AR interventions have an impact on learning effectiveness in STEM disciplines at a broader level. Such keen interest in the role of AR technology in learning can be explained by the novelty of this technology. However, as the learning intervention can exercise heterogeneous impact on different learners (Tetzlaff et al., 2021), we believe that considering individual differences at the phase of the foundational exploration of AR effectiveness in learning could have a big influence on promoting a more individualized approach to instructional effectiveness research. This gap could be addressed by increasing awareness of the necessity for ATI research within the AR community. This, for instance, could entail providing ideas for the types of individual differences to examine, suggesting appropriate tests for assessing these differences, and outlining methodological approaches for incorporating ATI perspectives into AR studies as well as a study design framework that will encompass the individual variables. In the next section, we attempt to ideate the prospective research agenda for further elaboration of individual differences in AR research.

4.1.1. Statistically modeling the affordances of AR in terms of individual differences

The first gap suggested was related to methodological challenges associated with the analysis of individual differences in the effects of AR in STEM. We see this gap as consisting of two issues. First, most of the reviewed studies used small sample sizes. Based on the recommendation of Cronbach and Snow (1981) to have at least $N = 100$ learners in each condition in ATI research, all but one studies in our review have to be considered small-sample studies (i.e., with sample sizes below this number). The reason for ATI being so sample size-intensive is that the critical analytical step in such analyses is the estimation of the moderation effect that represents the interaction between the individual characteristic × AR condition in focus in comparison to another learning condition. It is well-known that moderation effects typically require between four to eight times as large samples as analyses of simple main effects (e.g., condition comparisons; Blake & Gangestad, 2020). If sample sizes are smaller, statistical power is low, which undermines finding effects reliably but also being able to confirm null effects (Blake & Gangestad, 2020).

In the case of a limited sample size, various statistical options are available to improve statistical power. For example, Bayesian estimation of moderation effects can increase statistical power through introducing theoretical knowledge of reasonable effect sizes into specifications of parameter prior distributions (McCarthy & Masters, 2005; Van de Schoot et al., 2014). Another option is to use analytic approaches that bridge the gap between qualitative and quantitative approaches in small samples, such as fuzzy set qualitative comparative analysis (fsQCA; Ragin et al., 2006). Qualitative comparative analysis, in general, unites the qualitative and quantitative paradigm divide (Marx et al., 2014). fsQCA, being the most advanced of the comparative qualitative analyses, which can be utilized for smaller sample sizes (down to $N < 50$), is also notable for recognizing non-linear relationships between variables and accommodating asymmetry in the data (Geremew et al., 2024). This means that fsQCA can be used when the data exhibit asymmetry or skewness. Another characteristic, particularly useful for individual differences research is that fsQCA assigns each case the individual membership scores that can be partial (any score from 0 to 1), signifying that the individual characteristics are more nuanced and can be “moderately present” or “scarcely present”. These characteristics coupled with the fsQCA robustness for smaller sample sizes could make this analysis a perfect solution to tackle the methodological challenge related to the smaller sample sizes. Another prominent characteristic of fsQCA is that it treats the multitude of individual differences alongside the various learning outcomes that occur, which helps to manifest the complexity of the educational context learner characteristics interaction (Ling et al., 2021; Reinhold et al., 2020).

Another similar approach that has been introduced by Tetzlaff et al. (2023) is based on latent profile analysis. In this approach, learners' profiles across multiple individual characteristics are modelled and then their interactions with different conditions are examined. By reducing the information from multiple variables into one profile variable, this approach can also increase statistical power and help find interactions between multiple learner characteristics and instructional conditions.

4.1.2. Broadening the variety of individual differences considered in AR research

Another gap was the limited exploration of the variety of individual differences that could be considered in AR studies. To support the inclusion of a larger variety of individual differences variables relevant to STEM learning in future research, we assembled Table 6. This table provides an overview of variables that have been shown to affect STEM learning. In addition, the table provides suggestions for measures that might be used to assess the respective variables. The present table could be used for reference in AR study design to make an informed decision on which exact individual variable to include, the rationale for its inclusion and the possible measurement instruments to use.

Table 6
List of individual differences to measure in AR research.

Individual difference	Possible measurement	Why include in a study
Inductive reasoning (IR)	Raven's Progressive Matrices (Raven, 2003)	IR plays a crucial role in predicting STEM performance in schools (Stender et al., 2018; Venville & Oliver, 2015).
Relational reasoning	Test of relational reasoning (TORR) (Alexander et al., 2016)	Relational reasoning is a fundamental cognitive skill that underlies STEM performance (Alexander, 2017).
Prior knowledge	n/a due to its domain-specificity	Prior knowledge is key to predicting learning outcomes (Simonsmeier et al., 2022). However, to advance the understanding of how prior knowledge influences learning outcomes, it is essential to differentiate between and describe knowledge types in prior knowledge assessment. This involves distinguishing between and describing declarative and procedural knowledge, for instance, as well as specifying the content of different knowledge elements (Hofer & Schalk (in preparation)
Representational competence	STEM topic-specific (Küchemann et al., 2021 RCFI – for vector fields; Klein et al., 2017 – KiRC for kinematics)	The ability to use and interpret various representations in science has an impact on conceptual knowledge development (Rau, 2017).
Working memory capacity	Complex-span tasks, updating tasks, and binding tasks (Wilhelm et al., 2013)	Working memory capacity is strongly associated with math performance, which in turn is key to STEM achievement (Berkowitz et al., 2022).
Spatial ability	Mental rotations test (Peters et al., 1995) Visualization of Rotations (Revised PSVT: R) (Guay, 1980)	Substantial evidence highlighting the importance of spatial ability within STEM (Buckley et al., 2018 – for systematic literature review).
Motivational-affective variables (e.g., self-concept, interest, engagement, self-regulation etc.)	Motivated Strategies for Learning Questionnaire (MSLQ) (Garcia & Pintrich, 1996)	Math and science motivational beliefs are positively associated with STEM achievement (Jiang et al., 2020).

4.2. Learning in technology-afforded environments triad for AR

By studying individual differences as predictors, we can gain insight into how learners' characteristics, such as spatial ability and working memory capacity, influence learning within specific learning environments or with certain teaching methods. However, focusing on learner characteristics from an ATI perspective grants a deeper understanding of how the effectiveness of learning environments and teaching methods (as compared to control conditions) varies between different learners. Adding the learner characteristics in the AR learning effectiveness research would also allow to identify the learners for whom AR might be less effective - or even detrimental.

Since our analysis revealed that even the few available findings offer limited opportunities to derive consistent evidence regarding the role of the individual characteristics due to the difference in the study design and measurements used, we believe a more unified approach to incorporating individual differences in the AR research should be adopted. This way researchers can accumulate stronger evidence, helping to uncover existing "blind spots" and new opportunities for optimizing AR-based instruction.

Therefore, drawing insights from this systematic review, we adapted the Learning in Technology-Afforded Environments Triad to the AR context (L-Tech Triad for AR), emphasizing the importance of integrating an ATI perspective in study design. That means, that alongside including the four parameters, we expanded on earlier in this work - AR affordances, instructional design that considers the environmental affordances, learning outcomes, and individual variables - in the study design, it is also essential to consider the ATI methodological requirements. That is, the presence of the control group in the study design, a sufficiently large sample size within the study for an adequate statistical power or statistical models that can improve statistical power (Ling et al., 2021; McCarthy & Masters, 2005; Reinhold et al., 2020) are required to further conduct interaction analysis.

This framework (Fig. 3) is intended to serve as a guide for future research endeavors aimed at examining the learning effects of AR technology incorporating an ATI perspective, thus enabling a more robust examination of the educational potential of AR.

4.3. Limitations

There are some limitations to note. The literature review does not include studies focusing on comparing AR with other media (e.g., VR). This is partly attributed to the focus of the review, i.e., taking a classical ATI perspective on the AR studies, that presupposes the presence of a control group and the intervention group (in our case, AR). The individual differences examined in the between-media comparison studies might provide interesting questions for future research. The second limitation is strict compliance to the traditional ATI analysis approach, where individual differences are regarded as moderators. This focus brought us to overlook the potential roles of individual differences as mediators or covariates. Gaining an understanding of the role of individual differences as mediators or covariates could help develop other insights of the impact learner characteristics have on learning with AR, but as discussed in detail, without moderation analysis in experimental-control group comparisons, these approaches do not allow extracting effects of learner characteristics that are specific to AR. In addition, we focused on the role of individual differences in AR research specifically within STEM fields. Consequently, the studies done in the other fields such as foreign language learning, art and special education, were excluded which may have led to the omission of some relevant contributions in ATI-AR research.

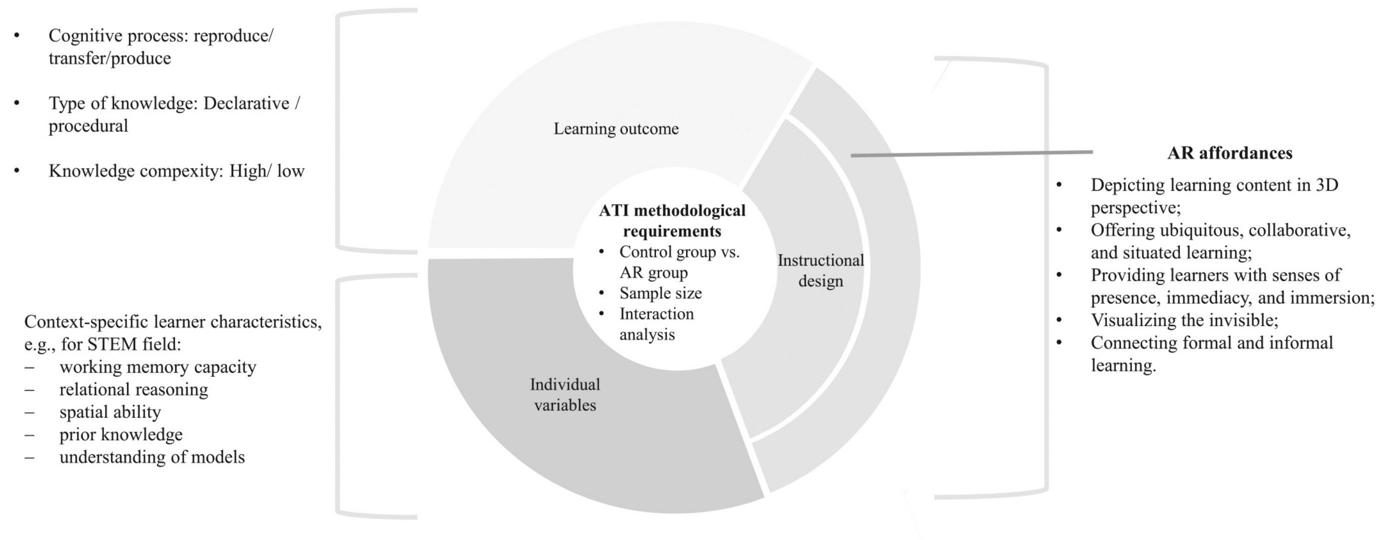

Fig. 3. Learning in Technology-Afforded Environments Triad: a study design framework for AR environments.
Note. This figure shows the four components of study design with the use of AR for STEM education (L-Tech Triad for AR).

5. Conclusion

In light of the ever-growing interest in STEM disciplines, the demonstrated efficacy of AR in cultivating positive attitudes towards its integration may serve as a potent catalyst for fostering more enjoyment of STEM learning among students (Wang et al., 2023; Yu et al., 2017). We have synthesized a decade of research on consideration of individual differences, AR affordances and learning outcomes in AR research carried out in STEM. Our findings suggest that at present, there is little attention to the role of the learners' characteristics from an ATI perspective. However, our systematic review also maps out possible solutions to promote a wider inclusion of the learners' characteristics in AR research. The insights gained may support the emergence of more interest and expansion of the research focus from the overall effectiveness of AR for learning to the individual learner, their aptitudes, whether these can be reinforced or compensated for by means of employing AR technology in the learning process or whether the use of AR can have a detrimental effect on learning on the individual level. As a result, the findings of this review may inform future research agendas and guide the design of studies in the STEM field from an ATI perspective, utilizing the L-Tech Triad for technology-afforded environments.

Educational relevance

The current systematic review, to the best of our knowledge, is the first attempt to look into the extent to which differential effectiveness depending on learners' characteristics is considered in research into the use of Augmented Reality (AR) in STEM studies. The study maps out the exemplary methodological practices and proposes the Learning in Technology-Afforded Environments Triad (L-Tech Triad) to further guide the study design in STEM with the use of AR.

Credit authorship contribution statement

Zoya Kozlova: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Katharina M. Bach:** Writing – review & editing, Validation, Formal analysis. **Peter A. Edelsbrunner:** Writing – review & editing, Formal analysis. **Sarah I. Hofer:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgements

We would like to thank Prof. Chang, Dr. Binali and their team for their assistance in the initial phases of the paper selection procedure.

References

Aebersold, M., Voepel-Lewis, T., Cherara, L., Weber, M., Khouri, C., Levine, R., & Tait, A. R. (2018). Interactive anatomy-augmented virtual simulation training. *Clinical Simulation in Nursing, 15*, 34–41. https://doi.org/10.1016/j.ecns.2017.09.008

Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior, 57*, 334–342. https://doi.org/10.1016/j.chb.2015.12.054

Alexander, P. A. (2017). Relational reasoning in STEM domains: A foundation for academic development. *Educational Psychology Review, 29*, 1–10. https://doi.org/10.1007/s10648-016-9383-1

Alexander, P. A., Dumas, D., Grossnickle, E. M., List, A., & Firetto, C. M. (2016). Measuring relational reasoning. *The Journal of Experimental Education, 84*(1), 119–151. https://doi.org/10.1080/00220973.2014.963216

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., & Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses - Theoretical background and empirical results. *British Journal of Educational Technology, 51*(2), 611–628. https://doi.org/10.1111/bjet.12900

Azuma, R. T. (1997). A survey of augmented reality. *Presence Teleoperators and Virtual Environments, 6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Bacovic, M., Andrijasevic, Z., & Pejovic, B. (2022). STEM education and growth in Europe. *Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-021-00817-7

Barmaki, R., Yu, K., Pearlman, R., Shingles, R., Bork, F., Osgood, G. M., & Navab, N. (2019). Enhancement of anatomical education using augmented reality: An empirical study of body painting. *Anatomical Sciences Education, 12*(6), 599–609. https://doi.org/10.1002/ase.1858

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612. https://doi.org/10.1037/0033-2909.128.4.612

Beheshti, E., Kim, D., Ecanow, G., & Horn, M. S. (2017, May). Looking inside the wires: Understanding museum visitor learning with an augmented circuit exhibit. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1583–1594).

Benhamdi, S., Babouri, A., & Chiky, R. (2017). Personalized recommender system for e-Learning environment. *Education and Information Technologies, 22*, 1455–1477. https://doi.org/10.1007/s10639-016-9504-y

Benyon, D. (2012). Presence in blended spaces. *Interacting with Computers, 24*(4), 219–226. https://doi.org/10.1016/j.intcom.2012.04.005

Berkowitz, M., Edelsbrunner, P., & Stern, E. (2022). The relation between working memory and mathematics performance among students in math-intensive STEM programs. *Intelligence, 92*, 101649. https://doi.org/10.1016/j.intell.2022.101649

Bernacki, M. L., Greene, M. J., & Lobczowski, N. G. (2021). A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose(s)? *Educational Psychology Review, 33*(4), 1675–1715. https://doi.org/10.1007/s10648-021-09615-8

Bhagat, K. K., Liou, W. K., Spector, J. M., & Chang, C. Y. (2019). To use augmented reality or not in formative assessment: A comparative study. *Interactive Learning Environments, 27*(5–6), 830–840. https://doi.org/10.1080/10494820.2018.1489857

Bhagat, K. K., Yang, F. Y., Cheng, C. H., Zhang, Y., & Liou, W. K. (2021). Tracking the process and motivation of math learning with augmented reality. *Educational Technology Research and Development, 69*(6), 3153–3178. https://doi.org/10.1007/s11423-021-10066-9

Bhutoria, A. (2022). Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a human-in-the-loop model. *Computers and Education: Artificial Intelligence, 3*, Article 100068. https://doi.org/10.1016/j.caeai.2022.100068

Blake, K. R., & Gangestad, S. (2020). On attenuated interactions, measurement error, and statistical power: Guidelines for social and personality psychologists. *Personality and Social Psychology Bulletin, 46*(12), 1702–1711. https://doi.org/10.1177/014616722091336

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: Handbook II*. David McKay.

Bogomolova, K., van der Ham, I. J. M., Dankbaar, M. E. W., van den Broek, W. W., Hovius, S. E. R., van der Hage, J. A., & Hierck, B. P. (2020). The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: A double-center randomized controlled trial. *Anatomical Sciences Education, 13*(5), 558–567. https://doi.org/10.1002/ase.1941

Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning, 38*(1), 285–303. https://doi.org/10.1111/jcal.12553

Buchner, J., & Kerres, M. (2023). Media comparison studies dominate comparative research on augmented reality in education. *Computers & Education, 195*, Article 104711. https://doi.org/10.1016/j.compedu.2022.104711

Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review, 30*(3), 947–972. https://doi.org/10.1007/s10648-018-9432-z

Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J. C. (2021). Effects of learning physics using Augmented Reality on students' self-efficacy and conceptions of learning. *British Journal of Educational Technology, 52*(1), 235–251. https://doi.org/10.1111/bjet.13020

Cao, W., & Yu, Z. (2023). The impact of augmented reality on student attitudes, motivation, and learning achievements - A meta-analysis (2016–2023). *Humanities and Social Sciences Communications, 10*(1), 1–12. https://doi.org/10.1057/s41599-023-01852-2

Cen, L., Ruta, D., Al Qassem, L. M. M. S., & Ng, J. (2019). Augmented immersive reality (AIR) for improved learning performance: A quantitative evaluation. *IEEE Transactions on Learning Technologies, 13*(2), 283–296. https://doi.org/10.1109/TLT.2019.2937525

Chang, H. Y., Binali, T., Liang, J. C., Chiou, G. L., Cheng, K. H., Lee, S. W. Y., & Tsai, C. C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-) experimental studies to investigate the impact. *Computers in Education*. https://doi.org/10.1016/j.compedu.2022.104641

Chang, Y. S., Hu, K. J., Chiang, C. W., & Lugmayr, A. (2019). Applying mobile augmented reality (AR) to teach interior design students in layout plans: Evaluation of learning effectiveness based on the ARCS model of learning motivation theory. *Sensors, 20*(1), 105. https://doi.org/10.3390/s20010105

Chao, J., Chiu, J. L., DeJaegher, C. J., & Pan, E. A. (2016). Sensor-augmented virtual labs: Using physical interactions with science simulations to promote understanding of gas behavior. *Journal of Science Education and Technology, 25*(1), 16–33. https://doi.org/10.1007/s10956-015-9574-4

Chao, W. H., & Chang, R. C. (2018). Using augmented reality to enhance and engage students in learning mathematics. *Advances in Social Sciences Research Journal, 5*(12), 455–464. https://doi.org/10.14738/assrj.512.5900

Chen, C.-H. (2020). Impacts of augmented reality and a digital game on students' science learning with reflection prompts in multimedia learning. *Educational Technology Research and Development, 68*(6), 3057–3076. https://doi.org/10.1007/s11423-020-09834-w

Chen, C.-H., Chan, W.-P., Huang, K., & Liao, C.-W. (2022). Supporting informal science learning with metacognitive scaffolding and augmented reality: Effects on science knowledge, intrinsic motivation, and cognitive load. *Research in Science & Technological Education, 41*(4), 1480–1495. https://doi.org/10.1080/02635143.2022.2032629

Chevalère, J., Cazenave, L., Berthon, M., Martinez, R., Mazenod, V., Borion, M. C., … Huguet, P. (2021). Computer-assisted instruction versus inquiry-based learning: The importance of working memory capacity. *PLoS One, 16*(11), Article e0259664. https://doi.org/10.1371/journal.pone.0259664

Chu, H. C., Chen, J. M., Hwang, G. J., & Chen, T. W. (2019). Effects of formative assessment in an augmented reality approach to conducting ubiquitous learning activities for architecture courses. *Universal Access in the Information Society, 18*, 221–230. https://doi.org/10.1007/s10209-017-0588-y

Cronbach, L. J., & Snow, R. E. (1981). Aptitudes and instructional methods: A handbook for research on interactions. *Ardent Media*.

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740

Edelsbrunner, P. A., Malone, S., Hofer, S. I., Küchemann, S., Kuhn, J., Schmid, R., … Lichtenberger, A. (2023). The relation of representational competence and conceptual knowledge in female and male undergraduates. *International Journal of STEM Education, 10*(1), 44. https://doi.org/10.1186/s40594-023-00307-3

Edelsbrunner, P. A., & Thurn, C. M. (2023). Improving the utility of non-significant results for educational research: A review and recommendations. *Educational Research Review, 100590*. https://doi.org/10.1016/j.edurev.2023.100590

Ekstrom, R. B. (1976). *Kit of factor-referenced cognitive tests*. Educational Testing Service.

Elford, D., Lancaster, S. J., & Jones, G. A. (2022). Exploring the effect of augmented reality on cognitive load, attitude, spatial ability, and stereochemical perception. *Journal of Science Education and Technology, 31*(3), 322–339. https://doi.org/10.1007/s10956-022-09957-0

Feldon, D. F., Jeong, S., & Clark, R. E. (2021). Fifteen common but questionable principles of multimedia learning. In R. E. Mayer, & L. Fiorella (Eds.), *The cambridge handbook of multimedia learning* (3rd ed., pp. 25–40). Cambridge University Press. https://doi.org/10.1017/9781108894333.005.

Ferrer-Torregrosa, J., Torralba, J., Jimenez, M. A., García, S., & Barcia, J. M. (2015). ARBOOK: Development and assessment of a tool based on augmented reality for anatomy. *Journal of Science Education and Technology, 24*(1), 119–124. https://doi.org/10.1007/s10956-014-9526-4

Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem-based learning: The effects on learning achievement and attitude in physics education. *Computers & Education, 142*, Article 103635. https://doi.org/10.1016/j.compedu.2019.103635

Forehand, M. (2010). Bloom's taxonomy. *Emerging Perspectives on Learning, Teaching, and Technology, 41*(4), 47–56.

Furió Ferri, D., Juan, M., Segui, I., & Vivó Hernando, R. A. (2015). Mobile learning vs. traditional classroom lessons: A comparative study. *Journal of Computer Assisted Learning, 31*(3), 189–201. https://doi.org/10.1111/jcal.12071

Garcia, T., & Pintrich, P. R. (1996). Assessing students' motivation and learning strategies in the classroom context: The motivated strategies for learning questionnaire. In M. Birenbaum, & F. J. R. C. Dochy (Eds.), *Evaluation in education and human services: Vol. 42. Alternatives in assessment of achievements, learning processes and prior knowledge*. Dordrecht: Springer. https://doi.org/10.1007/978-94-011-0657-3_12.

Geremew, Y. M., Huang, W. J., & Hung, K. (2024). Fuzzy-set qualitative comparative analysis as a mixed-method and analysis technique: A comprehensive systematic review. *Journal of Travel Research, 63*(1), 3–26. https://doi.org/10.1177/00472875231168619

Guay, R. (1976). *Purdue spatial visualization test*. Educational Testing Service.

Guay, R. B. (1980). *Spatial ability measurement: A critique and an alternative*.

Habig, S. (2020). Who can benefit from augmented reality in chemistry? Sex differences in solving stereochemistry problems using augmented reality. *British Journal of Educational Technology, 51*(3), 629–644. https://doi.org/10.1111/bjet.12891

Han, J., Liu, G., & Zheng, Q. (2023). Prior knowledge as a moderator between signaling and learning performance in immersive virtual reality laboratories. *Frontiers in Psychology, 14*, Article 1118174. https://doi.org/10.3389/fpsyg.2023.1118174

Hidayat, R., & Wardat, Y. (2023). A systematic review of augmented reality in science, technology, engineering and mathematics education. *Education and Information Technologies, 29*(8), 1–26. https://doi.org/10.1007/s10639-023-12157-x

Hofer, S. I. & Schalk, L. (in preparation). Facilitating instructional planning and constructive alignment: The pragmatic instructional alignment taxonomy (PIAT).

Hofer, S. I., & Reinhold, F. (submitted) Scaffolding of Instructional Strategies: ATI Effects in. Maths?.

Hsiao, H.-S., Chang, C.-S., Lin, C.-Y., & Wang, Y.-Z. (2013). Weather observers: A manipulative augmented reality system for weather simulations at home, in the classroom, and at a museum. *Interactive Learning Environments, 24*(1), 205–223. https://doi.org/10.1080/10494820.2013.834829

Hu, X., Goh, Y. M., & Lin, A. (2021). Educational impact of an augmented reality (AR) application for teaching structural systems to non-engineering students. *Advanced Engineering Informatics, 50*, Article 101436. https://doi.org/10.1016/j.aei.2021.101436

Huang, C. Y., Chou, Y. Y., Chen, C. H., & Tsai, Y. H. (2022). Applying activity system-based process model in augmented reality-based learning for natural science course in elementary school. *Mobile Information Systems, 2022*(1), Article 9579766. https://doi.org/10.1155/2022/9579766

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education, 123*, 109–123. https://doi.org/10.1016/j.compedu.2018.05.002

Jackson, D., Kaveh, H., Victoria, J., Walker, A., & Bursztyn, N. (2019). Integrating an augmented reality sandbox challenge activity into a large-enrollment introductory geoscience lab for nonmajors produces no learning gains. *Journal of Geoscience Education, 67*(3), 237–248. https://doi.org/10.1080/10899995.2019.1583786

Jiang, S., Simpkins, S. D., & Eccles, J. S. (2020). Individuals' math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Developmental Psychology, 56*(11), 2137–2150. https://doi.org/10.1037/dev0001110

Jones, C., Khalil, D., Mander, K., Yeoh, A., & Moro, C. (2022). Providing dementia education with augmented reality: A health sciences and medicine feasibility pilot study. *Research in Learning Technology, 30*, 2668. https://doi.org/10.25304/rlt.v30.2668

Kao, G. Y. M., & Ruan, C. A. (2022). Designing and evaluating a high interactive augmented reality system for programming learning. *Computers in Human Behavior, 132*, Article 107245. https://doi.org/10.1016/j.chb.2022.107245

Ke, F., & Hsu, Y. C. (2015). Mobile augmented-reality artifact creation as a component of mobile computer-supported collaborative learning. *The Internet and Higher Education, 26*, 33–41. https://doi.org/10.1016/j.iheduc.2015.04.003

Kell, H. J. (2018). Noncognitive proponents' conflation of "cognitive skills" and "cognition" and its implications. *Personality and Individual Differences, 134*, 25–32. https://doi.org/10.1016/j.paid.2018.05.025

Keller, J. M. (2010). *Motivational Design for Learning and Performance: The ARCS model approach*. Springer: New York, NY. https://doi.org/10.1007/978-1-4419-1250-3

Khan, T., Johnston, K., & Ophoff, J. (2019). The impact of an augmented reality application on learning motivation of students. *Advances in Human-Computer Interaction, 2019*. https://doi.org/10.1155/2019/7208494

Klein, P., Müller, A., & Kuhn, J. (2017). Assessment of representational competence in kinematics. *Physical Review Physics Education Research, 13*(1), Article 010132. https://doi.org/10.1103/PhysRevPhysEducRes.13.010132

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and Germane cognitive load. *Frontiers in Psychology, 8*(1997). https://doi.org/10.3389/fpsyg.2017.01997

Kohl, C., McIntosh, E. J., Unger, S., Haddaway, N. R., Kecke, S., Schiemann, J., & Wilhelm, R. (2018). Online tools supporting the conduct and reporting of systematic reviews and systematic maps: A case study on CADIMA and review of existing tools. *Environmental Evidence, 7*, 1–17. https://doi.org/10.1186/s13750-018-0115-5

Krüger, J. M., Buchholz, A., & Bodemer, D. (2019, December). Augmented reality in education: Three unique characteristics from a user's perspective. In *Proc. 27th Int. Conf. on Comput. in Educ* (pp. 412–422). https://doi.org/10.58459/icce.2019.481

Krüger, J. M., Palzer, K., & Bodemer, D. (2022). Learning with augmented reality: Impact of dimensionality and spatial abilities. *Computers and Education Open, 3*, Article 100065. https://doi.org/10.1016/j.caeo.2021.100065

Küchemann, S., Malone, S., Edelsbrunner, P., Lichtenberger, A., Stern, E., Schumacher, R., … Kuhn, J. (2021). Inventory for the assessment of representational competence of vector fields. *Physical Review Physics Education Research, 17*(2), Article 020126. https://doi.org/10.1103/PhysRevPhysEducRes.17.020126

Kühl, T., Fehringer, B. C., & Münzer, S. (2022). Unifying the ability-as-compensator and ability-as-enhancer hypotheses. *Educational Psychology Review, 34*(2), 1063–1095. https://doi.org/10.1007/s10648-021-09650-5

Kumar, S., Sahoo, S., Lim, W. M., Kraus, S., & Bamel, U. (2022). Fuzzy-set qualitative comparative analysis (fsQCA) in business and management research: A contemporary overview. *Technological Forecasting and Social Change, 178*, Article 121599. https://doi.org/10.1016/j.techfore.2022.121599

Lai, A. F., Chen, C. H., & Lee, G. Y. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology, 50*(1), 232–247. https://doi.org/10.1111/bjet.12716

Lee, E. A. L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education, 79*, 49–58. https://doi.org/10.1016/j.compedu.2014.07.010

Ling, Y., Zhu, P., & Yu, J. (2021). Which types of learners are suitable for augmented reality? A fuzzy set analysis of learning outcomes configurations from the perspective of individual differences. *Educational Technology Research and Development, 69*, 2985–3008. https://doi.org/10.1007/s11423-021-10050-3

Maeda, Y., Yoon, S. Y., Kim-Kang, G., & Imbrie, P. K. (2013). Psychometric properties of the revised PSVT: R for measuring first year engineering students' spatial ability. *International Journal of Engineering Education, 29*(3), 763–776.

Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. *Multimedia Tools and Applications, 83*(6), 1–49. https://doi.org/10.1007/s11042-023-15986-7

Marx, A., Rihoux, B., & Ragin, C. (2014). The origins, develop[1]ment, and application of qualitative comparative analysis: The first 25 years. *European Political Science Review, 6*(1), 115–142. https://doi.org/10.1017/s1755773912000318

McCarthy, M. A., & Masters, P. I. P. (2005). Profiting from prior information in bayesian analyses of ecological data. *Journal of Applied Ecology, 42*(6), 1012–1019. https://doi.org/10.1111/j.1365-2664.2005.01101.x

McNeal, K. S., Ryker, K., Whitmeyer, S., Giorgis, S., Atkins, R., LaDue, N., … Pingel, T. (2020). A multi-institutional study of inquiry-based lab activities using the Augmented Reality Sandbox: Impacts on undergraduate student learning. *Journal of Geography in Higher Education, 44*(1), 85–107. https://doi.org/10.1080/03098265.2019.1694875

Mendez-Lopez, M., Juan, M. C., Molla, R., & Fidalgo, C. (2022). Evaluation of an augmented reality application for learning neuroanatomy in psychology. *Anatomical Sciences Education, 15*(3), 535–551. https://doi.org/10.1002/ase.2089

Murphy, K. R. (2012). Individual differences. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 48–67). New York, NY: Oxford University Press.

Nagayo, Y., Saito, T., & Oyama, H. (2022). Augmented reality self-training system for suturing in open surgery: A randomized controlled trial. *International Journal of Surgery, 102*, Article 106650. https://doi.org/10.1016/j.ijsu.2022.106650

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education, 71*, 56–76. https://doi.org/10.1016/j.compedu.2013.09.011

Organisation for Economic Co-operation and Development. (2006). *Personalising education*. OECD Publishing.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Medicine, 18*(3), Article e1003583. https://doi.org/10.1371/journal.pmed.1003583

Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *Revue Internationale de Psychologie Sociale, 31*(1), 1–23.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition, 28*(1), 39–58. https://doi.org/10.1006/brcg.1995.1032

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. University of Michigan.

Preacher, K. J., & Sterba, S. K. (2019). Aptitude-by-treatment interactions in research on educational interventions. *Exceptional Children, 85*(2), 248–264. https://doi.org/10.1177/0014402918802803

Quinn, C. M., Reid, J. W., & Gardner, G. E. (2020). S+ T+ M= E as a convergent model for the nature of STEM. *Science & Education, 29*, 881–898. https://doi.org/10.1007/s11191-020-00130-w

Ragin, C. C., Drass, K. A., & Davey, S. (2006). *Fuzzy-set/qualitative comparative analysis 2.0* (Vol. 23(6), pp. 1949–1955). Tucson, Arizona: Department of Sociology, University of Arizona.

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review, 29*(4), 717–761. https://doi.org/10.1007/s10648-016-9365-3

Raven, J. (2003). Raven progressive matrices. In *Handbook of nonverbal assessment* (pp. 223–237). Boston, MA: Springer US.

Ray, R. (2015). *STEM education and economic performance in the American states*. Munich: University Library of Munich (MPRA Paper 65517).

Reinhold, F., Strohmaier, A., Hoch, S., Reiss, K., Böheim, R., & Seidel, T. (2020). Process data from electronic textbooks indicate students' classroom engagement. *Learning and Individual Differences, 83*, Article 101934. https://doi.org/10.1016/j.lindif.2020.101934

Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology, 102*(3), 254. https://doi.org/10.1037/apl0000151

Sahin, D., & Yilmaz, R. M. (2020). The effect of Augmented Reality Technology on middle school students' achievements and attitudes towards science education. *Computers & Education, 144*, Article 103710. https://doi.org/10.1016/j.compedu.2019.103710

Sahin, M., Caliskan, S., & Dilek, U. (2015). *Physics Anxiety Rating Scale (PARS) [database record]. APA PsycTests*. https://doi.org/10.1037/t51406-000

Schmidt, D., Baran, E., Thompson, A., Koehler, M., Mishra, P., & Shin, T. (2009, March). Examining preservice teachers' development of technological pedagogical content knowledge in an introductory instructional technology course. In *Society for Information Technology & Teacher Education International Conference* (pp. 4145–4151). Association for the Advancement of Computing in Education (AACE).

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelter, J. (2012). *Skalen zur erfassung des schulischen selbstkonzepts: SESSKO*. Hogrefe.

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education, 8*(4), 485–503. https://doi.org/10.1007/s40692-021-00188-7

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist, 57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700

Snow, R. E. (1977). Individual differences and instructional theory. *Educational Researcher, 6*(10), 11–15.

Snow, R. E. (1980). Aptitude, learner control, and adaptive instruction. *Educational Psychologist, 15*(3), 151–158.

Sorby, S., Veurink, N., & Streiner, S. (2018). Does spatial skills instruction improve STEM outcomes? The answer is 'yes'. *Learning and Individual Differences, 67*, 209–222. https://doi.org/10.1016/j.lindif.2018.09.001

Spector, J. M. (2014). Emerging educational technologies: Tensions and synergy. *Journal of King Saud University, Computer and Information Sciences, 26*(1), 5–10. https://doi.org/10.1016/j.jksuci.2013.10.009

Spector, J. M. (2018). The potential of smart technologies for learning and instruction. *International Journal of Smart Technology & Learning, 1*(1), 21–32. https://doi.org/10.1504/IJSMARTTL.2016.078163

Stender, A., Schwichow, M., Zimmerman, C., & Härtig, H. (2018). Making inquiry-based science learning visible: The influence of CVS and cognitive skills on content knowledge learning in guided inquiry. *International Journal of Science Education, 40*(15), 1812–1831. https://doi.org/10.1080/09500693.2018.1504346

Sun, K. T., & Chen, M. H. (2019). Utilizing free augmented reality app for learning geometry at elementary school in Taiwan: Take volumetric measurement of compound body for example. *International Journal of Distance Education Technologies, 17*(4), 36–53. https://doi.org/10.4018/IJDET.2019100103

Swiss Academic Software GmbH. (2018). Citavi (version 6). https://www.citavi.com (computer software).

Tarng, W., Lin, Y. S., Lin, C. P., & Ou, K. L. (2016). Development of a lunar-phase observation system based on augmented reality and mobile learning technologies. *Mobile Information Systems, 2016*(1), Article 8352791. https://doi.org/10.1155/2016/8352791

Tarng, W., Ou, K. L., Lu, Y. C., Shih, Y. S., & Liou, H. H. (2018). A sun path observation system based on augment reality and mobile learning. *Mobile Information Systems, 2018*(1), Article 5950732. https://doi.org/10.1155/2016/8352791

Tarng, W., Ou, K. L., Yu, C. S., Liou, F. L., & Liou, H. H. (2015). Development of a virtual butterfly ecological system based on augmented reality and mobile learning technologies. *Virtual Reality, 19*(3), 253–266. https://doi.org/10.1007/s10055-015-0265-5

Tarng, W., Tseng, Y. C., & Ou, K. L. (2022). Application of augmented reality for learning material structures and chemical equilibrium in high school chemistry. *Systems, 10*(5), 141. https://doi.org/10.3390/systems10050141

Tetzlaff, L., Edelsbrunner, P., Schmitterer, A., Hartmann, U., & Brod, G. (2023). Modeling interactions between multivariate learner characteristics and interventions: A person-centered approach. *Educational Psychology Review, 35*(4), 112. https://doi.org/10.1007/s10648-021-09674-8

Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing personalized education: A dynamic framework. *Educational Psychological Review, 33*, 863–882. https://doi.org/10.1007/s10648-020-09570-w

Thees, M., Altmeyer, K., Kapp, S., Rexigel, E., Beil, F., Klein, P., … Kuhn, J. (2022). Augmented reality for presenting real-time data during students' laboratory work: Comparing a head-mounted display with a separate display. *Frontiers in Psychology, 13*, Article 804742. https://doi.org/10.3389/fpsyg.2022.804742

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior, 108*, Article 106316. https://doi.org/10.1016/j.chb.2020.106316

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *American Journal of Physics, 66*(4), 338–352. https://doi.org/10.1119/1.18863

Tsai, C. Y., & Lai, Y. C. (2022). Design and validation of an augmented reality teaching system for primary logic programming education. *Sensors, 22*(1), 389. https://doi.org/10.3390/s22010389

Tytler, R. (2020). STEM education for the twenty-first century. In J. Anderson, & Y. Li (Eds.), *Advances in STEM educationIntegrated approaches to STEM education* (pp. 41–56). Springer.

Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & Van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development, 85*(3), 842–860. https://doi.org/10.1111/cdev.12169

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*(1), 118–130. https://doi.org/10.1016/j.chb.2010.07.038

Venville, G., & Oliver, M. (2015). The impact of a cognitive acceleration programme in science on students in an academically selective high school. *Thinking Skills and Creativity, 15*(2015), 48–60. https://doi.org/10.1016/j.tsc.2014.11.004

Wang, W. T., Lin, Y. L., & Lu, H. E. (2023). Exploring the effect of improved learning performance: A mobile augmented reality learning system. *Education and Information Technologies, 28*(6), 7509–7541. https://doi.org/10.1007/s10639-022-11487-6

Weng, C., Otanga, S., Christianto, S. M., & Chu, R. J. C. (2020). Enhancing students' biology learning by using augmented reality as a learning supplement. *Journal of Educational Computing Research, 58*(4), 747–770. https://doi.org/10.1177/0735633119884213

Weng, C., Rathinasabapathi, A., Weng, A., & Zagita, C. (2019). Mixed reality in science education as a learning support: A revitalized science book. *Journal of Educational Computing Research, 57*(3), 777–807. https://doi.org/10.1177/0735633118757017

Wild, K. P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15*, 185–200.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*, 433. https://doi.org/10.3389/fpsyg.2013.00433

Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

Wu, P. H., Hwang, G. J., Yang, M. L., & Chen, C. H. (2018). Impacts of integrating the repertory grid into an augmented reality-based learning design on students' learning achievements, cognitive load and degree of satisfaction. *Interactive Learning Environments, 26*(2), 221–234. https://doi.org/10.1080/10494820.2017.1294608

Xu, W. W., Su, C. Y., Hu, Y., & Chen, C. H. (2022). Exploring the effectiveness and moderators of augmented reality on science learning: A meta-analysis. *Journal of Science Education and Technology, 31*(5), 621–637. https://doi.org/10.1007/s10956-022-09982-z

Yoon, S., Anderson, E., Lin, J., & Elinich, K. (2017). How augmented reality enables conceptual understanding of challenging science content. *Journal of Educational Technology & Society, 20*(1), 156–168.

Yoon, S. A., & Wang, J. (2014). Making the invisible visible in science museums through augmented reality devices. *TechTrends, 58*, 49–55.

Yu, K. C., Sahami, K., & Dove, J. (2017). Learning about the scale of the solar system using digital planetarium visualizations. *American Journal of Physics, 85*(7), 550–556. https://doi.org/10.1119/1.4984812

Yu, S., Liu, Q., Ma, J., Le, H., & Ba, S. (2022). Applying augmented reality to enhance physics laboratory experience: Does learning anxiety matter? *Interactive Learning Environments, 31*(10), 6952–6967. https://doi.org/10.1080/10494820.2022.2057547

Zhang, J., Huang, Y. T., Liu, T. C., Sung, Y. T., & Chang, K. E. (2020). Augmented reality worksheets in field trip learning. *Interactive Learning Environments, 31*(1), 4–21. https://doi.org/10.1080/10494820.2020.1758728

Zumbach, J., von Kotzebue, L., & Pirklbauer, C. (2022). Does augmented reality also augment knowledge acquisition? *Journal of Educational Computing Research, 60*(5), 1325–1346. https://doi.org/10.1177/07356331211062945