REVIEW ARTICLE

Comparing Multiple Theories about Learning with Physical and Virtual Representations: Conflicting or Complementary Effects?

Martina A. Rau[1]

Abstract

In most STEM instruction, students interact with visual representations, which can be presented in either in a physical or a virtual mode or in a blended form that combines both modes. While much research has compared the effects of physical and virtual representations on students' learning, the field is far from being able to predict when and why one representation mode is more effective than the other. One reason why making such predictions is particularly difficult is that multiple different theories have been used to explain differences between representation modes. The goal of this article is twofold. First, it presents a survey of the literature to examine which theoretical perspectives have been used to motivate comparisons of representation modes and what predictions they make about their effectiveness. A review of 54 articles reveals five theoretical perspectives: physical engagement, cognitive load, haptic encoding, embodied action schemas, and conceptual salience. While the first two make general predictions about the effectiveness of representation modes, the last three make concept-specific predictions. Second, this article compares these predictions to examine how they conflict and align. This comparison identified several conflicts between theories that predict opposite effects, as well as several alignments where theories make the same predictions but based on different mechanisms. Further, this comparison revealed common confounds in experimental designs of the reviewed studies. The article concludes with recommendations for research to address the identified conflicts and with recommendations for instructors and designers of blended technologies for appropriate choices of representation modes.

✉ Martina A. Rau
marau@wisc.edu

[1] Department of Educational Psychology, University of Wisconsin–Madison, Madison, WI, USA

Introduction

Instruction increasingly uses blended educational technologies that combine physical and virtual experiences, commonly referred to as augmented/mixed reality or blended technologies (e.g., Antle et al. 2009; Johnson-Glenberg et al. 2014; Olympiou and Zacharia 2012). These developments have revived a century-old debate about whether physical interactions with visual representations enhance learning (e.g., Deboer 1991; Huxley 1897). Visual representations are objects that stand for another object, phenomenon, or idea while using perceptually similar features to their referent (Rau 2017; Schnotz 2014). Physical representations are tangible objects that students construct and manipulate with their hands. Virtual representations are displayed on a digital screen and are typically manipulated via mouse, keyboard, or touchscreen input. Blended technologies may combine these experiences by having students manipulate physical objects that affect changes on a virtual screen (e.g., Fjeld et al. 2007). The rationale underlying blended technologies is that physical and virtual representations have complementary advantages that are best leveraged when they are combined (e.g., Antle et al. 2009). However, effective combinations require knowledge about the specific advantages of each representation mode for students' learning. The goal of this article is to review prior research on differential effects of physical and virtual modes on cognitive learning outcomes in science, technology, engineering, and math (STEM) domains.

While research indeed shows that physical and virtual representations have complementary advantages (Olympiou and Zacharia 2012), the field is far from being able to predict when and why each representation mode is effective. Making such predictions is particularly difficult because multiple theoretical perspectives—which have been examined by mostly separate lines of research—yield different predictions for the effectiveness of physical versus virtual representation modes. These differing predictions result from the fact that different theories focus on different types of learning mechanisms. However, in realistic learning settings, effects of physical and virtual representations result from a combination of multiple learning mechanisms. To understand realistic effects of different representation modes, we must understand how these different mechanisms combine. Comparing predictions by different theories will therefore yield new predictions about how different learning mechanisms may interact with one another, for example, whether or under which conditions they might cancel each other out or amplify one another. Understanding interactions among these mechanisms may provide guidance for practitioners about how to choose or combine physical and virtual representations. Further, reviewing whether research has examined interactions between multiple learning mechanisms will highlight gaps in prior research that can guide future research. In sum, this article investigates (1) what predictions different theoretical perspectives make about the effectiveness of physical and virtual representations and (2) whether these predictions conflict or align with each other.

To address research question 1, this article reviews literature on learning with physical and virtual representations and identifies theoretical perspectives that prior research has used to motivate comparisons of these representation modes. To address research question 2, this article compares the predictions that result from these different theoretical perspectives and examines whether research has empirically contrasted predictions that are based on the different theories.

Definitions

Representations are objects that stand for something else—a *referent* (Peirce et al. 1935), which can be another object (e.g., a bathroom sign stands for an actual bathroom) or a concept (e.g., a line in a graph may stand for projected revenue increases). Representations can be *internal* objects that a person imagines or *external* objects that a person encounters in the world (Rau 2017; Schnotz 2014). Further, a common distinction is between symbolic and visual representations (Ainsworth 2006; Schnotz 2014). Symbolic representations such as text and equations have arbitrary mappings to the referent. Visual representations have similarity-based mappings to the referent. This STEM article focuses on external visual representations.

Visual representations differ in the degree to which they are static or interactive (Ainsworth 2008; Rau 2017). Static representations do not change (e.g., a picture of a line graph in a textbook), whereas interactive representations change in response to students' manipulations (e.g., a tool that allows students to change the slope and intercept of a line graph). Interactive representations also differ in terms of how much they constrain students' manipulations. For example, an animated line graph may highly constrain interactions if students can only play, pause, and change viewing speed of a mathematical process. A slightly lesser degree of constraint may allow students to manipulate the slope of the line. A lesser degree of constraint yet may allow students to draw lines and curves that represent various mathematical functions. This article focuses on interactive representations that involve some manipulation of visual features of the representation itself (i.e., beyond playing and pausing an animation).

Finally, visual representations differ by representation mode, that is, whether they are physical or virtual (de Jong et al. 2013; Zacharia et al. 2008). Physical representations are composed of tangible objects that students manipulate by hand. Virtual representations are presented on a screen and typically manipulated via text, mouse, or touchscreen input. With the advent of blended technologies that combine physical and virtual representations, the distinction is no longer dichotomous. Instead, it is useful to distinguish which components of a representation are physical or virtual (e.g., a student may manipulate a physical object to affect changes that are presented virtually). This article focuses on studies that compare representations that are either purely physical or purely virtual as well as studies that compare blended representations that differ in terms of which of their components are physical or virtual.

Methods

Search for Articles

To address the research questions, I used the following methodological approach. I searched research databases for articles published in journals and books, including ERIC, EBSCO, and PsycINFO and Google Scholar using the keywords "physical," "virtual," "tangible," "blended," "augmented," or "mixed," paired with "visuals," "representations," or "manipulatives" without restricting the range of years. The search was conducted in March 2019. I selected articles that presented primary studies that compared the effects of different representation modes (i.e., virtual vs. physical), combinations of modes (e.g., blending of physical and virtual vs. pure physical), or sequences of modes (e.g., physical then virtual vs. virtual then physical) on students' learning of STEM content knowledge. Studies involving virtual labs were included if they contained interactive visual representations and if they compared virtual and

physical components. Dissertations were included if they presented studies that had not been published elsewhere. When meta-reviews came up in the search, the primary studies they cited were included in the review. This search yielded 54 articles (see Table 1 for an overview).

Focus of the Reviewing Process

In reviewing these articles, I identified the theoretical perspectives they used to motivate the comparison between representation modes. That is, while the articles may have included additional theories to motivate other aspects of the study (e.g., the investigation of visual representations in general), I focused on those theories used to motivate the comparison of representation modes. Further, while several articles investigated representation modes in collaborative learning settings, my review focused on arguments pertaining to individual learning. Note that this review does not focus on results of individual studies about the relative effectiveness of representation modes, as this has been covered extensively elsewhere (e.g., Carbonneau et al. 2013; de Jong et al. 2013; Moyer-Packenham and Westenskow 2013). Finally, the review focused on arguments about cognitive learning processes and learning outcomes and therefore excluded arguments pertaining to practicality (e.g., virtual representations are cheaper) or conventions (e.g., physical labs are common practice).

Theoretical Perspectives

Five theoretical perspectives were identified: physical engagement, cognitive load, haptic encoding, embodied action schemas, and conceptual salience. Table 2 shows the prevalence of these perspectives across the articles. While 14 articles used only one perspective to motivate the study, most articles referred to multiple theoretical perspectives. Further, Table 2 shows which theoretical perspectives were used in conjunction to motivate the studies in the reviewed articles.

Physical Engagement

Twenty-three articles motivated the comparison of representation modes by referring to the potential of physical representations to engage students physically with the learning materials (see Table 2). While two of these articles used physical engagement as the sole motivator of the comparison, nine articles used this theory in conjunction with cognitive load, 12 with haptic encoding, 7 with embodied action schemas, and 16 with conceptual salience (see Table 2).

Overview of the Theoretical Perspective

Physical engagement perspectives on learning with representations originate in research on "hands-on" educational activities from as early as the late nineteenth century, before virtual representations were available (Huxley 1897). This early research suggests that direct contact with the physical environment engages students more thoroughly with the learning content than reading books or listening to lectures could (Huxley 1897). Building on this early work, scholars have argued that physical representations allow for kinesthetic interactions that are motivating and engage students' interest (Deboer 1991; Flick 1993). An increase in interest and motivation, in turn, leads to deeper processing of the concepts and thereby yields higher

Table 1 Overview of articles included in this review

Study	Perspectives	Study design	Domain
(Atanas 2018)	Conc	Quasi experiment	Physics
(Bakker et al. 2012)	Phys, Conc, Emb	Case study	Physics
(Barrett et al. 2015)	CL, Conc, Hapt	Randomized experiment	Chemistry
(Baturo et al. 2003)	Conc, Hapt	Case study	Math
(Burris 2010)	Phys, Conc, Hapt	Quasi experiment	Math
(Chien et al. 2015)	Conc, Emb	Randomized experiment	Physics
(Chini et al. 2012)	CL, Conc	Quasi experiment	Physics
(Cuendet et al. 2012)	Phys, CL, Emb	Randomized experiment	Math
(Doias 2013)	Phys, CL, Hapt	Quasi experiment	Math
(Dori and Barak 2001)	Conc	Quasi experiment	Chemistry
(Drickey 2000)	Conc	Quasi experiment	Math
(Finkelstein et al. 2005)	CL, Conc	Quasi experiment	Physics
(Gire et al. 2010)	Conc	Quasi experiment	Physics
(Han 2013)	Phys, Hapt	Experiment	Computer science
(Horn et al. 2012)	Hapt	Quasi experiment	Computer science
(Huppert et al. 2002)	Conc	Quasi experiment	Biology
(Jaakkola et al. 2011)	CL, Conc	Randomized experiment	Physics
(Katsioloudis et al. 2015)	Conc	Quasi experiment	Engineering
(Kim 1993)	Phys, Conc	Quasi experiment	Math
(Klahr et al. 2007)	Phys, Conc	Randomized experiment	Engineering
(Lee and Chen 2015)	Conc	Quasi experiment	Math
(Magruder 2012)	CL, Conc	Quasi experiment	Math
(Manches et al. 2010)	Phys, CL, Conc, Hapt, Emb	Randomized experiment	Math
(Manches et al. 2009)	CL, Conc, Hapt, Emb	Case study	Math
(Melcer et al. 2017)	Phys, Hapt, Emb	Randomized experiment	Computer science
(Melcer and Isbister 2018)	Phys	Randomized experiment	Computer science
(Moyer-Packenham et al. 2013)	Conc	Randomized experiment	Math
(Ozgun-Koca and Edward 2011)	Phys, Conc	Survey study	Math
(Olympiou and Zacharia 2012)	Conc, Hapt	Randomized experiment	Physics
(Pan 2013)	Phys, CL, Conc, Emb	Quasi experiment	Engineering
(Pyatt and Sims 2012)	Phys, CL, Conc	Randomized experiment	Chemistry
(Rau 2017)	Conc, Hapt	Case study	Chemistry
(Renken and Nunez 2013)	CL, Conc	Randomized experiment	Physics
(Schneider and Blikstein 2018)	Phys	Randomized experiment	Biology
(Schneider et al. 2016)	Phys, Conc, Hapt	Randomized experiment	Logistics
(Skulmowski et al. 2016)	CL, Hapt, Emb	Randomized experiment	Biology
(Smith and Puntambekar 2010)	Conc, Hapt	Quasi experiment	Physics
(Stull et al. 2013)	Conc, Hapt	Randomized experiment	Chemistry
(Stull and Hegarty 2016)	CL, Hapt	Randomized experiment	Chemistry
(Stusak et al. 2015)	Hapt	Randomized experiment	Math
(Suh and Moyer 2007)	CL, Conc	Quasi experiment	Math
(Sung et al. 2015)	Conc	Quasi experiment	Math
(Toth et al. 2009)	Cognitive load, Conc	Quasi experiment	Biology
(Triona and Klahr 2003)	Phys, Conc, Hapt	Randomized experiment	Physics
(Wang and Tseng 2018)	Phys, CL, Conc, Hapt	Quasi experiment	Chemistry
(Yannier et al. 2015)	Phys, Hapt, Emb	Randomized experiment	Physics
(Yannier et al. 2016)	Phys, Conc, Hapt, Emb	Randomized experiment	Physics
(Yuan et al. 2010)	Conc	Quasi experiment	Math
(Zacharia and Constantinou 2008)	Phys, CL, Conc	Randomized experiment	Physics
(Zacharia et al. 2012)	Phys, Conc, Hapt	Randomized experiment	Physics
(Zacharia and Michael 2016)	Conc, Hapt	Randomized experiment	Physics
(Zacharia and Olympiou 2011)	Phys, CL, Conc, Hapt	Randomized experiment	Physics
(Zacharia et al. 2008)	Phys, CL, Conc	Randomized experiment	Physics

Phys physical engagement, *CL* cognitive load, *Hapt* haptic encoding, *Emb* Emb, *Conc* conceptual salience

cognitive learning outcomes. For example, Doias (2013), referencing Montessori (1966), states that students "learn and retain information best when they can manipulate objects with their own hands." Further, physical representations offer concrete experiences that are connected to realistic contexts (Clements 1999; Goldstone and Son 2005). Realistic contexts provide perceptually rich sensory-motor experiences that allow for intuitive processing in the sense that students can connect them to prior experiences (Goldstone and Son 2005). Making such connections allows students to embed newly learned content in existing knowledge structures, which strengthens memory. It has been argued that concrete experiences that allow for such intuitive processing are necessary for students to understand abstract concepts (e.g., Bruner 1966; Dienes 1961; Wolfe 2001). Building on this literature, scholars have argued that virtual representations deprive students of hands-on experiences and may therefore be detrimental to learning (Scheckler 2003).

In sum, physical engagement perspectives posit that concrete experiences with physical representations increase engagement with to-be-learned concepts and thereby enhance learning outcomes.

Treatment of the Theoretical Perspective by the Reviewed Articles

The review of the selected articles revealed a disparity in whether they referred to the physical engagement perspective in a positive or a negative way. Articles that compared virtual representations to virtually enhanced physical representations (e.g., tangible user interfaces) often referred to physical engagement in a positive way to motivate the use of blended modes. For example, Cuendet et al. (2012) state that physical representations "increase students' engagement" (p. 99) with the content. Pyatt and Sims (2012) and Schneider and Blikstein (2018) mention that physical representations can increase active engagement with target concepts and may thereby enhance learning. In contrast, studies that compared virtual representations to ordinary physical representations (i.e., without virtual enhancements) often mentioned that the physical engagement perspective has not held up against empirical evidence and describe it as outdated. For example, Han (2013) Jaakkola et al. (2011), and Zacharia and Olympiou (2011) mention this perspective in the context of a review of prior research that found no evidence that virtual representations restrict learning because they do not provide hands-on experiences (e.g., Triona and Klahr 2003).

Table 2 The first column shows the total number of articles referencing each theoretical perspective

	Prevalence of theory	Physical engagement	Cognitive load	Haptic encoding	Embodied schemas	Conceptual salience
Physical engagement	23	*2*	9	12	7	16
Cognitive load	19	9	*0*	8	5	15
Haptic encoding	24	12	8	*2*	6	16
Embodied schemas	10	7	5	6	*0*	6
Conceptual salience	42	16	15	16	6	*10*

The remaining columns give an overview of theoretical perspectives that were used in conjunction with one another to motivate the studies in the reviewed articles. The diagonal (in italics) corresponds to articles that used only one theoretical perspective. Because a given theory might be used in conjunction with multiple other theories, the first column is not the sum of the remaining columns

In summary, the physical engagement perspective suggests that physical representations are generally more effective than virtual representations. While this perspective is prevalent across the reviewed articles, articles focusing on blended modes tend to present a positive view on this perspective, whereas other articles mention it in a negative way to compare it to other perspectives that they view more applicable.

Cognitive Load

Nineteen articles motivated the comparison of representation modes by referring to the potential of virtual representations to reduce cognitive load (see Table 2). No article used cognitive load as the sole motivator of the comparison. Nine articles referred to this theory together with physical engagement, eight with haptic encoding, five with embodied action schemas, and 15 with conceptual salience (see Table 2).

Overview of the Theoretical Perspective

Cognitive Load Theory (Chandler and Sweller 1991; Sweller et al. 1998) builds on findings that human capacity for cognitive processing is limited (Miller 1956). It is generally assumed that humans can hold 7 ($\pm$ 2) chunks in working memory before cognitive capacity is exceeded (Miller 1956). If processing of learning material exceeds this capacity, students experience cognitive overload, which can hinder learning. Specifically, Cognitive Load Theory distinguishes three types of cognitive load (Sweller et al. 1998). Intrinsic load is attributed to the difficulty level of the learning material; that is, the more chunks need to be processed simultaneously to understand the content, the higher the intrinsic load of the material. Germane load results from the construction of new schemas and the integration of new content in existing knowledge structures. Extraneous load results from processing demands imposed by the design of the material; that is, the more distracting the design of the material is—for instance, if it includes "seductive" details that could distract from the learning content—the higher its extraneous load. Further, visual and verbal modalities are processed in parallel via different channels (Chandler and Sweller 1991). Because each channel has its own capacity, the addition of visual information to verbal information makes more efficient use of students' working memory capacity.

In sum, cognitive load perspectives posit that learning can be enhanced if learning materials, including visual representations, are designed in ways that reduce the risk of cognitive overload.

Treatment of the Theoretical Perspective by the Reviewed Articles

In referring to Cognitive Load Theory, research on physical and virtual representations typically refers to two design principles that seek to reduce the risk of cognitive overload (see Mayer 2005, 2009, 2010; Mayer and Moreno 2003 for a detailed description of the principles).

First, the contiguity principle recommends to design learning materials so that students do not have to split their attention between multiple sources of information (Mayer 2005, 2009; Sweller et al. 1998; see Schroeder and Cenkci 2018 for a recent meta-analysis). When visual information is presented separately from verbal information (e.g., a book with a picture on a different page than the text describing it), students have to engage in visual search processes to

establish mappings between the two sources. Because such search processes are cognitively demanding and are not relevant to the target concepts, separate presentation increases extraneous load, which could cause cognitive overload. Further, with greater distance between the sources, students are altogether less likely to actively establish mappings between them.

This principle has implications for the comparison of physical and virtual representations when considering how they are typically presented along with other instruction. When students use virtual representations that are displayed on a digital screen, instruction on how to manipulate them is typically presented on the screen as well. By contrast, when students interact with physical representations, instruction is often presented on a paper sheet or on a digital screen, which requires students to split their attention between two sources. Hence, physical representations have a higher risk of inducing split attention effects that can impede learning. For example, Barrett et al. (2015) and Lee and Chen (2015) describe the capability of virtual representations to integrate multiple resources as one of their advantages.

Second, the coherence principle recommends eliminating surface features, that is, perceptual details that are not relevant to the target concepts and that could distract students (Mayer 2005, 2009; see Rey 2012 for a meta-analysis). Such "seductive" details can increase extraneous load because it may not be immediately obvious to students that they are irrelevant and because it requires cognitive effort to ignore them (Chandler and Sweller 1991; Sweller et al. 1998). Because physical representations have richer, more concrete features that may be distracting, they may increase extraneous cognitive load compared to virtual representations, which may in turn hinder learning (Goldstone and Son 2005; Kaminski and Sloutsky 2013; Kaminski et al. 2009). For example, Magruder (2012) suggests that virtual representations have the same informational content and require the same transformation as physical representations while being less distracting. Zacharia and Olympiou (2011) describe ways in which virtual representations can eliminate irrelevant details and thereby reduce cognitive load compared to their physical counterparts. Stull et al. (2013) suggest that virtual representations make it easier to identify key features of the representations, which reduces cognitive load. Jaakkola et al. (2010) and Pyatt and Sims (2012) provide examples of potentially distracting visual features in physical circuits, such as the colors of wires, or other distracting factors of their manipulation, such as tangled wires.

Further, virtual representations can be augmented with highlights that focus students' attention, thereby reducing cognitive load that is associated with the search for relevant information (Wang and Tseng 2016). Indeed, several of the reviewed articles cite studies showing that advantages of virtual over physical representations are due to increased cognitive efficiency and attention to relevant features (Barrett et al. 2015; Durmus and Karakirik 2006; Yuan et al. 2010). The inclusion of rich, concrete features may pose a particular challenge for young students. For example, Manches et al. (2009) and Manches et al. (2010) refer to research by Uttal et al. (1997) to suggest that it is particularly cognitively demanding for children to interpret physical objects to stand for something other than themselves.

Finally, the interactions with physical representations themselves may involve actions that are unrelated to the to-be-learned concepts, thereby increasing cognitive load. For example, Suh and Moyer (2007) suggest that the cognitive effort associated with manipulating physical representations may be too high for students and can impede their ability to relate the representations to abstract concepts. In contrast, virtual representations can automate routine tasks and thereby reduce cognitive load (Toth et al. 2009).

However, not all articles seem to agree with this interpretation of Cognitive Load Theory. For example, Cuendet et al. (2012) suggest that physical representations are more natural and may therefore reduce cognitive effort necessary to manipulate them, which in turn reduces cognitive load. Others (e.g., Chini et al. 2012; Schneider et al. 2016; Zacharia et al. 2012) draw on an extended version of Cognitive Load Theory that accounts for haptic encoding (see below) to argue that physical representations can reduce the risk of cognitive overload.

In sum, while there is some disagreement on how to interpret Cognitive Load Theory in predicting differences between physical and virtual representations, most of the reviewed articles use this theory to highlight general advantages of virtual representations because they can be more easily integrated with other learning materials and exclude distracting features.

Haptic Encoding

Twenty-three articles motivated the comparison of representation modes with the capability of physical representations to offer haptic encodings of information (see Table 2). In two of these articles, haptic encoding was the sole motivator of the comparison. Twelve articles referred to this theory together with physical engagement, eight with cognitive load, six with embodied action schemas, and 16 with conceptual salience (see Table 2).

Overview of the Theoretical Perspective

Research on haptic encoding proposes that physical representations provide haptic cues for learning, that is, students can encode the target concepts directly through the sense of touch, in addition to the visual sense (Magana and Balachandran 2017; Zaman et al. 2012). The availability of haptic cues in physical representations allows students to make more explicit connections between the perceived environment and the target concepts, compared to virtual representations (Shaikh et al. 2017; Skulmowski et al. 2016). Haptic encoding perspectives describe three mechanisms of how the availability of haptic cues can enhance learning. First, if students can connect visual as well as haptic features to concepts, they have more retrieval cues than if they can only connect visual features to concepts. This increases their ability to remember the given concept and to make further connections to it later on. Second, haptic experiences contribute to perceptual grounding of abstract concepts. Perceptual grounding describes how abstract concepts originate in concrete experiences that become increasingly stylized (Goldstone et al. 1997; Harnad 1990). That is, initial experiences with a concept are strongly tied to sensorimotor experiences of the concept (e.g., lifting objects requires energy) until they gradually become more and more abstract (e.g., understanding energy as a form of strength). Third, haptic cues can increase cognitive capacity. This line of reasoning draws on updated version of Cognitive Load Theory. While Baddeley (1992)'s original working memory model proposed that visual and verbal modalities can be processes in parallel, recent modifications (e.g., Baddeley 2012) also include a haptic modality. Since each modality has its own working memory, adding a haptic modality increases students' overall cognitive capacity and thereby reduces the risk of cognitive overload.

Haptic encoding perspectives fundamentally differ from physical engagement perspectives. While physical engagement perspectives refer to general engagement with the learning materials that is not specific to the target concepts (i.e., if a physical representation is more motivating, that is true regardless of which concept a student is learning), haptic encoding perspectives focus on how specific concepts are encoded in a physical representation. If a

physical representation provides haptic cues for one concept but not for another, learning benefits would be expected only for the concept for which haptic cues are available and only if students explicitly attend to these cues.

In sum, haptic encoding perspectives focus on advantages of haptic cues in providing retrieval cues, in offering opportunities for perceptual grounding, and in reducing risk of cognitive overload. Hence, physical representations can enhance learning, provided they offer haptic cues for the target concepts.

Treatment of the Theoretical Perspective by the Reviewed Articles

The way in which the reviewed articles refer to haptic encoding perspectives reflects their focus on connecting haptic cues and target concepts. For example, Barrett et al. (2015) and Stull and Hegarty (2016) describe advantages of physical representations for learning of spatial concepts as a result of their direct encoding of spatial information that is accessible via touch (e.g., bond angles in chemical molecules). Some authors further emphasize that the presence of haptic cues alone is not sufficient, but that students' manipulations of the haptic features need to be associated with the to-be-learned concepts (Manches et al. 2010; Melcer et al. 2017).

The reviewed articles describe a variety of advantages of haptic encoding. First, several articles emphasize advantages for memory and recall. For example, Olympiou and Zacharia (2012) and Wang and Tseng (2018) suggest that physically feeling the materials and procedures performed in laboratories experiments provides richer experiences that may enhance learning. Similarly, Stusak et al. (2015) suggest that processing information both through a visual and a haptic channel enhances memorability.

Second, a few articles refer to advantages of haptic encoding in terms of providing perceptual grounding for abstract concepts. For example, Han (2013), Zacharia and Michael (2016), and Zacharia and Olympiou (2011) describe physical experiences of concepts as the foundation of understanding. Similarly, Zacharia et al. (2012) suggest that physical touch can serve as an anchor that grounds conceptual understanding. Such haptic experiences may also improve students' ability to mentally manipulate the representation later on when it is no longer physically present (Yannier et al. 2015).

Third, several articles describe advantages of haptic encoding in terms of increased cognitive capacity. For example, Zacharia et al. (2012) propose that physical representations increase overall cognitive capacity because they allow students to utilize their haptic working memory. Similarly, Chini et al. (2012) suggest that physically touching representations offers an additional processing pathway that increases cognitive capacity and thereby decreases cognitive load. Also Schneider et al. (2016) suggest that haptic information provided by physical representations can decrease cognitive load. Some of the reviewed articles relate this argument to the distributed cognition literature. For example, Manches et al. (2010) propose that physical representations allow offloading memory into physical materials, which can lead to a reduction in cognitive load. Note that this argument is distinct from the cognitive load perspective described above in that the reviewed articles emphasize that cognitive load advantages of haptic encoding only come into play if the haptic cues encode the target concepts. For example, Skulmowski et al. (2016) provide a nuanced discussion of how physical representations may on the one hand increase cognitive load because they require more motor actions, but how they may on the other hand reduce cognitive load if these motor actions are directly associated with the to-be-learned concepts.

In sum, this perspective suggests that physical representations are not generally more effective than virtual representations but that their advantages depend on whether they allow students to explicitly process haptic cues that encode specific target concepts. If physical representations offer haptic cues for a concept, they may enhance memory, provide perceptual grounding of the concept, and reduce cognitive load.

Embodied Action Schemas

Ten articles motivated the comparison of representation modes by referring to the potential of physical representations to activate embodied action schemas (see Table 2). None of these articles used embodied action schemas as the sole motivator of the comparison. Seven articles used this theory in conjunction with physical engagement, five with cognitive load, six with haptic encoding, and six with conceptual salience (see Table 2).

Overview of the Theoretical Perspective

Embodied action theory suggests that body movements influence cognition (Glenberg 2010; Glenberg et al. 2013; Wilson 2002). According to this theory, cognition evolved to facilitate our interactions in the real world by mentally simulating effects of our actions (Glenberg 1997). Similarly, all higher-order thinking (e.g., thinking about abstract concepts) can be viewed as a mental simulation of body actions. The use of embodied metaphors in our day-to-day language is an illustrative example of embodied action theory. Embodied metaphors tie abstract concepts to body actions, thereby grounding the concepts in real-world experiences. For example, the phrase "that made an impression on me" relates an emotional experience to a physical experience of an imprint (Lakoff and Johnson 1980, p. 127). More formally, embodied metaphors are implicitly acquired action schemas that result from sensory-motor experiences of our body movements and interactions in the world (Lakoff and Johnson 1980). Embodied action schemas can be invoked through speech (e.g., using metaphors as in the example above) or through body movements (e.g., moving one's hands upward can invoke concepts related to increase, improvement, or happiness) (Black et al. 2012; Johnson-Glenberg et al. 2014). The more embodied the experience of a concept through physical interactions, the higher learning outcomes regarding that concept (Johnson-Glenberg et al. 2014).

According to this theory, learning means that students form mental simulations that are grounded in such embodied action schemas (Abrahamson and Lindgren 2014; Clark 2013). Research on embodied action schemas shows that learning can indeed be enhanced by moving the body in ways that are synergistic with mental simulations of target concepts (Hayes and Kraemer 2017). For example, when students learn about growth functions, they mentally simulate increase that is grounded in upward movements, which may be enhanced by physically moving their hand upwards.

While this perspective is related to the perceptual grounding argument described under haptic encoding, it is distinct from it in the sense that it does not require students to be explicitly aware of the connection between the body action and the to-be-learned concepts. Hence, it describes an implicit, often nonverbal mechanism. Indeed, even seemingly unrelated body movements affect cognition. For example, being instructed to enact shapes helps students understand geometry concepts even if they are not aware of the relation between their enactments and the concepts (Nathan and Walkington 2017).

Further, embodied action schema and physical engagement perspectives are somewhat related in that both propose that moving the body can enhance learning. However, they are distinct in that the former suggests that only movements that invoke action schemas that align with the target concept positively affect learning, whereas the latter proposes effects independent of which concept is learned. Indeed, in a recent review, Duijzer et al. (2019) use the involvement of the student's body as a key dimension to organize embodied cognition studies. Specifically, they distinguish studies in which students move their own bodies and studies where students observe movements of others. In the latter case, they refer to neural mirroring mechanisms (see Anderson 2010; Gallese and Lakoff 2005 for overviews) to explain why observing movements of others can activate action schemas and thereby ground concepts in physical experiences. Nevertheless, in each case, the movement needs to be related to the target concept. A review by Skulmowski and Rey (2018) further underlines the importance of this relation. They compared studies in which the relation between concept and movement was weak or strong. They found that stronger relations yielded higher learning outcomes. Further, they examined the role of the amount of body movements in embodied learning. While high amounts of body movement can compensate for weak relations between movement and concept, they can also increase cognitive load and thereby reduce learning outcomes.

In sum, embodied action schema perspectives propose that body actions implicitly ground abstract concepts in real-world experiences. Therefore, an effect of a particular representation mode results from its capacity to ground a specific concept in body actions.

Treatment of the Theoretical Perspective by the Reviewed Articles

Many of the reviewed articles refer to embodied action schemas to advocate for physical over virtual representations. For example, Pan (2013) refers to the link between sensorimotor action and cognition to argue that physical representations can "tap into cognition at a very primal level and may provide a more unconscious understanding" of the to-be-learned concepts (p. 8). Several articles refer to the potential of physical representations to activate sensorimotor states that can enhance learning. For example, Skulmowski et al. (2016) interpret the embodied cognition literature's tenet that abstract thinking is a reenactment of sensorimotor perceptual states as suggesting that physical representations have motor affordances that may enhance learning. Similarly, Yannier et al. (2015) reference the literature on embodied action schemas to argue that the research showing that mind and body are integrated during learning suggests that bodily activity can support cognition. In addition, several of the reviewed articles refer directly to the activation of embodied action schemas. For example, Bakker et al. (2012) and Melcer et al. (2017) motivate the design of tangible user interfaces in their capability to engage students in recurring sensorimotor experiences that activate embodied metaphors. Yannier et al. (2016) propose that interactions with physical representations can trigger affordances for action that activate embodied schemas.

This view stands in contrast to research on embodied cognition, which suggests that it is not the mode of the representation that matters but how students move their bodies when interacting with the representation. Studies show how virtual representations that are manipulated by movements that invoke synergistic embodied metaphors enhance students' learning of target concepts more so than those manipulated by less synergistic movements (Segal et al. 2014). Further, representations that are designed so that they require students to move their bodies in ways that are more synergistic to

target concepts have been shown to be more effective (Abrahamson and Lindgren 2014; Antle et al. 2009; Bamberger and diSessa 2003).

Yet another view in the reviewed articles is that imagined movements may suffice to leverage embodied action schemas and that physical movement may not be necessary. For example, King and Smith (2018) refer to research on embodied action schemas showing that imagined actions are neurologically correlated with brain activation resulting from physical actions and may hence yield the same benefits for learning. Similarly, Manches et al. (2009, 2010) describe research showing that viewing gestures performed by instructors can also activate embodied action schemas to explain why virtual representations may have advantages for learning over physical ones.

In sum, this theoretical perspective suggests that (actual, observed, or imagined) actions used to manipulate representations may affect learning more so than the representation mode itself. Specifically, representations that engage students in body actions that invoke embodied schemas that are synergistic with the to-be-learned concepts should enhance learning. Because physical and virtual representations are manipulated via different movements, this theory may nevertheless explain effects of representation modes.

Conceptual Salience

Forty-two articles motivated the comparison of representation modes by referring to their capability of making concepts salient (see Table 2). In ten of these articles, conceptual salience was the sole motivator of the comparison. Sixteen articles used this theory in conjunction with physical engagement, 15 with cognitive load, 16 with haptic encoding, and six with embodied action schemas (see Table 2).

Overview of the Theoretical Perspective

Conceptual salience perspectives build three lines of prior research. First, they build on information processing accounts of cognition (Miller 1956), Multimedia Learning Theory (Mayer 2005, 2009) and the Integrated Theory of Picture Comprehension (Schnotz 2005; Schnotz and Bannert 2003), which hold that students have to explicitly attend to information in order to process it in working memory. Further, the design of external representations can affect the likelihood that students attend to certain information (Mayer and Moreno 1998; Schnotz 2005; Schnotz and Bannert 2003). For example, highlighting a specific feature of a representation makes it more likely that students pay attention to it, thereby making this feature more salient. If a salient feature carries meaningful information about the to-be-learned concept, it is *conceptually salient*.

Second, they build on affordance theory (Gibson 1997), which proposes that perception and action are invariably intertwined. Students do not "objectively" see representations but instead subjectively see representations as allowing for certain actions that help them achieve certain goals. That is, students do not perceive the objects themselves but their affordances for action. According to affordance theory, even if a physical and virtual representation provides the same information, they are perceived differently because they naturally afford different actions. Further, according to Olympiou and Zacharia (2012), virtual representations emerged to address a need to complement physical representations and therefore were designed to affordances that complement (and hence differ from) physical representations.

Third, conceptual salience perspectives build on research on conceptual change, which suggests that students have to be confronted with events that challenge their preconceptions (diSessa 2014; Vosniadou 1994). Interactive representations provide the grounds for students to explore events that challenge their thinking and compare these events to their own conceptions (Olympiou and Zacharia 2012). Because physical and virtual representations differ in terms of which aspects of events they make salient, they differ in their effectiveness to induce conceptual change about a particular concept or event.

Research on conceptual salience builds on a large number of studies that have compared virtual and physical representations (Chini et al. 2012; Klahr et al. 2007; Yuan et al. 2010; Zacharia and Constantinou 2008). These studies did not find conclusive evidence that either representation mode is generally more effective. However, the pattern of results suggests that the effectiveness of representation modes depends on whether they make the target concept salient; that is, whether they can draw students' explicit attention to the concept. Experimental evidence for this interpretation comes from a study that determined a priori which representation mode made which target concept more salient (Olympiou and Zacharia 2012). Results showed that students indeed benefited more from the representation mode that had advantages in terms of conceptual salience.

In sum, the main tenet of conceptual salience perspectives is that effects of representation modes are concept-specific and depend on whether they make a concept salient.

Treatment of the Theoretical Perspective by the Reviewed Articles

In contrast to findings that effects of representation modes are specific to the target concept, the review of the identified articles revealed that many articles used this theoretical perspective to argue for the effectiveness of a particular representation mode. On the one hand, several authors emphasize the capability of virtual representations to make concepts salient. For example, several authors suggest that one advantage of virtual representations is that they can make unobservable phenomena visible and thereby conceptually salient (Chien et al. 2015; Drickey 2000; Finkelstein et al. 2005; Gire et al. 2010; Lee and Chen 2015; Moyer-Packenham and Westenskow 2013; Pyatt and Sims 2012; Renken and Nunez 2013; Yannier et al. 2016; Yuan et al. 2010). Burris (2010) extends this argument to connections to symbolic representations and suggests that virtual representations afford easier connections to abstract symbols. Further, the ability of virtual representations to provide immediate feedback on interactions can help students attend to relevant visual features (Magruder 2012; Sung et al. 2015).

Several authors connect this latter argument to Cognitive Load Theory. For example, Chini et al. (2012), Smith and Puntambekar (2010), and Toth et al. (2009) suggest that the capability of virtual representations to highlight features and constraining certain interactions enhances students' ability to notice conceptually relevant information. On the other hand, some authors argue that physical representations have advantages when it comes to conceptual salience. For example, Schneider et al. (2016) and Stull et al. (2013) emphasize advantages of physical representations make spatial information salient. Others have argued that physical representations are more concrete and detailed, which can make information more salient (Stusak et al. 2015).

Overall, however, there seems to be an agreement that different representation modes may have complementary advantages in making concepts salient—especially among the recent publications within the reviewed articles. While some argue that the advantages of

physical and virtual representations depend on how they display information relevant to the content being taught (Barrett et al. 2015), the overarching consensus seems to be that displaying information is not sufficient to make concepts salient, but that students need to interact with them (e.g., Jaakkola et al. 2011; Manches et al. 2009; Olympiou and Zacharia 2012; Triona and Klahr 2003; Wang and Tseng 2018; Zacharia et al. 2012). That is, physical and virtual representations afford different types of interactions, which allows them to make different concepts salient. For example, Olympiou and Zacharia (2012) argue that interacting with physical representations of experiments allows for measurement errors—hence making this concept more salient—whereas virtual representations afford cleaner, more controlled interactions and hence allow making concepts of systematic variation more salient.

In sum, this perspective suggests that it is not the representation mode itself that affects learning. Rather, regardless of the mode, a representation should be more effective if it engages students in actions that make the target concept more salient.

Comparison of Theoretical Perspectives

Figure 1 summarizes and compares the mechanisms and scope of each theoretical perspective. This also highlights several conflicts between the perspectives. First, Fig. 1 illustrates that physical engagement and cognitive load perspectives predict general effects on learning outcomes that are not specific to a particular concept (i.e., the arrows in Fig. 1 point at the full circle that stands for general learning outcomes). Specifically, physical engagement perspectives make unspecific predictions that are based on the assumption that students' increased engagement with learning content enhances learning outcomes, without fully specifying the underlying mechanisms. Cognitive load perspectives mainly propose advantages of virtual representations because they can be more easily designed to reduce extraneous cognitive load that physical ones. By contrast, haptic encoding, embodied action schemas, and conceptual salience perspectives make predictions that are specific to the target concept (i.e., the arrows point at specific concepts).
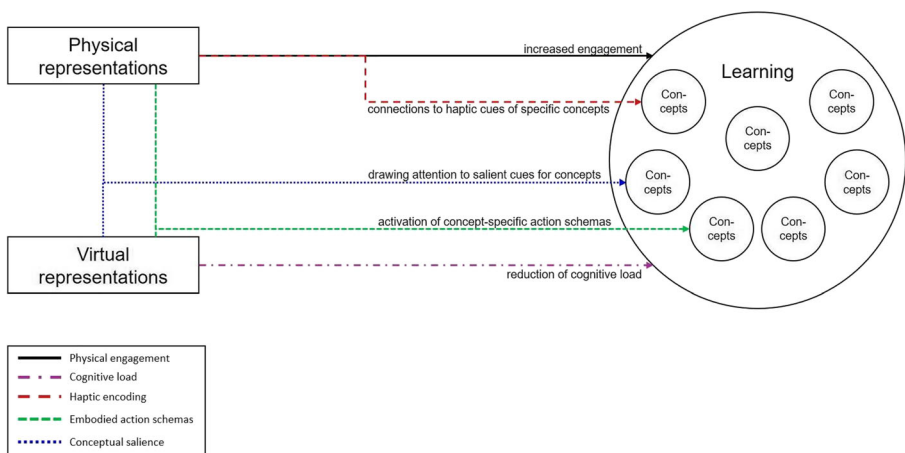

Fig. 1 Summary and comparison of the different theoretical perspectives

It seems that this conflict has mainly been resolved in recent literature, as the consensus is that there is no representation mode that is consistently better than the other (see above). However, this does not mean that physical engagement and cognitive load effects do not occur. Rather, it appears that they are not strong enough to override the concept-specific effects described by the other theoretical perspectives. For example, if a physical representation contains seductive details that increase cognitive load but makes the target concept more salient, it may still yield higher learning outcomes than a virtual representation that does not include seductive details and fails to make the target concept salient. Hence, this conflict highlights shortcomings of the physical engagement and cognitive load perspectives; namely, that they ignore the consensus of prior research that found concept-specific effects of representation modes. Likewise, the concept-specific theories should not ignore general effects due to physical engagement and cognitive load mechanisms.

A second conflict exists between the physical engagement and cognitive load perspectives. Both predict effects of representation modes per se but in opposite directions. As illustrated in Fig. 1, the physical engagement perspective predicts advantages of physical representations, whereas the conceptual load perspective predicts advantages of virtual representations. While it is possible that both perspectives describe mechanisms that co-occur when students work with physical and virtual representations, we know little about how these mechanisms interact.

When examining how the reviewed studies treated this conflict, I found that only one of the 54 reviewed studies (Skulmowski et al. 2016) tested for interactions between designs that reduce cognitive load (integrated labels) and representation mode (blended vs virtual). The authors found an interaction effect, such that integrated labels enhanced learning for the blended mode but not for the virtual mode. Even though this study did not compare pure physical and virtual modes, it suggests that physical engagement and cognitive load mechanisms may interact. In 18 of the 54 reviewed studies, the designs pertaining to cognitive load were the same between modes, 16 articles did not provide enough information about this issue, and 20 confounded instructional designs that reduce cognitive load with representation mode. That is, over one third of the studies confounded cognitive load and representation mode in their experimental designs. Confounds typically resulted from virtual representations being integrated with instructional text to reduce cognitive load, whereas physical representations were not. When such confounds exist, it is possible that effects other than the representation mode itself may account for the effects in these studies. Hence, these conflicts highlight shortcomings of prior studies that have not teased apart these two mechanisms. Future research should systematically examine to what degree physical engagement and cognitive load mechanisms may counteract one another and to what extent either of these mechanisms drive differences between representation modes.

A third conflict exists between the embodied schemas and conceptual salience perspectives. As illustrated in Fig. 1, neither predicts an advantage of the representation mode per se. Instead, both predict effects that are specific to the target concept and that depend on students' actions. However, the mechanisms they describe are fundamentally different. The embodied schemas perspective predicts action effects based on implicit mechanisms that students are not necessarily aware of. By contrast, the conceptual salience perspective predicts action effects based on explicit mechanisms that require students' attention to how the representations show concepts. Because of these different underlying mechanisms, these predictions often conflict in practice. An action that implicitly invokes embodied schemas without requiring students' awareness may at the same time reduce the saliency of the concept. For example, consider a student learning about constant functions using an interactive coordinate graph. If she has to

move her hand sideways to plot the dots, this movement may induce an embodied metaphor related to equality (Lakoff and Johnson 1980), which is synergistic to the concept that in a constant function, the y-value is equal for all x-values. At the same time, however, the sideways movement means that the student does not have to pay attention to the y-value each time she plots a dot. Therefore, the action that induces a synergistic embodied schema does not make the concept salient. By contrast, if the student had to perform an intermediate action (e.g., pick up a pin to place on a physical board that shows the coordinate graph), she would have to find the y-value each time, which makes it more salient that the y-values are equal. However, this action no longer induces a synergistic embodied schema.

When inspecting the reviewed articles for this conflict, I found that only one study used an intervention that induced embodied schemas (Bakker et al. 2012). However, in this study, students were explicitly asked to relate their actions to the target concept, so that the intervention did not exclusively manipulate embodied action schemas. Indeed, in most research on learning with representations that is based on embodied schema theory, the interventions are designed to help students explicitly connect embodied schemas to concepts (Abrahamson and Lindgren 2014; Segal et al. 2014). Further, while activation of embodied schemas is effective without such explicit connections, connecting actions to concepts enhances the effectiveness of this activation (Nathan and Walkington 2017; Nathan et al. 2014). However, prior research has typically focused on only one representation mode at a time, such as virtual only (Segal et al. 2014), physical only (Dackermann et al. 2016), or blended only (Howison et al. 2011). Consequently, it is unclear to what extent the embodied schemas mechanism by itself accounts for possible differences between representation modes. Further, in cases where embodied action schemas and conceptual salience conflict because—as is the case in many representations that were not purposefully designed to help students connect embodied schemas to concepts—we do not know whether the two mechanisms cancel each other out or if one is stronger than the other and consequently prevails. Hence, this conflict highlights a shortcoming of embodied action schema perspectives to explain the extent to which its mechanisms are independent of explicit connection making to concepts as pertaining to physical and virtual representations.

In addition, Fig. 1 highlights two aligning predictions. Both the physical engagement and the haptic encoding perspectives predict mode effects in favor of physical representations. However, because the haptic encoding perspective makes concept-specific predictions, this alignment only applies to physical representations that offer haptic cues for the given concept. The inspection of the reviewed articles showed that none of the selected articles tested whether physical representations yield higher learning outcomes *only* for those concepts for which they provide haptic cues. None of the reviewed studies inspected whether null effects could be explained by a physical representation's lack of haptic cues. Only one study came close to examining differential effects of haptic cues. Manches et al. (2010) found that haptic cues of physical representation modes afford different problem-solving strategies for partitioning tasks than virtual representations. Even though this was not directly assessed in their study, the authors argue that differences in problem-solving strategies could yield different conceptual outcomes. Hence, this alignment reveals a shortcoming of haptic encoding perspectives in that the concept-specific scope their predictions has not been empirically tested.

A second alignment exists between the haptic encoding and conceptual salience perspectives. As illustrated in Fig. 1, both perspectives predict that physical representations with haptic cues for a concept enhance learning. They do so for similar yet slightly different reasons. The haptic encoding perspective explains this advantage of physical representations with their

ability to provide direct, bodily experiences of the target concept. By contrast, the conceptual salience perspective suggests that physical experiences can draw students' attention to specific concepts—but this is just one of multiple features that could do so (e.g., visual cues of a physical representation could also make a concept salient). The inspection of the reviewed articles showed that arguments about haptic encoding and conceptual salience were often confounded. For example, Olympiou and Zacharia (2012) tested concept-specific effects but did not distinguish whether a representation provided haptic cues for the concept and/or made it more salient. Similarly, Wang and Tseng (2018) qualitatively inspected concept-specific effects. However, even though they used haptic encoding to motivate their research, their hypotheses about mode effects were purely based on conceptual salience, and they did not discuss whether haptic cues could account for concept-specific effects. Finally, Zacharia et al. (2012), Zacharia and Michael (2016), and Zacharia and Olympiou (2011) tested concept-specific effects based on conceptual salience, which included some reasoning about how haptic cues can make concepts more salient. However, their analyses did not distinguish effects that were due to a physical representation making a concept salient through haptic cues from those due to other characteristics such as affordances for interactions or visual features. Hence, this alignment highlights a shortcoming of conceptual salience perspectives, which have yet to compare the effectiveness of different types of conceptual cues. Future research could address this shortcoming by systematically comparing—for example—haptic cues, visual cues, and their combination, to test which has the strongest effect on students' learning.

Finally, Fig. 1 highlights that multiple mechanisms together may account for the effects of physical and virtual representations on students' learning. Rather than making predictions based on one theoretical perspective, a unified theory would suggest that a given representation mode is most effective if it engages students in multiple of the described mechanisms. For example, a physical representation may be more effective if it is designed to activate concept-specific embodied action schemas in addition to offering haptic cues for the concept, compared to a version of the physical representation that only provides haptic cues. Similarly, a virtual representation may be more effective if it draws attention to salient cues while also reducing cognitive load, compared to a version of the virtual representation that only focuses on reducing cognitive load. However, prior research has not tested whether these mechanisms are additive and whether or when they might interfere with one another. Hence, thus far, the suggestion that multiple mechanisms together may enhance students' learning is solely based on theoretical considerations.

Discussion

This review investigated what predictions different theoretical perspectives make about the effectiveness of physical and virtual representations (research question 1). To this end, I examined how prior research has motivated comparisons of representation modes, focusing on theoretical perspectives used to describe learning with physical, virtual, or blended representations. This revealed five theoretical perspectives that describe different mechanisms through which representation modes may affect learning and that yield different predictions for learning outcomes. Two of the perspectives make general predictions in favor of one or the other representation mode (i.e., the physical engagement perspective favors physical representations; the cognitive load perspective favors virtual representations). Three of them make

concept-specific predictions: one predicts a mode effect (i.e., the haptic encoding perspective favors physical representations for specific concepts), and two predict an effect not of representation mode but of the actions students use to manipulate them (i.e., the embodied action schemas perspective favors actions that invoke embodied schemas that are synergistic to the concept, and the conceptual salience perspective favors actions that draw students' attention to concepts).

An interesting contrast exists between the prevalence of the two perspectives that made concept-specific action predictions. On the one hand, the conceptual salience perspective is by far the most dominant perspective across the reviewed papers: it occurred in 42 of 54 reviewed articles. It was used as a sole perspective in ten papers and used equally frequently with the physical engagement, cognitive load, and haptic encoding perspectives. The dominance of this perspective reflects a consensus in the literature that neither physical nor virtual representations are superior but that their effectiveness depends on the concepts they illustrate. On the other hand, the embodied schemas perspective is the least dominant perspective: it occurred in 10 of 54 articles. It was never used as the sole perspective and about equally often used in conjunction with any of the other perspectives. The sparsity with which this perspective was used to motivate comparisons of representation modes is surprising for two reasons. First, this perspective is rarely used even though it matches the consensus that no mode is generally superior. Second, the infrequent use of this perspective in the context of comparisons of representation modes stands in contrast to research on the design of visual representations that has extensively made use of this perspective. One reason why the conceptual salience perspective is more prevalent than the embodied schemas perspective may be that the former describes explicit mechanisms through which students connect features of the representations to concepts, whereas the latter describes implicit mechanisms that do not require students attention or awareness. When multiple mechanisms may be at play, it may be easier to assess those that are more overt and accessible, for example, through verbal protocols. By contrast, implicit mechanisms are more difficult to assess because they are not verbally accessible.

The remaining perspectives were moderately frequent. The haptic encoding perspective occurred in 24 of 54 articles and was used as the sole perspective only twice. Most frequently, this perspective was used in conjunction with the conceptual salience perspective. This reflects the fact that the mechanisms described by the haptic encoding perspective match the mechanism described by the conceptual salience perspective because both rely on explicit processes. Similarly, the physical engagement perspective was used in 23 of 54 articles, three times as the sole perspective, and most frequently in conjunction with conceptual salience. Again, the mechanism is explicit because students can report on what features of a representation catch their attention, which aligns with the explicit nature of the conceptual salience perspective. Further, it is worth noting that—with the exception of articles that focused on blended representation modes—most articles contrasted this perspective to the conceptual salience perspective to illustrate that a pure physical engagement perspective is somewhat outdated because it does not account for concept-specific effects. Finally, the cognitive load perspective was used in 19 of 54 articles, was never used as the sole perspective, and was also most frequently used in conjunction with conceptual salience. The common co-occurrence likely results from the fact that both perspectives describe mechanisms through which visual features draw students' explicit attention to relevant concepts. In sum, in response to research question 1, this review shows that comparisons of representation modes are mostly motivated by a combination of multiple perspectives that describe mostly explicit mechanisms.

In addition, this review investigated whether predictions by the different mechanisms conflict or align with each other (research question 2). To this end, I compared predictions made by the given theoretical perspectives and examined whether research has empirically contrasted predictions that are based on the different theories. This comparison revealed three conflicting predictions. First, the perspectives differ in scope, that is, whether they predict general or concept-specific effects. While the literature has converged on concept-specific mechanisms, general mechanisms may still be at play when students work with representation modes. Second, physical engagement and cognitive load perspectives make opposite predictions. In the reviewed articles, they were often confounded, which complicates the interpretation of mode effects. Third, embodied schema and conceptual salience mechanisms often have opposite directions because actions that invoke synergistic embodied schemas often do not require explicit conceptual processing. None of the reviewed articles described studies that distinguished between these mechanisms, so that it is unclear how they interact with one another.

Further, the comparison revealed two aligned predictions. First, in cases where physical representations have haptic cues for a concept, the haptic encoding and physical engagement perspectives make identical predictions in favor of physical representations. Second, in cases where a haptic cue makes a concept salient, the haptic encoding and conceptual salience perspectives made identical predictions in favor of physical representations. None of the reviewed articles distinguished whether advantages of physical representations were due to physical engagement, conceptual salience, or haptic encoding mechanisms.

In sum, in response to research question 2, the comparison of the different theoretical perspectives suggests that multiple mechanisms co-occur while students learn with physical and virtual representations. This highlights a need for a unifying theory that would specify exactly how these mechanisms interact with one another. Specifically, in cases where the predictions conflict, we know very little about which mechanism outweighs the other and therefore explains the identified effects. In cases where the predictions align, we do not know whether the mechanisms are additive or which mechanism is the dominant one. A unifying theory would not only explain these interactions but also yield specific predictions for practitioners to effectively combine physical and virtual representations.

Implications for Research

This review has several implications for future research, summarized in Table 3. First, because multiple mechanisms can explain advantages of physical or virtual representations, researchers should be aware of potential confounds when comparing representation modes. Most prominently, this review showed that cognitive load mechanisms were often not controlled for, which makes it difficult to interpret differences between representation modes. For example, in light of prior research suggesting that physical representations can increase cognitive load because they are perceptually richer than virtual representations, research on blended representations should account for cognitive load effects when comparing blended to virtual representations. Overall, researchers should ensure their experimental designs do not confound different mechanisms that have been shown to account for effects of representation modes.

Second, because multiple mechanisms are likely at play when students learn with physical and virtual representations, future research should move beyond focusing on just one mechanism and address open questions about how multiple mechanisms interact, with an eye towards building a unified theory of learning with physical and virtual representations.

Table 3 Overview of implications for future research

Implication	Gap in prior research	Recommendation for future research
Confounds	Prior research has often confounded multiple mechanisms that might account for effects of representation modes	Future research should control for different mechanisms when comparing representation modes
Interactions among mechanisms	Prior research has mostly focused on a single mechanism of learning with representation modes	Future research should investigate how the different mechanisms of learning with representation modes interact with one another
Include implicit mechanisms	Prior research has mostly focused on explicit learning mechanisms pertaining to effects of representation modes	Future research should include implicit mechanisms (e.g., embodied action schemas) into research designs
Role of physical engagement	Prior research on blended representations takes a different view on physical engagement than prior research on physical/virtual representations	Future research should reconcile different views on the impact of physical engagement on cognitive learning outcomes

Specifically, this review identified cases where mechanisms align and cases where they conflict. When they align, open questions exist about whether the different mechanisms are additive. For example, is there added benefit to adding a haptic cue to a physical representation that already makes the target concept salient through visual features? When the mechanisms conflict, it remains unknown whether they cancel each other out or whether one mechanism is stronger and therefore prevails. For example, if a physical representation invokes a synergistic embodied schema but fails to make the concept salient, whereas the reverse is true for a virtual representation, which mode is more effective? Systematic comparisons of representation modes that isolate which learning mechanism accounts for learning advantages could resolve these questions. Blended representation modes offer new venues to isolate effects of particular physical and virtual features because they can strategically vary which aspects of students' interactions are physical or virtual.

Third, the dominance of perspectives that focus on explicit mechanisms ignores a large body of research on embodied action schemas, which suggests that implicit mechanisms also play an important role in students' learning with physical and virtual representations. Hence, research should take implicit mechanisms into account. To this end, research may assess implicit, nonverbal learning processes, for example, based on eye gaze or gestures.

Finally, this review revealed a gap in how research on physical/virtual modes and research on blended modes treats the physical engagement perspective. While the former casts this perspective as outdated, the latter uses it more optimistically to motivate the use of blended representations. This suggests that future research should examine the impact of physical engagement on cognitive learning outcomes.

Implications for Instruction

Instructors face a difficult choice when it comes to selecting appropriate representation modes for their students. Likewise, developers of blended educational technologies have to weigh multiple considerations when deciding which components should be presented in the physical or virtual mode. While this review showed that recent research appears to have reached a consensus that effects of representation modes are concept-specific, this does not invalidate prior research that established general mode effects through physical engagement and cognitive load mechanisms.

Further, multiple perspectives describe different types of concept-specific mechanisms. Instructors and developers should be aware of these perspectives to make educated choices. Until research addresses open questions about how the different mechanisms interact, these choices can only be based on a few heuristics, summarized in Table 4.

First, if a physical representation makes a concept more salient than a virtual representation and offers haptic cues for the concept, instructors and developers need to ensure that it does not increase extraneous cognitive load. For example, physical representations are often used in ways that require students to split their attention between the representation and other instructional materials. Further, they can distract students from relevant visual features because they may contain more seductive details than their virtual counterparts. If necessary, they could consider modifying how the physical representation is incorporated with other instruction so as to minimize extraneous cognitive load. For example, providing verbal instead of written instructions for using the representation can reduce split attention effects. Further, focusing students' attention on relevant visual features, for example, through verbal instructions or pointing gestures, may reduce the risk of students getting distracted by seductive details. If it is not possible to decrease cognitive load demands, the instructor has to weigh the intrinsic and germane cognitive load of the learning experience (e.g., difficulty of the material) against the increased cognitive load of the physical representation and against the possibility of losing its haptic and conceptual benefits if students instead used a virtual representation.

Second, if a representation has conceptual advantages, instructors and developers may want to ensure that it does not invoke antagonistic embodied action schemas (e.g., by requiring hand movements that invoke a conflicting schema). If it does, for a physical representation, they may consider making modifications so as to disrupt body movements that could invoke misleading schemas. For example, if a horizontal movement would be more favorable for a given concept, arranging materials so that students pick up pieces of the physical representation they are assembling with a horizontal rather than vertical movement might enhance learning. Similarly, interactions with a virtual representation could be modified so that students click buttons in a certain sequence that induces a horizontal movement.

Finally, purposefully combining representation modes may enhance students' learning because different learning mechanisms may be more or less relevant to particular content. In an instructional lesson, this can be done by pairing representation modes with specific content. In a blended learning environment, this can be done by manipulating which components are

Table 4 Overview of heuristics for instruction

Heuristic	Research basis	Recommendation for instruction
Reduce extraneous cognitive load for physical representations	Physical representations may increase extraneous cognitive load through seductive details and split attention	Reduce extraneous cognitive load by focusing attention to relevant visual features and through verbal instruction
Reduce antagonistic body movements for representations with conceptual advantages	Body movements that invoke antagonistic embodied action schemas may reduce learning	Modify representations so that students move bodies in synergistic ways or disrupt antagonistic movements
Purposefully combine physical and virtual representations	Different representation modes offer complementary advantages by engaging students in different learning mechanisms	Consider the representation's potential for physical engagement, cognitive load, haptic encoding, embodied action schemas, and conceptual salience for the target concept

offered in a physical or virtual mode. To this end, instructors and designers should carefully consider which concepts are best enhanced through physical and virtual representations, pair them accordingly, and switch between modes as these considerations change throughout the learning experience. Further, they may examine how students react toward physical experiences, for instance, whether they connect physical representations more readily to concrete experiences in ways that could help them learn. If this is the case, instructors and developers could strategically choose physical representations for concepts for which they also have haptic, conceptual, and embodied advantages. By contrast, virtual representations may be particularly helpful when reducing the risk cognitive overload is particularly important, for instance, when the content is particularly complex, when it is important to emphasize specific conceptually relevant features, or when they can activate embodied schemas. Nevertheless, whether or not a physical, virtual, or blended mode is most effective for the given learning content likely depends not only on the mode itself but also on the specific design of the given representation.

Limitations

This review article should be interpreted in the context of several limitations. First, it focused on cognitive learning outcomes. It did not take into account affective or motivational outcomes such as interest in the subject matter or enjoyment of the instructional activities. Yet, it is possible that representation modes that increase enjoyment also motivate students to interact with them more in the future, which could in turn affect long-term cognitive outcomes. Relatedly, this review focused on learning mechanisms that impact cognitive learning outcomes. Therefore, I did not take into account arguments about the practicality of physical vs. virtual representations, such as that virtual representations are more easily accessible than physical ones that are often more expensive. This review also did not take into account arguments pertaining to conventional practices, such as the fact that physical ball-and-stick models are commonly used in chemistry and that gaining experience with these physical representations is therefore a goal in and of itself. Third, this review focused on individual learning. Therefore, arguments about advantages of collaborating with shared physical resources were not taken into account. Finally, this review focused on STEM domains. It is possible that the role of representation modes for non-STEM fields, such as the use of interactions with physical artifacts and virtual experiences in domains such as history or arts, is fundamentally different than in STEM.

Conclusion

The availability of blended instructional materials that combine physical and virtual representations has drawn renewed attention to comparisons of representation modes. This review revealed that research generally considers five different mechanisms that make different predictions about when and why each mode is effective. Further, this review revealed specific cases when predictions align and conflict, which yields new directions for future research that can systematically investigate which learning mechanisms account for mode effects. Such research will yield important directions for instructors who face a practical decision about which representation mode to use for instructional activities. This review also showed that

research lacks attention to a large body of literature that has focused on implicit embodied mechanisms. While the embodied schemas perspective does appear in several of the reviewed articles, the research designs rarely reflect this perspective. Consequently, we know little about the interplay between implicit and explicit mechanisms. Finally, further research pending, this review yields some practical heuristics that may help instructors and designers of blended technologies to choose representation modes wisely and modify how they integrate them with other instructional materials.

References

Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 358–376). New York: Cambridge University Press.

Ainsworth, S. (2006). DeFT: a conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183–198.

Ainsworth, S. (2008). How do animations influence learning? In D. H. Robinson & G. Schraw (Eds.), *Current perspectives on cognition, learning, and instruction: Recent innovations in educational technology that facilitate student learning* (pp. 37–67). Charlotte: Information Age Publishing Inc..

Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245–266.

Antle, A. N., Corness, G., & Droumeva, M. (2009). What the body knows: exploring the benefits of embodied metaphors in hybrid physical digital environments. *Interacting with Computers, 21*(1), 66–75.

Atanas, J. P. (2018). Is Virtual-Physical or Physical-Virtual Manipulatives in Physics Irrelevant within Studio Physics Environment?. Athens Journal of Education, 5(1), 29-42.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

Bakker, S., Antle, A. N., & Van Den Hoven, E. (2012). Embodied metaphors in tangible interaction design. *Personal and Ubiquitous Computing, 16*(4), 433–449.

Bamberger, J., & diSessa, A. (2003). Music as embodied mathematics: a study of a mutually informing affinity. *International Journal of Computers for Mathematical Learning, 8*(2), 123–160.

Barrett, T. J., Stull, A. T., Hsu, T. M., & Hegarty, M. (2015). Constrained interactivity for relating multiple representations in science: when virtual is better than real. *Computers in Education, 81*, 69–81.

Baturo, A. R., Cooper, T. J., & Thompson, K. (2003). Effective teaching with virtual materials: Years six and seven case studies. In N. A. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), 2003 Joint Meeting of PME and PME-NA (pp. 299-306). Honolulu, Hawaii.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. L. (2012). Embodied cognition and learning environment design. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (pp. 198–223). New York: Routledge Taylor & Francis Group.

Bruner, J. S. (1966). *Toward a theory of instruction*. Belknap: Cambridge.

Burris, J. T. (2010). *Third graders' mathematical thinking of place value through the use of concrete and virtual manipulatives.* (Doctor of Education), University of Houston, Houston, Texas.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology, 105*(2), 380–400.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Chien, K. P., Tsai, C. Y., Chen, H. L., Chang, W. H., & Chen, S. (2015). Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education, 82*, 191–201.

Chini, J. J., Madsen, A., Gire, E., Rebello, N. S., & Puntambekar, S. (2012). Exploration of factors that affect the comparative effectiveness of physical and virtual manipulatives in an undergraduate laboratory. *Physical Review Special Topics-Physics Education Research, 8*(1), 010113.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

Clements, D. H. (1999). 'Concrete' manipulatives, concrete ideas. *Contemporary Issues in Early Childhood, 1*(1), 45–60.

Cuendet, S., Bumbacher, E., & Dillenbourg, P. (2012). Tangible vs. virtual representations: When tangibles benefit the training of spatial skills. *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (pp. 99-108). ACM.

Dackermann, T., Fischer, U., Huber, S., Nuerk, H. C., & Moeller, K. (2016). Training the equidistant principle of number line spacing. *Cognitive Processing, 17*(3), 243–258.

de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science, 340*(6130), 305–308.

Deboer, G. (1991). *A history of ideas in science education*. New York: Teachers College Press.

Dienes, Z. P. (1961). *The Dienes M.a.B. multibase arithmetic blocks*. London: National Foundation for Educational Research.

diSessa, A. A. (2014). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 44–62). New York: Cambridge University Press.

Doias, E. D. (2013). *The effect of manipulatives on achievement scores in the middle school mathematics class.* (Doctor of Education), Lindenwood University, Saint Charles, Missouri.

Dori, Y. J., & Barak, M. (2001). Virtual and physical molecular modeling: Fostering model perception and spatial understanding. Educational Technology & Society, 4(1), 61-74.

Drickey, N. A. (2000). *A comparison of virtual and physical manipulatives in teaching visualization and spatial reasoning to middle school mathematics students*. Logan: Utah State University.

Duijzer, C., Van den Heuvel-Panhuizen, M., Veldhuis, M., Doorman, M., & Leseman, P. (2019). Embodied learning environments for graphing motion: A systematic literature review. *Educational Psychology Review, 31*, 597–629.

Durmus, S., & Karakirik, E. (2006). Virtual manipulatives in mathematics education: A theoretical framework. *The Turkish Online Journal of Educational Technology, 5*(1).

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., et al. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics—Physics Education Research, 1*, 1–8.

Fjeld, M., Fredriksson, J., Ejdestig, M., Duca, F., Bötschi, K., Voegtli, B., & Juchli, P. (2007). Tangible user interface for chemistry education: Comparative evaluation and re-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 805–808). New York: ACM.

Flick, L. B. (1993). The meanings of hands-on science. *Journal of Science Teacher Education, 4*(1), 1–8.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3–4), 455–479.

Gibson, J. J. (1997). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hillsdale: Erlbaum.

Gire, E., Carmichael, A., Chini, J. J., Rouinfar, A., Rebello, S., Smith, G., & Puntambekar, S. (2010). The effects of physical and virtual manipulatives on students' conceptual learning about pulleys. In K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Proceedings of the 9th International Conference of the Learning Sciences* (Vol. 1, pp. 937–943). International Society of the Learning Sciences.

Glenberg, A. M. (1997). What memory is for: Creating meaning in the service of action. *Behavioral and Brain Sciences, 20*(1), 41–50.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment 25 years of cognitive psychology. *Perspectives on Psychological Science, 8*(5), 573–585.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences, 14*(1), 69–110.

Goldstone, R. L., Schyns, P. G., & Medin, D. L. (1997). Learning to bridge between perception and cognition. *Psychology of Learning and Motivation, 36*, 1–14.

Han, I. (2013). Embodiment: A new perspective for evaluating physicality in learning. *Journal of Educational Computing Research, 49*(1), 41–59.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena, 42*(1), 335–346.

Hayes, J. C., & Kraemer, D. J. (2017). Grounded understanding of abstract concepts: The case of stem learning. *Cognitive Research: Principles and Implications, 2*(1).

Horn, M. S., Crouser, R. J., & Bers, M. U. (2012). Tangible interaction and learning: The case for a hybrid approach. Personal and Ubiquitous Computing, 16(4), 379-389. https://doi.org/10.1007/s00779-011-0404-2.

Howison, M., Trninic, D., Reinholz, D., & Abrahamson, D. (2011). The mathematical imagery trainer: From embodied interaction to conceptual learning. In G. Fitzpatrick & C. Gutwin (Eds.), *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI 2011)*. Vancouver, CA: ACM.

Huppert, J., Lomask, S. M., & Lazarowitz, R. (2002). Computer simulations in the high school: Students' cognitive stages, science process skills and academic achievement in microbiology. International Journal of Science Education, 24(8), 803-821.

Huxley, T. H. (1897). Scientific education: Notes of an after-dinner speech. In T. H. Huxley (Ed.), *Collected essays: Science and education* (Vol. 3, pp. 111–133). New York: Appleton.

Jaakkola, T., Nurmi, S., & Lehtinen, E. (2010). Conceptual change in learning electricity: Using virtual and concrete external representations simultaneously. In L. Verschaffel, E. De Corte, T. de Jong, & J. Elen (Eds.), *Use of representations in reasoning and problem solving: Analysis and improvement* (pp. 133–152). New York: Routledge.

Jaakkola, T., Nurmi, S., & Veermans, K. (2011). A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts. *Journal of Research in Science Teaching, 48*(1), 71–93.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology, 106*(1), 86–104.

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology, 105*(2), 351–363.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2009). Concrete instantiations of mathematics: A double-edged sword. *Journal for Research in Mathematics Education, 40*(2), 90–93.

Katsioloudis, D. P., Dickerson, D. D., Jovanovic, D. V., & Jones, M. (2015). Evaluation of Static Vs. Dynamic Visualizations for Engineering Technology Students and Implications on Spatial Visualization Ability: A Quasi-Experimental Study. Engineering Design Graphics Journal, 79(1).

Kim, S. Y. (1993). The relative effectiveness of hands-on and computer-simulated manipulatives in teaching seriation, classification, geometric, and arithmetic concepts to kindergarten children. (Ph.D.), University of Oregon, Eugene, OR.

King, B., & Smith, C. P. (2018). Mixed-reality learning environments: What happens when you move from a laboratory to a classroom? *International Journal of Research in Education and Science, 4*(2), 577–594.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching, 44*(1), 183–203.

Lakoff, G. J., & Johnson, M. (1980). *Metaphors we live by.* Chicago-London: University of Chicago Press.

Lee, C. Y., & Chen, M. J. (2015). Effects of worked examples using manipulatives on fifth graders' learning performance and attitude toward mathematics. *Journal of Educational Technology & Society, 18*(1), 264–275.

Magana, A. J., & Balachandran, S. (2017). Students' development of representational competence through the sense of touch. *Journal of Science Education and Technology, 26*(3), 332–346.

Magruder, R. L. (2012). *Solving linear equations: A comparison of concrete and virtual manipulatives in middle school mathematics.* (Doctor of Education), University of Kentucky, Lexington, KY.

Manches, A., O'Malley, C., & Benford, S. (2009). Physical manipulation: Evaluating the potential for tangible designs. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (pp. 77-84). ACM.

Manches, A., O'Malley, C., & Benford, S. (2010). The role of physical representations in solving number problems: A comparison of young children's use of physical and virtual materials. *Computers & Education, 54*(3), 622–640.

Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. Paper presented at the CHI-98 Workshop on Hyped-Media to Hyper-Media.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E. (2009). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 31–48). New York: Cambridge University Press.

Mayer, R. E. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 131–152). New York: Cambridge University Press.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.

Melcer, E. F., Hollis, V., & Isbister, K. (2017). Tangibles vs. mouse in educational programming games: Influences on enjoyment and self-beliefs. In G. Mark & S. Fussel (Eds.), *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems* (pp. 1901–1908). New York: ACM.

Melcer, E. F., & Isbister, K. (2018). Bots & (Main) frames: exploring the impact of tangible blocks and collaborative play in an educational programming game. In R. Mandryk & M. Hancock (Eds.), Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 266). New York, NY: ACM.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97.

Montessori, M. (1966). *Secret of childhood*. New York: Ballantine Books.

Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments, 4*(3), 35–50.

Moyer-Packenham, P., Baker, J., Westenskow, A., Anderson, K., Shumway, J., Rodzon, K., & Jordan, K. (2013). A study comparing virtual manipulatives with other instructional treatments in third- and fourth-grade classrooms. Journal of Educational Computing Research, 193(2), 25-39.

Nathan, M. J., & Walkington, C. (2017). Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive Research: Principles and Implications, 2*(9), 1–20.

Nathan, M. J., Walkington, C., Boncoddo, R., Pier, E. L., Williams, C. C., & Alibali, M. W. (2014). Actions speak louder with words: The roles of action and pedagogical language for grounding mathematical reasoning. *Learning and Instruction, 33*, 182–193.

Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education, 96*(1), 21–47.

Ozgun-Koca, S. A., & Edward, S., T. (2011). Hands-on, minds-on or both? A discussion of the development of a mathematics activity by using virtual and physical manipulatives. Journal of Computers in Mathematics and Science Teaching, 30(4), 389-402.

Pan, E. A. (2013). *The use of physical and virtual manipulatives in an undergraduate mechanical engineering (dynamics) course.* (Doctor of Philosophy), University of Virginia, Charlottesville, Virginia.

Peirce, C. S., Hartshorne, C., Weiss, P., & Burks, A. (1935). *Collected Papers of Charles Sanders Peirce* (Vol. I-VI). Cambridge: Harvard University Press.

Pyatt, K., & Sims, R. (2012). Virtual and physical experimentation in inquiry-based science labs: Attitudes, performance and access. *Journal of Science Education and Technology, 2*(1), 133–147.

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing stem learning. *Educational Psychology Review, 29*(4), 717–761.

Renken, M. D., & Nunez, N. (2013). Computer simulations and clear observations do not guarantee conceptual understanding. *Learning and Instruction, 23*, 10–23.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Psychology Review, 7*(3), 216–237.

Scheckler, R. K. (2003). Virtual labs: A substitute for traditional labs? *International Journal of Developmental Biology, 47*(2-3), 231–236.

Schneider, B., & Blikstein, P. (2018). Tangible user interfaces and contrasting cases as a preparation for future learning. *Journal of Science Education and Technology, 27*(4), 369–384.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2016). Using mobile eye-trackers to unpack the perceptual benefits of a tangible user interface for collaborative learning. *ACM Transactions on Computer-Human Interaction (TOCHI), 23*(6), 39.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. Learning and Instruction, 13(2), 141-156. https://doi.org/10.1016/S0959-4752(02)00017-8.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (pp. 49-69). New York, NY: Cambridge University Press.

Schnotz, W. (2014). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). New York: Cambridge University Press.

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: a meta-analysis. *Educational Psychology Review, 30*(3), 679–701.

Segal, A., Tversky, B., & Black, J. (2014). Conceptually congruent actions can promote thought. *Journal of Applied Research in Memory and Cognition, 3*(3), 124–130.

Shaikh, U. A., Magana, A. J., Neri, L., Escobar-Castillejos, D., Noguez, J., & Benes, B. (2017). Undergraduate students' conceptual interpretation and perceptions of haptic-enabled learning experiences. *International Journal of Educational Technology in Higher Education, 14*(1), 1–21.

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research: Principles and Implications, 3*(6).

Skulmowski, A., Pradel, S., Kühnert, T., Brunnett, G., & Rey, G. D. (2016). Embodied learning using a tangible user interface: the effects of haptic perception and selective pointing on a spatial learning task. *Computers & Education, 92*, 64–75.

Smith, G. W., & Puntambekar, S. (2010). Examining the combination of physical and virtual experiments in an inquiry science classroom. In C. Z. Zacharia, C. P. Constantinou, & G. Papadourakis (Eds.), *Proceedings of computer based learning in science* (pp. 153–163). Warsaw: OEIiZK.

Stull, A. T., & Hegarty, M. (2016). Model manipulation and learning: fostering representational competence with virtual and concrete models. *Journal of Educational Psychology, 108*(4), 509–527.

Stull, A. T., Barrett, T., & Hegarty, M. (2013). Usability of concrete and virtual models in chemistry instruction. *Computers in Human Behavior, 29*, 2546–2556.

Stusak, S., Schwarz, J., & Butz, A. (2015). Evaluating the memorability of physical visualizations. In B. Begole, J. Kim, K. Inkpen, & W. Woo (Eds.), *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3247–3250). New York: ACM.

Suh, J., & Moyer, P. S. (2007). Developing students' representational fluency using virtual and physical algebra balances. *Journal of Computers in Mathematics and Science Teaching, 26*(2), 155–173.

Sung, Y. T., Shih, P. C., & Chang, K. E. (2015). The effects of 3d-representation instruction on composite-solid surface-area learning for elementary school students. *Instructional Science, 43*(1), 115–145.

Sweller, J., van Merrienboër, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.

Toth, E. E., Morrow, B. L., & Ludvico, L. R. (2009). Designing blended inquiry learning in a laboratory context: a study of incorporating hands-on and virtual laboratories. *Innovative Higher Education, 33*(5), 333–344.

Triona, L. M., & Klahr, D. (2003). Point and click or grab and heft: comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction, 21*(2), 149–173.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: a new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18*(1), 37–54.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45–69.

Wang, T. L., & Tseng, Y. K. (2016). The comparative effectiveness of physical, virtual, and virtual-physical manipulatives on third-grade students' science achievement and conceptual understanding of evaporation and condensation. *International Journal of Science and Mathematics Education*, 1–17.

Wang, T. L., & Tseng, Y. K. (2018). The comparative effectiveness of physical, virtual, and virtual-physical manipulatives on third-grade students' science achievement and conceptual understanding of evaporation and condensation. *International Journal of Science and Mathematics Education, 16*(2), 203–219.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636.

Wolfe, P. (2001). *Brain matters: Translating research into classroom practice*. Alexandria: Association for Supervision and Curriculum Development.

Yannier, N., Koedinger, K. R., & Hudson, S. E. (2015). Learning from mixed-reality games: Is shaking a tablet as effective as physical observation? *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1045-1054). ACM.

Yannier, N., Hudson, S. E., Wiese, E. S., & Koedinger, K. R. (2016). Adding physical objects to an interactive game improves learning and enjoyment: evidence from Earthshake. *ACM Transactions on Computer-Human Interaction (TOCHI), 23*(4), 21–31.

Yuan, Y., Lee, C. Y., & Wang, C. H. (2010). A comparison study of polyominoes explorations in a physical and virtual manipulative environment. *Journal of Computer Assisted Learning, 26*(4), 307–316.

Zacharia, Z. C., & Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the physics by inquiry curriculum: the case of undergraduate students' conceptual understanding of heat and temperature. *American Journal of Physics, 76*(4), 425–430.

Zacharia, Z. C., & Michael, M. (2016). Using physical and virtual manipulatives to improve primary school students' understanding of concepts of electric circuits. In M. Riopel & Z. Smyrnaiou (Eds.), *New developments in science and technology education* (pp. 125–140). Switzerland: Springer International Publishing.

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction, 21*(3), 317–331.

Zacharia, Z. C., Olympiou, G., & Papaevripidou, M. (2008). Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature. *Journal of Research in Science Teaching, 45*(9), 1021–1035.

Zacharia, Z. C., Loizou, E., & Papaevripidou, M. (2012). Is physicality an important aspect of learning through science experimentation among kindergarten students? *Early Childhood Research Quarterly, 27*(3), 447–457.

Zaman, B., Vanden Abeele, V., Markopoulos, P., & Marshall, P. (2012). Editorial: The evolving field of tangible interaction for children: the challenge of empirical validation. *Personal and Ubiquitous Computing, 16*(4), 367–378.