How to make a good animation: A grounded cognition model of how visual representation design affects the construction of abstract physics knowledge

Zhongzhou Chen^{**} and Gary Gladding

Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

(Received 28 May 2013; published 3 April 2014)

Visual representations play a critical role in teaching physics. However, since we do not have a satisfactory understanding of how visual perception impacts the construction of abstract knowledge, most visual representations used in instructions are either created based on existing conventions or designed according to the instructor's intuition, which leads to a significant variance in their effectiveness. In this paper we propose a cognitive mechanism based on grounded cognition, suggesting that visual perception affects understanding by activating “perceptual symbols”: the basic cognitive unit used by the brain to construct a concept. A good visual representation activates perceptual symbols that are essential for the construction of the represented concept, whereas a bad representation does the opposite. As a proof of concept, we conducted a clinical experiment in which participants received three different versions of a multimedia tutorial teaching the integral expression of electric potential. The three versions were only different by the details of the visual representation design, only one of which contained perceptual features that activate perceptual symbols essential for constructing the idea of “accumulation.” On a following post-test, participants receiving this version of tutorial significantly outperformed those who received the other two versions of tutorials designed to mimic conventional visual representations used in classrooms.

DOI: [10.1103/PhysRevSTPER.10.010111](https://doi.org/10.1103/PhysRevSTPER.10.010111)

PACS numbers: 01.40.Fk, 01.40.Ha, 01.50.fh, 01.50.ht

I. INTRODUCTION

From simple sketches of field lines to elaborate 3D animations of stellar systems, visual representations play critical roles in all of physics. Aside from providing visuo-spatial information, such as indicating the direction of forces or showing the topology of circuits, visual representations can have a substantial influence on conceptual understanding. A good visual representation can instantly clarify an otherwise confusing concept to students, while a bad one may lead to serious misinterpretations.

For example, the “energy bar chart” by Van Heuvelen and Zou [1] represents kinetic and potential energy of a mechanical system in the form of a bar chart, helping students “to ‘see’ conservation of energy” [Fig. 1(a)]. Students receiving the “energy bar chart” were observed to focus more on qualitative reasoning and improved significantly in problem solving. Computer enhanced visual representation has also been shown in various cases to improve students’ understanding of electricity and magnetism concepts [2–4].

Computer animations can also increase the validity of assessment under certain conditions [5].

On the other hand, many instructors have noticed that the conventional representation of electromagnetic waves [Fig. 1(b)] has likely contributed to the formation of certain misconceptions, such as the belief that the fields only exist in the regions formed by the field vectors and the sine curve [Fig. 1(b)] [6–8]. It is also well known that introductory physics students often have great difficulty correctly interpreting the canonical velocity vs time and acceleration vs time diagram [9].

But what are the key elements that distinguish a good visual representation from a bad one? How can we create effective representations such as the “energy bar chart” for other physics concepts, and avoid bad ones like the graph for electromagnetic (EM) waves? Are conventional forms of visual representations ideal for teaching physics, and if not, what needs to be changed?

These questions do not have easy answers, as it is impossible to simply look at a visual representation and intuitively judge whether it is good or bad, or identify which elements need to be changed to make it better. The way we interpret any visual representation is influenced so much by our physics expertise that our intuitive judgment has been explicitly shown to be largely untrustworthy, i.e., figures and graphs that look completely fine to us can turn out to be confusing to students [10,11]. In order to overcome this biased intuition and obtain a truly objective criterion for evaluating (and improving) visual representations, we need a

^{*}Corresponding author.
zchen22@mit.edu

Now at Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

FIG. 1. Examples of visual representations used in physics teaching. (a) Example of an energy bar chart. Reproduced from Ref. [1]. (b) Example of conventional representation of electric and magnetic wave.

cognitive model that describes in detail how perception of visual representations influences the brain in constructing conceptual understandings of physics knowledge.

Several cognitive models of learning from verbal and visual representations exist for teaching much simpler content, such as the well-known multimedia learning model developed by Mayer [12], and the more recent dual-coding model by Schnotz [13,14]. While these models are successful in the case of teaching encyclopedic knowledge, such as how storms form or how the time difference between different locations on the globe is calculated, on a closer look it seems that they still fall short of explaining how deeper conceptual understanding of abstract knowledge can be influenced by visual representations, which is critical for learning physics.

What makes learning physics knowledge so difficult to explain is that most physics principles, such as conservation of energy, are abstract logical arguments that by definition should not be associated with any particular form of visual representation. Nor do most physics concepts, such as energy and momentum, have any perceivable visual feature. Logical arguments, on the other hand, are thought to be more closely related to verbal representations, such as words and math expressions. In that case, why should any particular form of visual representation be superior to another in teaching an abstract idea? What gives a bar chart more illustrative power than a pie chart or any type of chart that carries the same amount of information in explaining the argument of “initial energy equals final energy”?

As a result of this lack of theoretical understanding, those successful innovations in visual representation design remain as isolated individual cases, as we are unable to obtain any generalizable knowledge from their success that can be used to develop effective new representations. This is becoming a more serious problem as computer animations and online courses are becoming an integral part of physics education. Provided with the unprecedented visualization capability of computer technology, we devote an increasing amount of resources and effort into creating new visual materials in the hope of improving learning outcomes. Yet with a limited understanding of the underlying

cognitive mechanism, it is possible that our new creations appeal much more to ourselves than to our students. Furthermore, there is no obvious way to tell whether the additional visual complexity inherent in a computer animation is truly beneficial to learning or simply serves as an unnecessary visual distraction.

In this paper, we present an attempt to overcome this theoretical difficulty, proposing a possible cognitive mechanism that describes how the visual features in a representation can have a significant impact on the brain’s understanding of abstract physics arguments, which are traditionally thought to be best conveyed by math expressions and verbal descriptions.

The key factor that differentiates our approach from existing theoretical models is that our approach is based on a new and emerging branch of cognitive psychology: grounded cognition. As will be discussed in detail in Sec. III, the prominent advantage of grounded cognition lies in its ability to provide a clear connection between external perception and internal cognitive structures, such as categories and concepts. It allows us to understand the role of visual perception in constructing conceptual understanding from a novel perspective, giving rise to a new objective criterion for judging the instructional effectiveness of a given design of visual representation.

In this paper, we put our theoretical framework for designing effective representations to a preliminary test by conducting a clinical experiment in which subjects received animated tutorials that are different only by the details of visual representational design. It turns out that despite the relatively small difference between the treatments, visual representations designed according to grounded cognition principles were able to substantially improve subjects’ conceptual understanding, as reflected by their performance on a post-test following the treatment. Notably, subjects receiving higher quality visual representations were much more likely to engage in qualitative reasoning rather than equation hunting during problem solving.

We begin by briefly reviewing some of the most relevant existing multimedia learning theories in Sec. II, and explain why they fall short in determining the pedagogical value of

a particular representation. In Sec. III, we introduce the theoretical framework behind grounded cognition, namely the perceptual symbol system (PSS) framework developed by Barsalou, which is followed by our insight on the impact of visual representations on learning abstract knowledge. We then describe our clinical experiment and its results in Secs. IV and V, followed by a discussion in Sec. VI.

II. MULTIMEDIA LEARNING THEORIES

In physics education research (PER), there have been relatively few studies on how the design of visual representations affects learning outcomes; and to the best of our knowledge, there hasn't been a theoretical account of how physics knowledge is constructed out of visual and verbal representations. On the other hand, a number of theoretical models exist in the field of multimedia learning that describe knowledge construction based on external visual and verbal representations. The best known is Mayer's multimedia learning theory (MML), which provides a set of "design principles" for creating animated multimedia tutorials. Multimedia instructional materials created following these principles have been repeatedly shown to be significantly more effective than those that don't.

On a very coarse scale, Mayer's MML theory can be viewed as being based on two basic ideas. First, students learn better when more cognitive resources can be used to make sense of the verbal or pictorial material. Second, when auditory and visual signals are presented in coherence, the cognitive resources required to process these signals are minimized. In that case, more cognitive resources can be allocated to making sense of the materials. His multiple design principles serve as detailed instructions for how to avoid incoherence between verbal and visual signals, which will incur unnecessary cognitive load on the brain. Mayer's principles offer important guidelines in designing multimedia instruction. However, we will argue that when it comes to teaching more complicated and abstract knowledge such as physics, these principles alone are not enough to guide the creation of high quality visual representations.

The most prominent shortcoming of Mayer's model, as has already been pointed out by some researchers [14,15], is that it lacks the ability to distinguish the quality between two different visual (or verbal) representations for the same content in terms of learning outcomes. In other words, the model cannot tell if a 3D computer animation of an electric field is better or worse than a sketch of field lines, or whether showing realistic images of a battery and wires is better than a conventional circuit diagram. According to Mayer's principles, as long as the visual representation is presented in coherence with verbal explanations (preferably audio narration) at a reasonable pace, students should have enough cognitive resources to make sense of the material. Mayer's principles, therefore, serve as guidelines for coordinating verbal and visual representations, rather than guidelines for designing each individual representation.

For Mayer, this shortcoming did not pose much of a problem, as most of Mayer's "multimedia instructional messages" are about concrete objects that have well-defined visual features, such as brakes and pumps, or simple processes that have well-established symbols, such as a red arrow indicating the rising of hot air. Therefore, any visual representational design that makes intuitive sense to the instructor will most likely make good sense to the learner as well, without resulting in misunderstanding. In addition, these messages are focused on describing the actual physical processes, rather than teaching the abstract principles behind those processes.

The choice of a visual representation to be used in teaching college level physics, however, is not at all obvious. As has been noticed by some PER researchers [11], some physics figures that make perfect sense to the instructor may turn out to be very confusing for students. In addition, most physics concepts such as energy or electric potential do not have any obvious visual features, meaning that there can be many plausible ways to visually represent those concepts. As is mentioned in the introduction section, different forms of visualization could potentially lead to very different learning outcomes.

Strictly speaking, even if we rigorously comply with Mayer's principles in designing multimedia materials, we still run the risk of presenting a misleading graph and a confusing piece of text to a student in perfect harmony with each other, since Mayer's principles did not provide guidelines for how to design each individual representation. As a result, even though students will have ample cognitive resources for sense making, many will end up with a wrong idea. In other words, Mayer's model alone could not have guaranteed the creation of effective new representations, nor can it predict the relative effectiveness between a bar chart and other possible types of representation such as a pie chart or a computer animation representing the same scenario. Therefore, although Mayer's model is very useful in guiding certain aspects of multimedia instruction design, it is not suitable for our purpose of improving the quality of conventional physics representations.

A noteworthy alternative to Mayer's theory is the multimedia learning model developed by Schnotz [13,14,16], who is able to provide one clear principle for the design of visual representations based on his model. Schnotz outlines a more detailed description of the structure of knowledge obtained from multimedia learning. In his model, knowledge consists of two separate parts: a mental model that encodes visuo-spatial information, and a semantic model that represents the logical relations between different objects and entities. Roughly speaking, the mental model is similar to a mental regeneration of the perceived visual representation, and the semantic model arises mainly from comprehending verbal explanation. Schnotz argues that in some cases of problem solving, the solver will need to "read off" relevant visual-spatial

FIG. 2. Different types of graphical representation used in the experiments by Schnotz [16]. (a) Carpet diagram of time difference. (b) Circle diagram of time difference. Reproduced from Ref. [16].

information by examining the mental model (similar to generating and examining a mental image), and then incorporate that information into the semantic model for logical reasoning. Therefore, he argues that the quality of any visual representation depends on the easiness to mentally read off information from the graph.

In his own example, Schnotz studied the difference between two graphical representations of geographical time difference: the carpet diagram and the circle diagram (see Fig. 2). In his experiment, each participant was given two types of problems to solve, the time difference task and the circumnavigation task, after studying either a single piece of text or a piece of text combined with one of the two graphical representations. A typical time difference task is “What time and which day is it in Los Angeles when it is Tuesday 2 o’clock p.m. in Tokyo?” whereas a circumnavigation task is “Why did Magellan’s sailors believe that they arrived on a Wednesday after sailing around the world, although it was already Thursday?”

It is easy to see that reading off relevant information for the time difference task is easier using the carpet diagram, whereas the circle diagram naturally lends itself to solving the circumnavigation problem. In the experiment, Schnotz found that providing a task-inappropriate graph significantly hinders problem solving performance. However, providing a task appropriate graph was not found to significantly improve problem solving when compared to providing text only.

While Schnotz’s model does provide some guidance to the design of visual representation, its application is limited to cases in which the exact same graph used as an example can be applied to solve every new problem. For example, the time zone graph contains all the essential information to solve each and every time zone question, so upon encountering any new problem, the solver only needs to mentally regenerate the exact same graph and read off relevant information from it. In that case, the difficulty of solving

the problem solely depends on the difficulty to memorize the graph and correctly read off the relevant information.

While this is true for time-zone questions since time zones never change, this is hardly ever the case for most physics problems. For example, in conservation of energy problems, re-creating the exact same energy bar chart (or even visually similar ones) from a previous example is clearly not helpful for solving even a slightly different problem (such as one with friction vs one without friction). On the other hand, once the correct visual representation has been generated, reading off corresponding information is mostly trivial. Furthermore, one could even argue that for many physics problems, if a student has enough knowledge to create the correct mental graph for the problem situation, then that same understanding should already be sufficient for solving the problem. In that case, mentally generating a figure or graph becomes largely unnecessary. (This speculation gets some support from the observation by Van Heuvelen and Zou that students seldom recreate the bar chart during an exam. [1].)

In short, the difficulty facing the Schnotz model when applied to physics is the following: most physics knowledge comes in the form of abstract rules, which according to Schnotz’s model is represented by the semantic network which comes almost entirely from language comprehension. In this model the quality of visual representations has little, if any, impact on the semantic model, and therefore should have little effect on understanding the majority of physics knowledge. In conclusion, this model is still insufficient for guiding the design of visual representations in teaching physics.

A. Difficulties facing existing theories of multimedia learning

Neither Mayer’s nor Schnotz’s model, nor any existing learning theory we know of, could satisfactorily explain how different designs of visual representations affect the

construction of abstract physics knowledge in the mind. We suspect that this theoretical difficulty is not just an isolated issue in either multimedia learning or PER, but rather is an instance of a more profound problem in mainstream cognitive psychology of the past century. Namely, for quite a long time, mainstream cognitive psychology has not been very successful in explaining how conceptual knowledge arises from perception in the first place [17].

Until recently, most cognitive psychologists believed that concepts are represented in the brain by so-called “amodal symbols”: abstract symbols that are being processed by specialized neural circuits of the brain, independent of all the sensory-motor systems. For example, upon seeing a chair, the perception of that particular chair activates a “CHAIR” symbol, which represents the semantic meaning of the word “chair.” This symbol is independent of the perception of any particular chair, so it could be used to represent the abstract idea of a general chair. In long term memory, the CHAIR symbol is connected to other symbols such as “BACK,” “LEG,” and “SIT” through propositional relations such as *HAS*(CHAIR, BACK), and *SIT* (PERSON, CHAIR). Notice that these propositional structures form the “semantic network” in Schnotz’s model, representing nonperceptual logical relation. Since amodal symbols are generally believed to have a very close correspondence with words in language, Schnotz proposed that semantic network arises mainly from language comprehension. (Although Mayer didn’t include a detailed description of knowledge structure in his MML model, the dual-channel structure of his model seems to suggest a similar divide between visuo-spatial and semantic knowledge.)

Cognitive psychologists believed that amodal symbols serve as the internal “language” used by the human brain to represent conceptual knowledge, especially abstract knowledge, and carry out various cognitive functions such as categorization and deduction. Sensory-motor systems, on the other hand, merely serve to translate various perceptions into amodal symbols, and control the perceptual and motor domains according to the instruction of amodal symbols.

However, although theories based on amodal symbols have been widely used to explain a variety of cognitive phenomena, how these symbols are initially established in the brain has always remained an unanswered question. Although they are thought to be closely related to words, most researchers agree that the meaning of the words have to initially arise through various perceptual experiences. However, since amodal symbols are thought to be independent of perceptual domains, how the brain translates perceptual experience into amodal symbols becomes a rather intractable question.

It is easy to see that such a shortcoming of the amodal symbol view has a particularly serious impact on the understanding of learning, since learning is essentially the process of generating new amodal symbols (or forming

new connections) from external perceivable representations. If the concept of “energy” is represented by the amodal symbol of *ENERGY*, then what types of perceptual experience do we need to provide to a student for such a symbol to emerge in their brain and be correctly linked to its various properties such as “*CONSERVED*”? Without an adequate understanding of this fundamental process, it seems hopeless to gain any insight into how different designs in the perceptual features of a visual representation can affect the learner’s understanding of abstract knowledge.

In recent years, there has been a growing skepticism in the field of cognitive psychology as to whether amodal symbols exist at all [18,19]. Accumulating new experimental evidence seems to suggest that the sensory motor systems of the brain are heavily involved in a wide range of cognitive tasks, such as categorization and logical reasoning, that were once thought to be carried out solely through following abstract rules coded in amodal symbols [19–25].

In light of this new evidence, some researchers have proposed to abolish the concept of amodal symbols altogether by arguing that that all human cognition is actually “grounded” in the sensory-motor systems of the brain. According to some researchers in grounded cognition, evolution has enabled the human brain to cleverly utilize existing neural circuits in sensory-motor domains, to perform more advanced cognitive tasks such as understanding math and language [17,19].

One of the most significant advantages of a grounded cognition view of knowledge is that the connection between perception and internal knowledge structures now almost comes for free, since they are both represented by activation of neural circuits in much the same sensory-motor domains. Therefore, grounded cognition serves as an ideal theoretical framework for understanding the impact of visual representational design on learning. In the next section, we will briefly introduce one of the most well-developed theoretical frameworks in grounded cognition, namely Barsalou’s PSS theory [17,18,20], and discuss how it provides us with new insights into the design of visual representation in teaching physics.

The main purpose of the next section is to provide the reader with some basic background knowledge in grounded cognition for understanding the rest of the paper. For a more comprehensive review of grounded cognition, including the limitations and challenges it faces, we recommend review articles by Barsalou [17,18,20] and others [26].

III. A BRIEF INTRODUCTION TO PERCEPTUAL SYMBOLS SYSTEM

A. Perceptual symbols

According to PSS, the most fundamental building blocks of knowledge are perceptual symbols: records of neural activation in the sensory-motor system as a result of

perception. For example, as you are looking at this paper, various neural circuits are activated in the visual domain of the brain in response to the perceived color, size, edges, vertices, lines, and surfaces, while the motor domain is actively controlling your hand to hold this paper, feeling its weight and texture (or holding the mouse if you're viewing it on the screen). At any moment, your attention is only focused on a certain aspect of the perceptual experience. For instance, you might be focusing on looking at a few words, or on the overall length and shape of the paragraph, or on the feeling of holding the paper (or mouse) in your hand, but not all of them at the same time. No matter what aspect of perception is being selectively focused by the brain, the neural activation patterns resulting from that perception will have a high chance of being recorded in long term memory. Such a record of neural activation is referred to as a "perceptual symbol" in PSS.

The most important assumption of the PSS theory (and grounded cognition in general) that differentiates it from amodal symbol theories of cognition, is that PSS assumes that all concepts, even the most abstract ones such as "truth" or "justice," are represented solely by reactivation of various perceptual symbols. In order to be able to represent complex and abstract concepts, perceptual symbols must possess a number of unique and critical features that differentiates them from holistic perceptions. First of all, perceptual symbols are not records of holistic perceptions, but rather a partial aspect of the entire perception that the brain's attention is focused on. For example, when looking at a polygon, the brain may record neural activations resulting from perceiving lines and vertices, but neglects other aspect such as size, orientation, color, or even the number of edges and vertices (causing one to forget whether it was a hexagon or polygon that was being observed). When perceiving a bar code, neural activities resulting from perceiving black and white stripes have a high chance of being recorded, whereas the number of stripes and the width of each line are likely neglected. Other than perception and motor function, perceptual symbols can also arise from introspection, the internal activities of the brain such as emotion and cognitive operations (compare, search, transform, elaborate). Perceptual symbols of introspection are particularly important to the understanding of abstract concepts such as "truth" [25].

Finally, perceptual symbols are not stored in isolation. Rather, the relation between different perceptual symbols reflects the statistical probability of the perceptions being made simultaneously. A large body of experimental results supports the correlation of perceptual symbols (for a review, see [20]). For example, Tucker and Ellis [22] showed that perceiving a cup handle would inadvertently invoke the brain to simulate a grasping motion, to the extent that it affects an unrelated motion of the hand such as pressing a button. Chao and Martin [23] showed through functional magnetic resonance imaging that perceiving

pictures of tools activates the motor areas controlling hand movement.

Being both partial and connected gives perceptual symbols a "symbol"-like property, which allows multiple perceptual symbols to be activated at the same time to form a so-called "simulation," which serves to represent the meaning of concepts.

B. Simulations

According to grounded cognition, the process of thinking about a particular cognitive entity such as an object, an event, or an idea, is essentially the process of activating a coherent set of perceptual symbols that are related to that entity. In other words, the brain simulates, to a certain degree of detail, a perceptual experience of that object or event, in the absence of direct perception. For example, when trying to decide whether a certain property belongs to an object, such as whether or not horns belong to horses, people will (mostly unconsciously) simulate the experience of looking at a horse to verify if it has horns. Neuroimaging experiments have shown that when performing these verification tasks, perceptual domains responsible for processing the features become activated (property "sweet" activates the taste domain, while property "horn" activates the visual domain). Therefore, under the PSS framework, this activated set of perceptual symbols is referred to as a simulation [17,27].

It is essential to point out that a simulation in PSS is significantly different from the common notion of a mental simulation (see, for example, [28,29]) in several important aspects. First of all, the term "mental simulation" is often used to refer to the process of consciously imagining concrete objects and processes, such as the function of mechanical devices like ratchets, sinks, or pumps [30] or processes such as the motion of stellar systems [31]. Simulations in PSS, on the other hand, could function unconsciously, probably being unconscious more often than conscious [27]. For example, Zwaan and Madden [24] showed that unconscious perceptual simulation underlies language comprehension. In their experiment, subjects reading the sentence "John pounded the nail into the wall" are faster at identifying the picture of a horizontal nail than that of a vertical nail, whereas the opposite is true for subjects reading the sentence "John pounded the nail into the floor." Barsalou [17] points out that research on skill acquisition has found that conscious awareness falls away as automaticity develops during skill acquisition, leaving unconscious mechanisms largely in control.

In addition, a mental simulation must contain enough details to form conscious mental imagery of objects and processes, and is often a similar regeneration of a previous experience. A simulation, on the other hand, could contain only skeletal components of a visual image, since it is able to function unconsciously. For example, an unconscious simulation of a triangle may consist of perceptual symbols

of its shape, but not its specific orientation. We know from neuroanatomy of vision that distinct channels in the visual system process different dimensions such as shape and orientation. Furthermore, it has been shown that when people construct conscious visual images, they construct them sequentially, component by component [17]. Also, due to the componential property of perceptual symbols, a simulation almost never precisely represents any particular previous experience. Rather, perceptual symbols activated to construct a simulation most likely come from a number of different previous experiences, as the selective activation of perceptual symbols is influenced by a number of different factors such as body states, emotion, and context (see, for example, Boroditsky *et. al.* [32]).

Furthermore, while mental simulations contain predominantly visual spatial information, simulations could also contain perceptual symbols from motor motion and introspection, such as cognitive operation and feeling. Therefore, mental simulations can only represent concrete objects and processes, but simulations could in principle represent all human concepts and thoughts, both concrete and abstract [25,33].

In many previous studies, the role of visual representations, especially animations, are thought to facilitate the learner in generating mental simulations of a concrete system during problem solving. In these situations the role of visual representations are being restricted to enhancing the learning of concrete knowledge such as how a mechanical device works, or how planets rotate around stars. However, based on grounded cognition, we suggest that a more important and more general role of visual representation in learning is to facilitate the generation of simulations that can represent both concrete and abstract concepts. We will discuss this idea in more detail in the next section.

C. Concepts

We store in our brain a very large number of interrelated perceptual symbols, which allows us to generate a great variety of different simulations. Whenever one is able to generate a variety of simulations about something, be it an object, a process, or an idea, to a socially acceptable degree, then one is thought to have mastered that “concept.” Therefore, the PSS definition of a concept is the collection of interrelated perceptual symbols that allows the mind to generate a large number of different but related simulations.

Barsalou [17] refers to this collection of perceptual symbols as a “simulator,” which suggests that the sole function of this structure is to generate simulations. Each simulation serves as an instantiation of that concept in the particular context in which the concept is being activated. A newly generated simulation never exactly resembles a single previous experience, but rather includes perceptual symbols acquired from multiple previous experiences, which allows us to easily think about imaginary and fictitious entities that we had never experienced before.

Because of this, knowledge acquired from a single visual representation does not have to be restricted to that particular representation, but is able to transfer to other similar situations.

IV. CONSTRUCTING PHYSICS KNOWLEDGE FROM REPRESENTATION

How exactly does the PSS framework help us understand the impact of visual representation on learning abstract physics concepts?

First of all, PSS provides a more precise definition of “learning.” According to PSS, one is said to have “acquired a concept” if they are able to generate simulations about an entity or process to a “culturally acceptable degree” [17]. In our case, this means to be able to simulate physics concept to a degree that allows one to reason like an expert. Therefore, the process of learning can be viewed as training the brain to generate new simulations. This can be achieved either by acquiring new perceptual symbols through direct perceptual experience or by combining existing perceptual symbols or smaller scale simulations in a novel way.

Since the majority of physics concepts, such as force, energy, and electric field, are not directly perceivable, learning a physics concept predominantly involves activating existing perceptual symbols and combining them in novel ways to form new simulations. This would suggest that the goal of any type of instructional representation used in teaching physics, be it text, graph, or animation, should be to correctly activate a set of existing perceptual symbols, so that the brain can combine them into a new simulation. The question then becomes “how exactly can instructional representations activate perceptual symbols?”

The most common method, of course, is to talk about it, or more precisely, to present a representation that has already been associated with an existing simulation consisting of multiple perceptual symbols. For example, the word “cup” has already been associated with the simulation of a cup, therefore perception of either the written form or the aural form of the word activates perceptual symbols associated with the object cup; to be more precise, a symbol is associated with a concept, and the exact simulation that the concept generates upon perceiving the symbol is influenced by a number of background variables, such as contextual and epistemological factors. We refer to this method as the “symbolic method,” since the representation serves as a symbol for its meaning. The proper functioning of the symbolic method relies entirely on our previous knowledge about the symbol and the existing convention, e.g., this paper can only be understood by an English speaking person, who can generate correct simulations for every word.

Visual representations also utilize the symbolic method. For example, in circuit diagrams we use symbols to represent real circuit elements; in phasor diagrams we use arrows to represent the change in voltage and current in

an ac circuit. Interpreting these figures not only requires knowledge about the represented objects (circuit elements and forces), but also the convention that maps the symbols to the objects as well. Some forms of visual representation, such as the phasor diagram, require a considerable amount of learning of the underlying convention before a student can successfully interpret its meaning.

A distinct feature of the symbolic method is that detailed perceptual features of the representation have little to do with the meaning that it stands for. The word “cat” by no means resembles the look or feel of a real cat, and could be written in any color, in any recognizable font, yet still corresponds to the same concept. If the symbolic method were the only method of activating perceptual symbols, then the design of new visual representations would involve little more than arbitrarily creating a piece of drawing, explicitly stating that it represents a certain concept, and strictly following that convention thereafter. Furthermore, it would be rather pointless to devise any new visual representations, since well-defined conventional visual representations already exist for every physics concept. Yet we know from experience that this cannot be further from the truth.

The most important insight we are able to gain from adopting the grounded cognition framework, is that the perceptual features of a visual representation, such as color and shape, can directly activate the perceptual symbols via a different mechanism which we will refer to as the “perceptual method.”

The perceptual method is based on the simple fact that, in order to understand the meaning of any representation, be it visual or verbal, the learner must first perceive that representation using the perceptual domains of the brain. Yet at the same time, according to grounded cognition, the “meaning” behind that representation is also being processed in much the same perceptual domains of the brain. Therefore, the perceptual features of the representation must have a significant chance of influencing the perceptual symbols being activated.

There are abundant examples of (seemingly irrelevant) perceptual features of representations interfering with the comprehension of their meaning. The most well-known example is the Stroop effect [34,35]: the perceived color of the ink used to write color words, such as “red” written in blue ink and “blue” written in red ink, significantly affects the processing of their semantic meaning. Similar effects have been shown for symbols representing abstract concepts, such as numbers. Henik and Tzelgov [36] showed that the perceived size of written numbers affects people’s judgment of their represented magnitudes. For example, if the number 3 is written in a bigger font than 5, people are slower and more error prone in judging which number is greater. More recent research [37,38] seems to suggest that this is because the brain area involved in processing the

magnitude of the numbers partly overlaps with the area processing their perceived size.

If the perceptual features of a representation are designed so that they activate the same neural circuits that are involved or closely related to the processing of its intended meaning, such as writing a numerically smaller number in a perceptually smaller font (both activating the same neural circuit representing “smaller”), it is likely to result in a constructive interference between the perceptual method and the symbolic method. From this point of view, the energy bar chart, for example, may have been superior because it provides the visual perception of invariance, which probably activates the same neural circuits representing the more abstract concept of “conservation.” In that case, both the symbolic method and the perceptual method activate the same perceptual symbol of conservation, which can be integrated into the simulation of energy, resulting in a correct simulation of “energy conservation.”

On the other hand, a problematic visual representation contains perceptual features that are in conflict with the meaning that it was intended to represent, resulting in destructive interference between the two. For example, the conventional representation of EM wave displays a repeating pattern with very obvious spatial boundaries, yet was intended to represent electric and magnetic fields that fill infinitely large planes. As a result, the perceptual features of the graph accidentally activate a “spatially confined” perceptual symbol, while the verbal explanations of EM waves tries to activate the “extending to infinity” perceptual symbol.

In fact, interference between perceptual features of representations and their represented abstract meaning has already been demonstrated through a series of studies by Landy and Goldstone on math calculation [39–41]. For example, they have shown that our understanding of the order of precedence in math calculation, which is nothing but a set of abstract rules, very likely shares much of the same brain system responsible for detecting visuo-spatial proximity. For example, $2 + 3 \times 5$ is calculated faster and more accurately than $2 + 3 \times 5$. Furthermore, research on the use of concrete vs abstract symbols [42–46] has also demonstrated that including visual features that are irrelevant to the topic, such as using overconcrete symbols, can also hinder learning of abstract relations.¹

Aside from direct perception, the perceptual method may also function through so-called “perceptual inference” [20,27]. As previously mentioned, perceptual symbols

¹Some of the research (Uttal *et al.* [47]) suggests that the negative impact of concrete symbols on learning may be due to the fact that learners are more likely to treat a concrete symbol as an object in its own right, rather than a representation of an abstract concept. This is a mechanism that is different from the one described in the current paper. It is an interesting question as to which mechanism is dominate, but discussion of this issue is out of the scope of this paper.

FIG. 3. Example of activating perceptual symbols of “stronger” and “weaker” via perceptual inference using color saturation and color gradient. (a) More saturated and redder color induces a feeling of being stronger, while a more transparent and bluer color seems weaker. (b) Perceptual symbols representing “increasing magnitude” activated by a combination of color saturation, color gradient, and line thickness, in the context of electric potential distribution.

are closely associated with each other based on the statistical frequency of co-occurrence in the physical world. Therefore, making one perception can activate a number of closely associated perceptual symbols. This method of activating perceptual symbols is particularly useful in teaching physics, as most physics concepts are invisible.

For example, the relative magnitude of *electric potential* distribution in space is not directly observable. However, perceptual symbols related to “stronger” and “weaker” could be easily activated through perceptual inference. For almost everyone, saturated color seems stronger than transparent color [Fig. 3(a)], and a thicker line looks stronger than a weaker line [Fig. 3(b)]. Therefore, if we represent electric potential in space around a charged particle using both color saturation and line thickness of equipotential surfaces, both color and thickness serve to activate the stronger and weaker perceptual symbols, creating a visual representation of electric potential that looks stronger closer to the charge [Fig. 3(b)].

Compared to the symbolic method, which relies heavily on the learner’s previous knowledge, it is our hypothesis that the perceptual method should turn out to be relatively much more stable and precise, since perception is a more fundamental function of the brain. Even when perceptual inference is involved, the perceptual method should still be much less subjected to differences in knowledge background. This is because the associations between perceptual symbols are established through our direct experience with the physical world, which is largely invariant between different individuals (e.g., handles are universally subjected to be grasped by hand). Even though in some cases perception can also be significantly influenced by prior knowledge, in general it is still much more invariant

between individuals when compared to our interpretation of symbols. For example, although people speaking different languages are known to have different perceptions of color [48] (and presumably there should be differences in other types of perception as well), they should still be able to communicate relatively well with each other through sketching and gesturing, while being completely unable to understand a word of each other’s language.

However, this does not mean that the functioning of the perceptual method is not subjected to any expert-novice differences. In fact, we think that although the perceptual symbols activated by the perceptual method are largely the same, the way the brain selects and incorporates those perceptual symbols into the final simulation can be very different between experts and novices.

More precisely, as experts have much more comprehensive and robust content knowledge, they are able to quickly and reliably identify meaningful patterns in the representation (often without conscious processing), and exclude irrelevant perceptual features from further processing [49–51].

Novices, on the other hand, often lack the proper background knowledge to help them decide which perceptual features are essential to the content and which ones are irrelevant.² As a result, they often end up focusing on a number of conflicting perceptual symbols, which either prevents them from constructing a new simulation or causes them to construct an incorrect simulation of the concept being taught.

Learning difficulties among novices caused by defects in perceptual features of a representation often surprise experts, as experts tend to subconsciously ignore those features in the first place. We think that some of the common learning difficulties, such as equation hunting and rote learning, which are often attributed to more general reasons such as epistemological belief or study habits, might actually (at least partly) be caused by defects in the design of visual representations that are being overlooked by the instructor. In other words, we think that in some cases it is not that the students are unwilling to gain a deeper understanding of physics, but that they are actually unable to do so due to interference from the perceptual method.

A. The quality of instructional representation

These new insights on learning from representations enable us to give a theoretically more precise and practically more useful definition of the quality of any instructional representation, especially visual representations.

From a grounded cognition perspective, the goal of any instructional representation is to activate a correct set

²It is well known that when asked to sort physics problems according to similarity of solutions, novices sort by surface feature similarity and not the physics principles needed for the solution [52].

of perceptual symbols from which the learner is able to construct an expertlike simulation of the subject. At the same time, however, any representation will also inevitably activate certain irrelevant perceptual symbols (such as the size and color of the font in written text), that should not be included in the final simulation. If learners fail to properly filter those irrelevant perceptions, they will interfere with the construction of a final useful simulation, leading to failed learning. Therefore, we think that a meaningful way to evaluate the quality of a piece of an instructional material is the difference between the amount of relevant perceptual symbols it is able to activate in the learner's mind and the amount of irrelevant perceptual symbols it brings along. Simply put, the quality of a representation depends on the "net" amount of useful perceptual symbols it activates.

It is easy to notice that under such a definition, the quality of instructional material depends not only on the design of the material itself, but also on the background of its potential audience, as both the activation and the filtering of perceptual symbols depends not only on the learner's content knowledge, but also on more general aspects such as beliefs, attitude, and emotional states.

Furthermore, since perceptual symbols are activated via the symbolic method and the perceptual method separately, we could further define the symbolic (s) quality and perceptual (p) quality for a single piece of representation, as the net amount of perceptual symbols that could be activated by each method separately. The s quality of a representation depends on the learner's familiarity with the conventions used in it, for example, representation " $5 > 3$ " is of high s quality among people using Arabic numbers, but the s quality of "five greater than three" is only high among English speaking learners. The p quality of a representation depends on whether the perceptual feature of the representation agrees with the intended meaning, so " $5 > 3$ " is of lower p quality than " $5 > 3$." Generally speaking, the p quality of a representation is less dependent on individual differences among learners, since all of us use the same perceptual organs to perceive the same physical world.

Of course, such a definition of representation quality cannot be quantitatively measured, as the PSS framework behind it is a conceptual construct rather than a computational model. The most prominent value of introducing the concepts of s quality and p quality, lies in the fact that it highlights the perceptual aspect of representation design, by explicitly separating it from the symbolic aspect. Separating the two aspects of representation design readily resolves the theoretical difficulty of explaining the effectiveness of certain innovations in physics representation design.

Furthermore, the idea of p quality serves as a useful guideline to instructors seeking to improve existing representations or creating new representations of higher quality. As mentioned above, it is not that experts cannot perceive

conflicting perceptual features in a representation, or that perceiving those features do not activate conflicting perceptual features in experts' brains, but rather that the experts' minds are trained to automatically ignore these perceptual features. By explicitly calling for attention on the perceptual aspect of a representation, experts will be more likely to consciously notice conflicting perceptual features in their design, and find ways to improve the p quality.

In the rest of this paper, we report a clinical experiment in which we try to demonstrate the significance of the impact of p quality on learning outcome. Namely, we show that by increasing only the p quality of visual representations being used, we are able to achieve a sizable improvement on students' performance on a post-test, even if students are only exposed to the representation for a relatively short amount of time (~6 minutes).

V. EXPERIMENT

The clinical experiment presented here serves as a "proof of concept," demonstrating that by simply improving the p quality of the visual representation alone, learning outcome from a 5-minute multimedia tutorial can be significantly improved.

A. Topic: Electric potential difference

The physics concept that we chose for this experiment is electric potential difference. For any electric field created by a certain charge distribution, the electric potential difference between any two points in space, A and B, is defined as the negative integral of the electric field from one point to the other along any path. Mathematically, the potential difference between A and B, $V_A - V_B$, can be written as

$$V_A - V_B = \int_B^A -E \cdot dl.$$

Here, E is a vector function describing the electric field in space.

Students taking introductory level calculus-based college physics are frequently observed to have multiple difficulties with this expression, such as determining the limits, path, and sign of the integral. The most common difficulty is that students often chose to calculate the potential difference for any circumstance using the following equation:

$$V = k \frac{Q}{r}.$$

This equation is applicable only for a number of simple spherically symmetric cases in which the zero-potential reference point is set at infinity.

For the case shown in Fig. 4, where a conducting shell surrounds a point charge, students would often

FIG. 4. A thick conducting shell surrounding a charged particle.

incorrectly calculate the potential difference between point A and B as

$$V_A - V_B = k \frac{Q}{r_A} - k \frac{Q}{r_B}.$$

There are several possible explanations for this behavior. First of all, students may have a very limited conceptual understanding of electric potential, and as a result, adopt an equation hunting strategy. Since the aforementioned equation is mathematically simpler than the integral expression, it is a popular choice among equation-hunting students. Another possible reason may be that students treat electric potential as a local property that is created by the local electric field at that point, $k(Q/r^2)$, in much the same way as the mass of an object results in its weight; therefore, comparing the potential difference between two points is like comparing the weight difference of two objects, which should not depend on anything in between the two objects. The idea that potential difference is defined by accumulation of the electric field along a path is hard to get across to students. Even when students do use the integral expression to calculate the potential difference, they often have difficulty setting up the integral. In cases where the electric field distribution in space is discontinuous, such as in the presence of a conductor, many students fail to notice that they need to break the potential integral into different parts according to the electric field distribution. Others have difficulty in determining the limits to use for the integral, and often use infinity or zero as one of the limits regardless of the actual problem context. It is likely that students have memorized some superficial details of the mathematical expression from one or more previous example solutions (learned by rote), and simply “pasted” those details into the solution for the new problem, without fully understanding the physics meaning behind the math expressions.

In short, many students lack an adequate conceptual understanding of the concept of electric potential. In PSS terms, it means that they have not learned how to generate a functional simulation representing “potential difference

as accumulation of electric field in space” for common physics situations. This happens even after they have gone through all the instruction and practice on the topic in their physics class, which often includes lecture, homework, experiment, and discussion sessions.

We believe that this difficulty at least partly originates from the fact that conventional visual representations used in the teaching of electric potential have particularly low *p* qualities for students at this level. Therefore every time they are exposed to the material, instead of gaining more understanding on the subject, the conflicting perceptual features listed in the following section results in more confusion. So some students may have simply memorized more superficial surface features to compensate for the lack of understanding.

In the following section we will list in detail various conflicting perceptual features in the visual representations typically associated with the teaching of the concept, and demonstrate how the *p* qualities can be improved following grounded cognition principles. Since we are mainly interested in students’ understanding of the idea of integral as an accumulation process, not their mathematical ability to set up and carry out the integral for different geometry, we will restrict our study only to cases with spherical symmetry.

B. Visual representation design

1. Conflicting design in conventional representation

An example of some conventional visual representations frequently used in introducing the integral expression of potential difference near a point charge is shown in Fig. 5. The examples shown here are highly similar to the representations used in a number of popular physics textbooks [53–55]. Upon a closer inspection, it is easy to identify a number of perceptual features in these representations that are in conflict with the idea of “accumulation.”

First of all, the electric field in space is represented using field lines, which perceptually looks like a holistic, continuous substance that extends throughout the entire space. A simulation of the concept accumulation, on the other hand, most likely contains perceptual symbols obtained from experience with substances such as sand, money, or water that can be broken into smaller components and accumulated. The perceptual features of the holistic field line representation are in direct conflict with these perceptual symbols. Therefore, the perceptions of these holistic field lines need to be suppressed and filtered, in order for the brain to activate the proper perceptual symbols (fragmented smaller pieces) for constructing a simulation for accumulation.

Second, when illustrating the electric potential difference between two points, the conventional visual representation is simply two dots labeled with V_A and V_B . The dots and the labels at the two locations are visually very similar with the same size and color. This perception of uniformity is in direct conflict with the idea of a “difference” between the two locations.

FIG. 5. Visual representations frequently used in teaching electric potential. (a) Line and arrow representation of electric field created by a point charge Q . (b) Electric potential between two points indicated by crosses and labeled V_A and V_B , respectively. (c) Snapshot of an animation representing the integral from point B to point A . The integral is represented by a moving arrow going through multiple points.

Third, when the integral expression of electric potential difference $\Delta V_{AB} = \int_A^B -E \cdot dl$ is introduced, a line with an arrow is drawn from one point to another.³ More careful instructors would emphasize that dl stands for a small segment on which the electric field is treated as constant, and the integral means dividing the path from A to B into an infinite number of infinitesimal small pieces of dl , and adding up the electric field along each small segment of dl . In that case, the accompanying visual representation is usually an arrowed line going through a series of multiple points along the path.

Neither width nor color of the line changes between the starting point and the ending point, displaying no perceptual feature that may suggest an increase (or change) in magnitude or strength as the result of an accumulation process. In addition, the entity that is being accumulated, the electric field, is not visible in this form of visual representation. In other words, the arrowed line representation looks nothing like an accumulation process, and therefore should interfere with the generation of a simulation of accumulation on the students' side.

In short, the three conventional representations have rather poor p quality, as they contain very few perceptual features that can be related to either the idea of accumulation or the difference in electric potential caused by accumulation process, and on the contrary, contain several perceptual features that suggest the opposite. According to grounded cognition, these conflicting perceptual features must be processed in largely the same brain areas that are also responsible for interpreting the verbal and mathematical description of "accumulating electric field," resulting in frequent destructive interference between the symbolic method and perceptual method. We believe that undergraduate level students viewing these representations have much difficulty activating the correct set of perceptual symbols to

construct a proper physics understanding, and as a result, often turn to memorizing superficial surface features of the verbal mathematical expression, such as the limits used in the integral or the final math expression of $kq/r_A - kq/r_B$.

2. Perpetually enhanced design

We introduce here a new design of the visual representations used in teaching electric potential with much improved p quality. The design follows a simple principle: the representation should visually look like what we want to teach our students as much as possible, or in more technical terms, perceptual features in the representation should be designed to facilitate the activation of those perceptual symbols that are essential for constructing a proper simulation of the content. For example, when we talk about potential difference in space, the corresponding visual representation should visually look very different at different points in space, so that the visual perception facilitates rather than interferes with the simulation of difference.

We redesigned all three representations mentioned above: the "holistic" electric field, the "uniform" electric potential difference, and the "straight line with arrow" integral. Starting with the potential difference, we chose to represent the concept with equipotential surfaces of different color and thickness, as shown in Figs. 7. The reason we replaced the original "dot" representation with colored surfaces (or colored rings in 2D) is simply because small changes in thickness and color on different rings are much more visually salient than the same change on dots. This allows us to depict a small gradient of change in thickness in a limited space without having some of the rings being so thick that they overlap with each other. We use both the line thickness and color saturation to further enhance the perception of "difference" and "change" between the rings, to the extent that it is impossible for a normal person to overlook, as demonstrated in Fig. 7.

It should be pointed out that the direct perception of neither color nor thickness is relevant to the concept of

³To eliminate unwanted complexity, we will restrict our discussion to the case in which the two points are placed along the radial direction, which means that the most direct path for the integral is a straight line between the two surfaces.

FIG. 6. Snapshot of perceptually enhanced representation depicting potential difference and electric field (see also Fig. 7). The difference in color and thickness between the two rings represents the potential difference between locations A and B (see also Fig. 7). The radiating black arrows represent the local electric field between the two rings.

electric potential. As experienced university instructors, we feel very confident in saying that students in a calculus based electricity and magnetism course should know very well that equipotential lines, an imaginative concept invented for describing properties of electromagnetic fields, have neither thickness nor color of their own. This background knowledge should, according to PSS theory, allow the learner to pay much less attention to these aspects of the perception. As a result, the precise color and thickness of the rings are less likely to be stored as part of the concept in LTM. Even if they are being (unintentionally) stored in LTM, and later activated by as part of a new simulation generated upon encountering a new situation, that background knowledge should be sufficient to allow the student to identify these features as irrelevant. Put in layman's language, we are confident that a college student should always be able to recognize the question “what is the color

(or thickness) of this equipotential line” as being meaningless, even if they were shown colorful equipotential lines during instruction. (In fact, one cannot show a “colorless” representation, as it would be invisible.)

As potential differences are represented by colored rings, the local electric field could naturally be represented by radiating arrows in between those rings. The density and gray scale of the arrows roughly correspond to the average magnitude of the local electric field [compare between Figs. 7(i) and 7(ii)]. Compared to the holistic field line representation, this piecewise representation visually resembles a fragmented, “accumulatable” substance.

The introduction of the accumulation process and its integral expression is accompanied with computer animation showing colored equipotential rings appearing one after the other at the same pace of a moving black arrow indicating the path of integral (Fig. 7). Local electric field between each of the two rings appears briefly as the arrow moves from the previous ring to the next ring, appearing as the substance of accumulation. This animation, together with a gradient of increasing line thickness, color saturation (the amount of white in the color), as well as a color shift from blue to red, provide strong visual perception of accumulation that is impossible to miss by the eye.

One detail that is worth mentioning for this “perceptually enhanced” design is that the equipotential lines are separated equally in space, in order to represent the equal step size taken by the integral (dr), as well as creating a better visual perception of a gradual change in magnitude. In conventional physics representation, equipotential lines are usually drawn at intervals representing equal potential differences, i.e., at a decreasing spatial separation from outside to inside in the case of a spherically charged object. In the final multimedia tutorial presented to the participants, the equal potential representation will also be introduced and explained to participants.

3. Control for spatial information design

As the purpose of this experiment is to test if the difference in p quality of visual representations alone

FIG. 7. Snapshots of the perceptually enhanced animation for the integral process.

FIG. 8. Example representation of electric potential designed in control for the extra information design. The black arrow moves from point *B* to point *A*. Radiating black arrows representing the local electric field between two different rings appear briefly when the tip of the moving arrow touches the corresponding ring.

can cause a measurable difference in learning outcomes, ideally we want the different versions of visual representations used in the experiment to differ only in *p* quality and nothing else. However, replacing the dots in the conventional representation with equipotential surfaces also introduced additional spatial information into the perceptually enhanced representation as a byproduct. Although none of this additional information is relevant to answering any of the problems given as assessment, we still felt it necessary to create a third version of design that is informationwise exactly the same as the perceptually enhanced version.

This third type of visual representation is created by keeping all the equipotential surfaces in the perceptually enhanced design, while removing as much as possible the visual elements that contribute to the improvement in *p* quality, such as color, thickness, and the animation on the rings. The piecewise electric field representation and the animation of the moving arrow are both kept the same to avoid the representation from being visually too complicated. An example of the end product is shown in Fig. 8.

C. Outline of experiment

The procedure of the experiment is quite straightforward (Fig. 9). Participants were first asked to complete an online pretest before coming to the experiment. During the experiment, participants are divided into three groups, with each group being presented with a different version of the online tutorial using one version of the visual representation design (see below for details). The tutorial consists of two parts, with each part followed by a number of assessment problems. At the end of the experiment, students were asked to evaluate the effectiveness of the tutorial they received on a scale of 1 to 10.

FIG. 9. Procedural outline of the experiment.

D. Multimedia tutorials

We chose to present the instructional materials to the participants in the form of online multimedia tutorials: computer animation accompanied with audio narration. Three different versions of multimedia tutorials were created, each using a different visual representation design: The *control* (*Ctrl*) version uses the conventional design, the *experimental* (*Exp*) version uses the perceptually enhanced design, and the *control for information* (*CInf*) version uses the control for extra information design. (See Supplemental Material [56].)

It should be noted that although the “conventional design” used in the *Ctrl* versions closely resembles common visual representations used in textbooks, this does not mean that the animation merely consists of displaying a series of static figures. In fact, most of the animated effects such as zoom in or zoom out and appearing or disappearing of objects are identical across all three versions. The audio scripts of the three versions were exactly the same except that the *Exp* version and the *CInf* version contain a couple more sentences to introduce the equipotential surface. Animations were created with Flash. Synchronization between the audio narration and the visual animation are kept exactly the same in all three versions. For example, if

FIG. 10. Illustration of the two problems used in the multimedia tutorial. (a) Find the potential $V(0)$ at the center of a neutral conducting shell. The zero potential reference point is set at infinity. (b) A positive charge Q is enclosed in a conducting shell. What will happen to the potential difference between points A and B if the shell is enlarged?

in the *Ctrl* version a set of dots and an arrow appear on the screen while the audio narration says the word “integral” in one sentence, then a set of colored rings and the same arrow will appear at exactly the same point in the *Exp* version.

One detail worth mentioning is that in the *Exp* version the audio script never mentions the correspondence between color and thickness of the rings to the electric potential, since we do not want students to read off information from the visual representation. In other words, the meaning of the added perceptual features stayed implicit throughout the tutorial.

The tutorials consist of two parts, each part presenting a detailed solution to an example electric potential problem. The first problem [Fig. 10(a)] asks for the electric potential at the center of a conducting sphere with radius r carrying charge Q , setting the zero potential reference point at infinity. The explanation of this problem also serves as a review of all the basic concepts such as the definition of integral as the sum of infinitely small increments, the integral expression of electric potential difference, and the definition of electric potential with respect to a reference points.

In the second problem [Fig. 10(b)], a point charge with charge Q is enclosed at the center of a thick neutral conducting shell. Two points, A and B , are separated by the shell. The problem first asks for the potential difference between points A and B , then asks what would happen to V_{AB} if the radius of the shell is increased while keeping the thickness unchanged (assuming that A and B are still separated by the shell).

The duration of part I of the multimedia tutorial is 225 s for the *Ctrl* version, and 240 s for both the *Exp* and *CInf* versions. The increase in time is due to the introduction of equipotential surfaces used in the tutorial. The duration of part II of the tutorial is 159 s for all three versions.

E. Participants

Participants in this study were invited from undergraduate students enrolled in a calculus-based introductory E&M course in a large midwestern university.

A total of 74 students responded to the invitation. All of them were asked to complete an online pretest before coming to the experiment, which consisted of three basic questions on electric potential. The pretest was then graded before the experiment, with each problem worth 1 point. Potential participants were then divided into three groups, each experiencing one of the three different versions of the multimedia tutorial on the day of the experiment. Those three groups have identical average scores on the pretest. For participants who scored ≥ 2 out of 3 points in the pretest, their answers to the post-test problems were excluded from the final analysis, as they likely have a decent understanding of the subject.

A total of 61 students participated in the experiment, with 22 in the *Ctrl* group, 21 in the *Exp* group, and 18 in *CInf* group. After excluding the answers from those with high pretest scores, the remaining populations in the three groups are 18, 20, and 16 for groups *Ctrl*, *Exp*, and *CInf*, respectively. The average score on hour exam 1 (HEX1) is not significantly different among the three groups, with the *Ctrl* group slightly lower and *CInf* group slightly higher ($p = 0.2$), as shown in Table I. The HEX 1 score is chosen as an indicator for students’ background content knowledge, since all of the knowledge relevant to the problems in the post-test were learned before HEX1 and covered in HEX1.

F. Assessments

The pretest consists of three relatively easy electric potential problems, each containing two questions. (See Supplemental Material [56].)

The post-test used in this experiment consists of two parts (see Sec. V C). Participants were asked to complete

TABLE I. Average scores on Hex 1 for each group.

	<i>Ctrl</i>	<i>Exp</i>	<i>CInf</i>
HEX 1	65.77 ± 1.95	68.15 ± 1.88	71.19 ± 1.92

each part of the post-test after viewing the corresponding part of the tutorial (see Fig. 9). Post-test part I contains 8 problems, and part II contains 6 problems. The form of the problems consists of both open-ended questions and multiple-choice questions. For most of the multiple-choice questions, students are asked to also write down the reason for making the particular choice, except for PI-4 and PII-1 where students choose between different graphs and the reason for the choice is obvious.

Most of the problems ask for the electric potential in space, with the exception of PI-2, PI-5, PII-3, and PII-5, which ask about the electric field distribution in space. The main purpose of these problems is to serve as a reminder to participants for using the electric field in their solution to the potential problems. Therefore, the electric field problems contain a lot of hints, and should be very easy for our participants.

For electric potential problems, the ones at the beginning of the each part of the post-test are, in general, more straightforward and share more surface feature similarity with the example used in the tutorial (i.e., near transfer problems), such as the type of question asked (straightforward calculation of potential) or the type of electrical medium involved (conductor). The problems towards the end are more challenging, as they share less surface feature similarity with the examples (i.e., far transfer problems; there is more discussion on near and far transfer problems in Sec. V H).

It should be pointed out that the knowledge of equipotentials is not necessary for solving any of the electric potential problems, nor can we think of any way that the solver can benefit from that knowledge in solving any of the problems. In other words, students who received an “equipotential ring” representation do not have any obvious edge over students who received a “dots and lines” representation.

G. Procedural details

The experiment was carried out a couple of days before the final exam of the E&M course that the participants were taking. Participants were given one and a half hours to finish both the tutorial and the assessment. All participants finished the experiment within the given time, with most participants finishing within an hour, with most of the time spent on solving post-test problems (see Sec. V I 4 for more on time usage).

The multimedia tutorials were distributed to participants via a custom version of the SMARTPHYSICS multimedia player⁴. The same player is used weekly to distribute course materials to the participants, and therefore participants are very familiar with the operation of the player.

⁴To view the animated tutorials with the multimedia player used in the experiment, go to the Web address in Ref. [57] and use “test” as NetId to log into the player.

When viewing the tutorials, participants are allowed to freely pause and rewind at any time. When working on part one of the assessment, participants can go back and review part I of the multimedia tutorial, but cannot proceed to watch part II of the tutorial. Once they finished assessment part I and proceeded to viewing tutorial part II, they were explicitly instructed not to go back and change their answers to assessment part I. When working on assessment part II, participants were free to view both parts of the multimedia tutorials. The player sends the time stamp of each action (play, pause, rewind, stop) taken by each individual participant to a database on a remote server.

Participants received \$10 for participating in the experiment.

H. Predictions

If our interpretation of grounded cognition principles is correct, then participants in the *Exp* group who receive the perceptually enhanced design of visual representation should be more likely to construct a better simulation of electric potential. However, we were unsure as to how this superiority in conceptual understanding can affect participants’ problem solving abilities. One could argue that since the tutorials are fairly short and participants in the *Exp* group were exposed to a new representation for the first time, they might only achieve a slightly better understanding, and would only be able to outperform the control groups on easier, more straightforward problems. On the other hand, it is equally likely that participants in both control groups could obtain the correct answer for the easier questions based on their shallow understanding or utilizing equation hunting strategy. In that case, the difference in understanding can be best reflected in the performance on harder problems.

Therefore, to increase our chance of observing a difference, we put in the post-test a number of different problems with a variety of difficulty level (as perceived by physics experts), ranging from simple near transfer problems to more challenging far transfer problems. We hope that participants in the *Exp* group will outperform participants in both *Ctrl* group and *CInf* group on at least some of the problems with the right level of difficulty.

I. Results

As mentioned in Sec. V F, the assessment problems consist of electric-field problems and electric-potential problems. All of the subjects answered the electric field problems correctly. Since these problems serve as hints to students and are irrelevant to the purpose of the experiment, they are excluded from our analysis. Therefore, in the remainder of the paper, we will use “problems” to refer to only the electric potential problems.

1. Grading scheme

Students’ answers for each question on both parts of the post-test are first graded on a scale of 0 to 3. The grading

TABLE II. Rubrics for experiment 1 post-test problems.

Score	Criteria
0	Completely wrong or no answer at all.
1	Showed some signs of understanding but made major mistakes.
2	Incomplete reasoning or minor mistakes.
3	Completely correct.

scheme for each problem is based on students' calculation and written explanation of their reasoning (if applicable), according to the rubric listed in Table II.

Two individual graders, a physics faculty and an experienced physics TA, both members of the PER group, graded all the students' answers independently at first. They then discussed the grading with each other, and were able to reach a consensus on the scores assigned to each problem to within 1 point. We will report analysis results based on the average score of the two graders. Six questions relevant to electric potential were asked on post-test part I, in which PI-6 and PI-7 were designed to test different aspects of the same concept, and are graded as one problem. Therefore, the total points available for part I of the post-test is 15. Four questions on electric potential are given on post-test part II (12 points in total).

2. Post-test performance

The average post-test score, reported as the percentage of total points earned, is listed in Table III and plotted in Fig. 11.

As can be seen from the data, the average total score for the first part of the post-test slightly favors the *Exp* group, while the average score for the second part of the post-test is significantly different between the three groups, with *Exp* group having a higher average than both control groups. Using simple *t* tests to compare pairs of groups on the post-test part II yields $p = 0.002$ between the *Ctrl* and *Exp* group, and $p = 0.01$ between the *Exp* and *CInf* group. (Scores are reported as the percentage of total possible points for potential problems in each part.)

To further investigate the participants' post-test performance, we look at their performance on each individual problem. However, by doing so we have further decreased the sample size in the data, and thus differences between the two graders have a nontrivial impact on the average score

FIG. 11. Average post-test score for both parts of the post-test. Post-test scores are reported as the percentage of total available points earned.

on some of the problems. Yet it turned out that the majority of the differences in scoring lies between “1”s and “0”s (somewhat wrong answer vs completely wrong answer), which is the least interesting. In order to eliminate this unnecessary ambiguity, we chose to plot the percentage of correct answers (score = 3) in each group for every problem on the post-test, which both graders completely agreed upon, as shown in Fig. 12.

As can be seen from the data, the *Exp* group outperformed both control groups on all problems. Among the problems in part I, the *Exp* group outperformed both control groups on P1-1 ($p = 0.06$). P1-3 showed a similar but not significant trend ($p = 0.25$). On part II of the post-test, the *Exp* group performed significantly better than both control groups on PII-1 and PII-6 ($p < 0.01$). Interestingly, PII-2 and PII-3 have a higher correct rate, yet did not show much difference between the three groups.

3. Participants' written reasoning for PII-6

As can be seen in Fig. 12, two of the post-test problems, PII-1 and PII-6, showed the largest between-group differences. Unlike PII-1, which is a straightforward multiple choice question testing a piece of factual knowledge, PII-6 is a more sophisticated problem that requires a substantial amount of logical reasoning to solve. Therefore, it is interesting to explore the types of reasoning used by the participants in solving this problem, since this might shed light on the cause of the big difference in performance between the groups.

For PII-6, participants were asked to provide a written explanation to their answer, and we report here some of the

TABLE III. Students' total score on both parts of the post-test. Scores are reported as a percentage of points available in each part. The θ column reports the effect size between two groups; σ stands for the pooled standard deviation between the two groups.

	<i>Ctrl</i>	<i>Exp</i>	<i>CInf</i>	p (ANOVA)	$\theta_{Ctrl,Exp}$	$\theta_{Exp,CInf}$
Part I	(30.37 ± 5.85)%	(46.50 ± 5.95)%	(37.71 ± 6.11)%	0.16	0.63 σ	0.34 σ
Part II	(41.44 ± 3.43)%	(66.04 ± 5.32)%	(48.96 ± 6.24)%	0.003	1.23 σ	0.70 σ

FIG. 12. Percentage of correct answer (score = 3) for each electric potential problem in the post-test. (a) Percentage of correct answer for problems in part I of the post-test. (b) Percentage of correct answer for problems in part II of the post-test.

most popular types of reasoning identified by looking at their written explanations. It must be noted that our purpose here is to provide the reader with some qualitative sense of different types of reasoning provided by different groups, rather than presenting a quantitative statistical analysis, since participants' written explanations can sometimes be incomplete or ambiguous, and the sample size is very small. Therefore, we will only give rough estimates of numbers for each type of reasoning.

In PII-6, a charged particle is surrounded by a charged thin insulating shell with negligible thickness (Fig. 13). The problem asks what would happen to the potential difference across the shell if the radius of the insulating shell were decreased.

One of the correct strategies for solving PII-6 is to argue that by changing the radius of the shell, the potential integral covers more distance where there is a stronger electric field, thus increasing the total potential difference. The majority of participants who correctly answered PII-6 provided such a qualitative argument, among which most were in the *Exp* group. On the contrary, participants in both control groups often (wrongly) focused on the insulating shell itself, providing arguments such as “the shell is infinitely thin so there’s no area in which the E field is zero, so (potential difference) will not change.” This type of reasoning comprised roughly a third of all the wrong

explanations identified. Remarkably, only one participant in the *Exp* group gave this type of reasoning.

Alternatively, PII-6 can also be solved by evaluating the mathematical expression of the potential integral as a function of the radius r , which was already obtained in a previous problem (PII-4). While this strategy is feasible, it is difficult for most of the participants to carry out due to its mathematical complexity, resulting in very high error rates. The integral expression contains two terms representing the potential difference in and out of the shell, respectively. The magnitudes of the two terms change in an opposite direction as a function of r , with the outside-of-the-shell term being more dominant due to a larger electric field. However, most participants performing a mathematical evaluation only evaluated the change in one of the two terms in the integral expression, while neglecting the other. This accounts for roughly another one-third of the incorrect answers for PII-6. None of those participants showed any signs of trying to tie their mathematical expression back to the physics situation.

Remarkably, participants in the *Exp* group showed an overwhelming preference for qualitative argument over mathematical manipulation. Only one participant in this group was observed to perform such a mathematical manipulation, who correctly evaluated both terms.

4. Student rating of tutorial and viewing data

Given the observed performance differences between the different groups, an important question to ask is whether the tutorials given to the two control groups are perceived to be of relatively similar quality as the *Exp* tutorial as judged by the participants. Recall that participants were drawn from a population who received similar weekly instructional materials, and so they are in a position to render an opinion on the quality of the tutorials based on their previous experiences. If students in the control groups were confused by the tutorials, then we would be guilty of creating a “straw man” to strike down with our *Exp* tutorial.

To answer that question, we will examine two pieces of information. First, we look at students' ratings on the quality of the multimedia tutorial (Table . As previously

FIG. 13. Figure used for post-test problem PII-6. An insulating shell with positive charge $+q$ surrounds a positive point charge Q . The problem asks what will happen to the potential difference between A and B, V_{AB} , if the radius of the shell becomes smaller.

TABLE IV. Average students' rating of both parts of the tutorial.

	Part I rating	Part II rating
<i>Ctrl</i>	7.73 ± 1.75	8.59 ± 1.18
<i>Exp</i>	8.10 ± 1.48	8.52 ± 1.66
<i>CInf</i>	8.11 ± 1.75	7.83 ± 1.20

TABLE V. Average total time spent on each part of the multimedia tutorial.

Time (s)	Part I	Part II
<i>Ctrl</i>	337.17 ± 30.83	210.35 ± 13.32
<i>Exp</i>	316.71 ± 22.48	194.43 ± 11.38
<i>CInf</i>	290.00 ± 21.44	217.39 ± 20.77
<i>p</i> (ANOVA)	0.45	0.55

mentioned, students were asked to rate the “helpfulness” of each part of the tutorial on a scale of 1 to 10 (see Supplemental Material [56]).

All three groups gave almost identical average ratings (~ 8 out of 10) to the first part of the tutorial. The *Ctrl* group and *Exp* group also gave identical ratings to the second part of the tutorial, while the *CInf* group gave a slightly lower (but statistically insignificant, $p = 0.23$) rating. One student in *CInf* group explicitly commented that the equipotential rings were distracting in this part.

Second, we look at the time students spent watching each part of the tutorial. From previous experiences analyzing students' interaction with multimedia, we know that if a student is confused by certain parts of a multimedia tutorial, they will often tend to rewind through that part and watch it again. Therefore, we could estimate the quality of the tutorial simply by looking at the total time they spent on watching it, with more time on a segment (more rewinds) implying less clarity. Furthermore, rewinds caused by confusion tend to be centered on the part of the tutorial that is difficult to understand, whereas rewinds caused by other reasons (such as hardware or software error) generally do not overlap with each other. By simply counting the total

number of times each segment of the animation was being played, we are able to roughly estimate the relative difficulty of each segment of the tutorial.

The total time spent by an average participant watching each part of the tutorial is listed in Table V. As previously mentioned in Sec. V D, the duration of part I of the tutorial is ~ 230 s (slightly different between different versions), and the duration of part II of the tutorial is 159 s. Therefore, on average, participants spent about an additional 1 to 2 minutes replaying certain segments of each part of the tutorial.

No significant difference is observed among the three groups on time spent watching the tutorials. Note that the “watching time” reported here represents the time spent when the animation is playing. Participants can also pause the animation, and look at a particular frame on the screen. However, there is no way of knowing how much time was spent on watching static frames.

In Fig. 14, we plot the average number of times each 10 s-long segment of the tutorial is being watched by participants. As can be seen, the viewing pattern is essentially the same for the three groups of students. The beginning and end of part I of the tutorial received more average views than other parts of the tutorial, which is viewed a little over once on average.

VI. DISCUSSION

The results of the experiment showed a clear difference in the post-test performance between the experimental group and the two control groups. The observation of a significant difference in post-test performance is quite remarkable given that the tutorials were merely ~ 5 minutes long, and are almost identical in every other aspect except for the visual details in the representation. In fact, the differences between the treatments in this experiment are much smaller than in any previous study on multimedia learning, such as the ones done by Mayer [30,58] and Schnotz [14,16,59].

Mayer's central argument is that multimedia design should not incur unnecessary cognitive load on the learner. In the current experiment setting, all three versions of animation are derived from the same template that strictly follows Mayer's principle, so they should in theory require very similar amounts of cognitive load to process. One

FIG. 14. Average number of times students watched every 10 s segment of the animated tutorial.

could even argue that the *Exp* version of animation puts a heavier extraneous load on the visual channel as it displays many more visual elements that are, from a physics point of view, completely unnecessary for conveying the correct knowledge. Therefore, using only Mayer's model, it is impossible to give a convincing argument as to why these additional visual elements could facilitate learning.

One might argue that the colored rings may have facilitated problem solving by enabling the subjects in the *Exp* group to read off the relative magnitude of the potential distribution in space, as suggested by the Schnotz model. Even though we never explicitly mention the correspondence between visual elements (color and thickness) and physics quantities (potential gradient) in the tutorial, subjects could still have generated this information themselves through visual inspection and self-explanation [60]. The critical flaw to this argument, however, is that for almost all of our assessment problems, regenerating the same (or even a similar) mental image and read-off information from that image would be "suicidal" for solving the problem, as the physical situation has completely changed. To generate the correct mental image for the problem situation, one needs to understand how the change in field distribution leads to the change in potential distribution. Yet, that abstract understanding itself is sufficient for solving the problems, leaving the generation of mental image completely unnecessary. In other words, Schnotz's model is inapplicable to learning of abstract knowledge such as electric potential.

Further, in contrast to the common wisdom that animations tend to make instructions more "interesting" [58], thereby enticing students to spend more time, we observed no difference in time spent on watching the tutorials. It is also unlikely that participants from the two control groups underperformed because they didn't like their version of animation, or because they noticed that they are being given a tutorial of lower quality, as they rated their tutorials as highly as participants in the *Exp* group. Furthermore, saying that the added visual elements simply helped students memorize the material by providing more visual cues is an oversimplification, as participants are free to review the tutorial at any time they want during the experiment, and that students in all three groups spent about the same amount of time viewing the tutorials.

Therefore, the most plausible explanation for the observed differences in post-test performance is that the *Exp* version of the tutorial, with higher *p* quality, facilitated the participants' ability to construct a proper simulation representing the integral expression of electric potential by activating some of the essential perceptual symbols through the perceptual method.

As to the exact mechanism by which a simulation facilitates the learner in problem solving, we do not have a conclusive answer. However, we are able to gain some insight into this question by looking at the different reasoning provided by participants to PII-6.

What makes PII-6 particularly interesting is that it shares a lot of surface features with the example used in the second part of the tutorial, yet the underlying physics contains a subtle difference. Both problems involve a point charge surrounded by a spherical shell, and both ask about the change in potential difference across the shell when the radius of the shell changes (Fig. 15). Yet, the key to solving the example problem, which is to notice that the electric field (and therefore the potential difference) is zero inside the conducting shell, is completely irrelevant to solving PII-6. On the other hand, solving PII-6 requires participants to focus on the area inside and outside of the shell, which is only indirectly related to the solutions given in the tutorial. In short, compared to the example in the second part of the tutorial, PII-6 is a "far" transfer problem that requires the solver to focus on deep structure instead of surface features.

Not surprisingly, participants in the control groups seemed more likely to cue on the surface similarity between PII-6 and the tutorial example, wrongly focusing on the zero thickness of the insulating shell. Notice that the tutorial presented to the *Exp* group contains a lot more visual elements than the other two, and yet participants in this group are much less likely to be distracted by surface similarity. Instead, they are more likely to focus on the underlying relation between electric field and electric potential.

This seemingly counterintuitive observation is hard to explain if one assumes that abstract relations are represented by amodal symbols that are conveyed predominantly through language instead of perception (e.g., the Schnotz model). In that case, a change in the perceptual features should have no major impact on their ability to process the abstract aspect of the problem. On the other hand, from a grounded cognition point of view, abstract relations are also grounded in perceptual simulations, and perceiving the added visual elements in the *Exp* version of the animation would facilitate the construction of such a simulation. When the brain's attention is focused mostly on the internal simulation, it naturally neglects those peripheral visual features that are nonessential to the simulation, such as the

FIG. 15. Comparison between post-test problem PII-6 and example problem used in tutorial part 2. (a) Problem PII-6 involves an infinitely thin conducting shell charged with charge $+q$, shown as a black circle, surrounding a point charge (black dot in the middle). (b) Example problem in part 2 of the tutorial involves a thick conducting shell (gray circular ring) surrounding the same charged particle Q .

thickness of the conducting shell. In contrast, the low p quality animation of the other two tutorials interferes with the construction of a proper simulation, and in the absence of a simulation the brain is unable to determine the importance of perceptual features. In that case, it tends to focus on (and memorize) the most salient visual feature perceived, which in this case is likely to be the thick conducting shell.

Another interesting observation is that participants in the *Exp* group overwhelmingly prefer qualitative reasoning over mathematical manipulation when solving for PII-6. There is no reason to believe that participants in the *Exp* group possess fewer mathematical skills, as they are equally capable of setting up and evaluating the integrals in other problems. A more plausible explanation is that participants in the *Exp* group view qualitative reasoning as a much easier and more obvious method to approach this problem, whereas participants in both control groups do not—in fact, they showed proclivity toward mathematical evaluations. This observation is also consistent with previous research on students' self-explanation [60].

Physics educators notice on a daily basis that introductory level students frequently choose equation hunting or number crunching over qualitative physics reasoning as their preferred approach for problem solving. Our observation suggests that at least in some cases, it might not be that students are unwilling to engage in qualitative

reasoning, but that they are in fact unable to do so because of inadequate understanding. We see that in this case when a proper simulation is constructed with the help of high p -quality visual representations, participants feel comfortable abandoning math manipulation methods and adopting qualitative reasoning using language based arguments.

Aside from PII-6, the other problem with significant performance difference is PII-1, in which participants are asked to choose the correct graph that represents the distribution of potential difference along the radial direction of a system consisting of a charge and a thick conducting shell (Fig. 16). A closer look at the problem provides some insight into how a better simulation can facilitate students in problem solving.

The major difference between the correct choice, a, and the other distractors, lies in that all the distractors depict the value of electric potential in the region inside the conductor as zero instead of a constant.

The large between-group difference in the answer to this problem demonstrates how low p -quality representations can interfere with the proper functioning of the s method. Namely, even though we consistently and repeatedly explained and emphasized throughout the tutorial that the “potential difference,” rather than the value of “potential,” is zero inside a conductor, the majority of participants in both control groups still end up thinking that the value of potential is zero inside the conducting shell. In contrast,

FIG. 16. Post-test problem PII-1.

when the visual representation is designed to represent the idea that “the change in potential is zero,” most *Exp* group participants were able to correctly interpret the verbal explanation. Remarkably, not one student in either of the control groups complained that the audio narration is incongruent with the visual representation, nor were there any complaints from the *Exp* group on the audio narration being “ambiguous” or “misleading.”

This result is easy to understand within a grounded cognition framework, in which the meaning of language is being represented by perceptual simulations. The creation of simulations by the brain is easily influenced by perception, as simulation and perception share the same perceptual domains. For introductory level physics students, their brain has not yet learned to consistently generate different simulations in response to the words potential and potential difference. Therefore, the perception of a vacant area in the conductor could easily lead the brain into generating a simulation of “zero potential” instead of “zero potential difference.” On the other hand, when the same language is being interpreted under the background simulation of a general “accumulation process,” it becomes much easier to generate a new simulation in which the accumulation procedure stops over a certain region (thus remaining constant), whereas a simulation of the accumulated value suddenly dropping to zero goes against all our existing understanding of an accumulation process. Since participants in the *Exp* group have a much better chance of generating a simulation of accumulation by being exposed to a high p quality representation, they also have a better chance of correctly interpreting the meaning of zero potential difference.

In contrast, if one assumes that language is processed independently from visual perceptions, as is the case with both Mayer’s and Schnotz’s model, then the interference between verbal and pictorial representations is merely a second order effect. In that case, the degree of interference observed for this problem becomes quite difficult to explain.

Admittedly, due to the small sample size of this experiment, these observations on the detailed difference in problem solving behavior are more suggestive than conclusive. Nonetheless, when viewed as a whole, all of the experimental results suggest that relatively small changes in the design of visual representations can indeed have a substantial impact on students’ conceptual understanding of abstract physics concepts, leading to a detectable difference in problem solving performance. This type of impact is beyond the scope of existing multimedia learning models, and can be best explained by adopting a grounded cognition framework. In grounded cognition terms, the quality of a given representation consists of two largely independent components, the p quality and the s quality, and the p quality of a representation can have a direct impact on the brain’s ability to construct a proper simulation representing the abstract physics knowledge.

We think that the concept of p quality serves as a useful guideline for instructors seeking to create new visual

representations for teaching physics, as it reminds us of the importance of perceptual details that are often easily overlooked. Especially for those who seek to utilize the visual power of computer animation, our results serve as a useful criterion for distinguishing between beneficial visual effects and unnecessary visual distractors. Namely, visual perceptions that are aligned with the meaning of the content, such as using thicker line and darker color to tacitly represent “stronger,” can facilitate conceptual understanding, whereas those that are irrelevant to the content, such as drawing a shiny realistic battery instead of a battery symbol, while being aesthetically pretty, make little difference in learning outcomes and could become a source of distraction [61]. Although it is impossible to precisely determine what simulation any given student would create upon perceiving a certain visual representation, a representation with higher p quality will, according to PSS theory, make it more likely for the student to generate a simulation that is similar to that of the instructor’s, which leads to better conceptual understanding and improved problem solving performance.

As a final remark, aside from providing new insights on effective design of visual representations, this paper also serves as one of the first attempts to introduce some of the ideas in the burgeoning new field of grounded cognition into PER. Being a fundamental cognitive theory that bridges the gap between external perception and internal cognition, grounded cognition can potentially have a much more profound impact on our understanding of learning.

In addition, as a young and developing field, grounded cognition also faces certain challenges and raises many interesting new questions, which present great opportunities for future research.

For example, this lack of a well-developed theory poses a number of important questions, such as the following: How much experience is required to generate a proper simulator, and how do we most effectively provide those experiences to students? What are the factors that determine whether a particular perceptual symbol will be incorporated into a simulation? How exactly do visual representations influence the semantic or linguistic understanding of mathematical equations?

Another challenging yet important task is to determine if a grounded view of learning can be reconciled with existing cognitive learning theories. For example, we have noticed that the PSS framework, consisting of perceptual symbols and simulations, bears much similarity with the “knowledge in pieces” view, which also suggests that knowledge is built from more fundamental pieces originating from perceptual experience with the world. It would be interesting to see to what extent the two theories are in fact congruent with each other, and what kind of new insights we might be able to obtain by combining the two theories.

We hope that this work can serve as a primitive attempt, and can inspire more research along this direction.

ACKNOWLEDGMENTS

We would like to give special thanks to Dr. Jose Mestre for his invaluable advice on preparing the manuscript. We would also like to thank Dr. Brian Ross and Dr. Timothy Stelzer for insightful suggestions on experimental design, as well as Dr. Michael Scott for providing essential technical support for the experiment. This research is supported by NSF Grant No. DUE 08-17185.

-
- [1] A. Van Heuvelen and X. Zou, Multiple representations of work-energy processes, *Am. J. Phys.* **69**, 184 (2001).
  - [2] Y. J. Dori and J. Belcher, Learning electromagnetism with visualizations and active learning, in *Visualization in Science Education*, edited by J. K. Gilbert (Springer, New York, 2005), pp. 187–216.
  - [3] J. M. Casperson and M. C. Linn, Using visualizations to teach electrostatics, *Am. J. Phys.* **74**, 316 (2006).
  - [4] Z. Chen, T. Stelzer, and G. Gladding, Using multimedia modules to better prepare students for introductory physics lecture, *Phys. Rev. ST Phys. Educ. Res.* **6**, 010108 (2010).
  - [5] M. Dancy and R. Beichner, Impact of animation on assessment of conceptual understanding in physics, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010104 (2006).
  - [6] B. Ambrose, P. Heron, S. Vokos, and L. C. McDermott, Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena, *Am. J. Phys.* **67**, 891 (1999).
  - [7] N. Podolefsky and N. Finkelstein, Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020104 (2007).
  - [8] N. S. Podolefsky and N. D. Finkelstein, Analogical scaffolding and the learning of abstract ideas in physics: An example from electromagnetic waves, *Phys. Rev. ST Phys. Educ. Res.* **3**, 010109 (2007).
  - [9] M. Kozhevnikov, M. A. Motes, and M. Hegarty, Spatial visualization in physics problem solving, *Cogn. Sci.* **31**, 549 (2007).
  - [10] C. Wieman, The “curse of knowledge” or why intuition about teaching often fails, *APS News*, **16**, The Back Page (2007) [<http://www.aps.org/publications/apsnews/200711/backpage.cfm>].
  - [11] C. E. Wieman and K. K. Perkins, A powerful tool for teaching science, *Nat. Phys.* **2**, 290 (2006).
  - [12] R. E. Mayer, *Multimedia Learning* (Cambridge University Press, New York, 2001).
  - [13] W. Schnotz, Construction and interference in learning from multiple representation, *Learn. Instr.* **13**, 141 (2003).
  - [14] W. Schnotz, Towards an integrated view of learning from text and visual displays, *Educ. Psychol.* **14**, 101 (2002).
  - [15] S. Reed, Cognitive architectures for multimedia learning, *Educ. Psychol.* **41**, 87 (2006).
  - [16] W. Schnotz and C. Kürschner, External and internal representations in the acquisition and use of knowledge: Visualization effects on mental model construction, *Instr. Sci.* **36**, 175 (2008).
  - [17] L. W. Barsalou, Perceptual symbol systems, *Behav. Brain Sci.* **22**, 577 (1999).
  - [18] L. W. Barsalou, Grounded cognition: Past, present, and future, *Top. Cognit. Sci.* **2**, 716 (2010).
  - [19] M. L. Anderson, The massive redeployment hypothesis and the functional topography of the brain, *Philos. Psychol.* **20**, 143 (2007).
  - [20] L. W. Barsalou, Grounded cognition, *Annu. Rev. Psychol.* **59**, 617 (2008).
  - [21] T. Hansen, M. Olkkonen, S. Walter, and K. K. R. Gegenfurtner, Memory modulates color appearance, *Nat. Neurosci.* **9**, 1367 (2006).
  - [22] M. Tucker and R. Ellis, On the relations between seen objects and components of potential actions, *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 830 (1998).
  - [23] L. Chao and A. Martin, Representation of manipulable man-made objects in the dorsal stream, *NeuroImage* **12**, 478 (2000).
  - [24] R. A. Zwaan and C. J. Madden, Embodied sentence comprehension, in *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, edited by D. Pecher and R. A. Zwaan (Cambridge University Press, New York, 2005), pp. 224–245.
  - [25] L. W. Barsalou and K. Wiemer-Hastings, Situating abstract concepts, in *Grounding Cognition* (Ref. [24]), pp. 129–163.
  - [26] M. L. Anderson, Embodied cognition: A field guide, *Artif. Intell.* **149**, 91 (2003).
  - [27] L. W. Barsalou, Simulation, situated conceptualization, and prediction, *Phil. Trans. R. Soc. B* **364**, 1281 (2009).
  - [28] D. L. Schwartz and J. B. Black, Shuttling between depictive models and abstract rules: Induction and fallback, *Cogn. Sci.* **20**, 457 (1996).
  - [29] M. Hegarty, Constructing mental models of machines from text and diagrams, *J. Mem. Lang.* **32**, 717 (1993).
  - [30] R. E. Mayer and R. B. Anderson, Animations need narrations?: An experimental test of a dual-coding hypothesis, *J. Educ. Psychol.* **83**, 484 (1991).
  - [31] C. Rebetez, M. Betrancort, M. Sangin, and P. Dillenbourg, Learning from animation enabled by collaboration, *Instr. Sci.* **38**, 471 (2010).
  - [32] L. Boroditsky and M. Ramscar, The roles of body and mind in abstract thought, *Psychol. Sci.* **13**, 185 (2002).
  - [33] L. W. Barsalou, Abstraction as dynamic interpretation in perceptual symbol systems, in *Building Object Categories*, edited by L. Gershkoff-Stowe and D. Rakison, Carnegie Symposium Series (Erlbaum, Mahwah, NJ, 2005), pp. 389–431.
  - [34] C. M. MacLeod, Half a century of research on the Stroop effect: An integrative review, *Psychol. Bull.* **109**, 163 (1991).

- [35] R. J. Stroop, Studies of interference in serial verbal reactions, *J. Exp. Psychol.* **18**, 643 (1935).
- [36] A. Henik and J. Tzelgov, Is three greater than five: The relation between physical and semantic size in comparison tasks, *Mem. Cogn.* **10**, 389 (1982).
- [37] P. Pinel, M. Piazza, D. Le Bihan, and S. Dehaene, Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments, *Neuron* **41**, 983 (2004).
- [38] R. Cohen Kadosh and V. Walsh, Numerical representation in the parietal lobes: Abstract or not abstract?, *Behav. Brain Sci.* **32**, 313 (2009).
- [39] D. Landy and R. L. Goldstone, Formal notations are diagrams: Evidence from a production task, *Mem. Cogn.* **35**, 2033 (2007).
- [40] D. H. Landy, M. N. Jones, and R. L. Goldstone, How the appearance of an operator affects its formal precedence, *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (Cognitive Science Society, Austin, TX, 2008), pp. 2109–2114.
- [41] R. L. Goldstone, D. H. Landy, and J. Y. Son, The education of perception, *Top. Cognit. Sci.* **2**, 265 (2010).
- [42] D. M. Belenky and T. J. Nokes, Examining the role of manipulatives and metacognition on engagement, learning, and transfer, *J. Probl. Solving* **2**, 102 (2009).
- [43] D. W. Braithwaite and R. L. Goldstone, Integrating formal and grounded representations in combinatorics learning, *J. Educ. Psychol.* **105**, 666 (2013).
- [44] R. L. Goldstone and Y. Sakamoto, The transfer of abstract principles governing complex adaptive systems, *Cogn. Psychol.* **46**, 414 (2003).
- [45] V. M. Sloutsky, J. A. Kaminski A. F. Heckler, The advantage of simple symbols for learning and transfer, *Psychon. Bull. Rev.* **12**, 508 (2005).
- [46] J. Y. Son and R. L. Goldstone, Fostering general transfer with specific simulations, *Pragmat. Cognit.* **17**, 1 (2009).
- [47] D. H. Uttal, K. V. Scudder, and J. S. DeLoache, Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics, *J. Appl. Dev. Psychol.* **18**, 37 (1997).
- [48] J. A. Winawer, N. Witthoft, L. Wu, M. Frank, A. Wade, and L. Boroditsky, Effects of language on color discriminability, *J. Vis.* **3**, 711 (2003).
- [49] J. A. Winawer, N. Witthoft, L. Wu, M. Frank, A. Wade, and L. Boroditsky, Effects of language on color discriminability, *J. Vis.* **3**, 711 (2003).
- [50] W. Chase and H. Simon, Perception in chess, *Cogn. Psychol.* **4**, 55 (1973).
- [51] A. M. Lesgold, H. Rubison, P. Feltovich, R. Glaser, D. Klopfer, and Y. Wang, Expertise in a complex skill: Diagnosing x-ray pictures, in *The Nature of Expertise*, edited by M. T. H. Chi, G. Robert, and M. Farr (Erlbaum, Hillsdale, NJ, 1988), pp. 311–242.
- [52] M. T. H. Chi, P. J. Feltovich, and R. Glaser, Categorization and representation of physics problems by experts and novices, *Cogn. Sci.* **5**, 121 (1981).
- [53] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics* (John Wiley & Sons, Inc., New York, 2000), 6th ed.
- [54] R. Wolfson and J. Pasachoff, *Physics: With Modern Physics for Scientists and Engineers* (Addison-Wesley, Reading, MA, 1998), 3rd ed.
- [55] P. Tipler and G. Mosca, *Physics for Scientists and Engineers* (W. H. Freeman and Company, New York, 2000), 5th ed.
- [56] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevSTPER.10.010111> for pre- and post-tests as well as flash animation used in the experiment.
- [57] The smartPHYSICS multimedia player used in the experiment can be found at <http://research.physics.illinois.edu/PER/Chen/index.html>.
- [58] Mireille Betrancourt, The animation and interactivity principles in multimedia learning, in *The Cambridge Handbook of Multimedia Learning*, edited by R. E. Mayer (Cambridge University Press, New York, 2005), pp. 287–296.
- [59] W. Schnotz and T. Rasch, Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning, *Educ. Technol. Res. Dev.* **53**, 47 (2005).
- [60] M. T. H. Chi, Self-explaining expository texts: The dual processes of generating inferences and repairing mental models, in *Advances in Instructional Psychology*, edited by R. Glaser (Lawrence Erlbaum Associates, Mahwah, NJ, 2000), pp. 161–238.
- [61] V. M. Sloutsky, J. A. Kaminski, and A. F. Heckler, The advantage of simple symbols for learning and transfer, *Psychon. Bull. Rev.* **12**, 508 (2005).