Educational Psychologist

ISSN: 0046-1520 (Print) 1532-6985 (Online) Journal homepage: http://www.tandfonline.com/loi/hedp20

The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes

Michelene T. H. Chi & Ruth Wylie

To cite this article: Michelene T. H. Chi & Ruth Wylie (2014) The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes, Educational Psychologist, 49:4, 219-243, DOI: 10.1080/00461520.2014.965823

To link to this article: https://doi.org/10.1080/00461520.2014.965823

Published online: 28 Oct 2014.

EDUCATIONAL PSYCHOLOGIST, *49*(4), 219–243, 2014

ISSN: 0046-1520 print / 1532-6985 online
DOI: 10.1080/00461520.2014.965823

The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes

Michelene T. H. Chi and Ruth Wylie
Mary Lou Fulton Teachers College
Arizona State University

This article describes the ICAP framework that defines cognitive engagement activities on the basis of students' overt behaviors and proposes that engagement behaviors can be categorized and differentiated into one of four modes: *Interactive, Constructive, Active,* and *Passive*. The ICAP hypothesis predicts that as students become more engaged with the learning materials, from *passive* to *active* to *constructive* to *interactive*, their learning will increase. We suggest possible knowledge-change processes that support the ICAP hypothesis and address the limitations and caveats of the hypothesis. In addition, empirical validation for the hypothesis is provided by examining laboratory and classroom studies that focus on three specific engagement activities: note taking, concept mapping and self-explaining. We also consider how ICAP can be used as a tool for explaining discrepant findings, dictate the proper choice of a control condition, and evaluate students' outputs. Finally, we briefly compare ICAP to existing theories of learning.

Educators have long recognized that, although students can learn from receiving information passively, they fare much better by learning actively. "Active learning" is typically defined by educational researchers as learning that requires students to *engage* cognitively and meaningfully with the materials (Bonwell & Eison, 1991), to get "*involved* with the information presented, really *thinking about* it (analyzing, synthesizing, evaluating) rather than just passively receiving it" (King, 1993, p. 2). Thus, "active" students are considered to be cognitively "engaged" (Corno & Madinach, 1983), and we use those terms synonymously.

Despite the fact that "active learning" is defined as engaging *cognitively*, most of the research on engagement considers it from the motivational perspective (Blumenfeld, Kempler, & Krajcik, 2004; Pintrich & De Groot, 1990; Zimmerman, 1990), the behavioral perspective, or the emotional perspective. When engagement is discussed in motivational terms, it tends to mean the *precursor attitude* or *interest* in getting involved with the learning materials. Behavioral engagement refers generally to the notion of participating and addresses large-grained measures, such as how often students attend class or do homework, whereas emotional engagement encompasses measures of positive and negative reactions to teachers, classmates, academics, and so on. We are not focusing on the precursor stages of engagement embodied within the motivational perspective, or the large-grained behavioral and the emotional perspectives; instead, we are focusing on the amount of cognitive engagement that can be detected by smaller grained behavioral activities *while students learn.* Although there is some research on cognitive engagement, it emphasizes broad notions such as thoughtfulness and willingness to exert the necessary effort to succeed and master complex skills and ideas (Fredricks, Blumenfeld, & Paris, 2004).

In this article, the term "learning activities" is used as a broad term, referring to the large collection of instructional or learning tasks from which teachers or educational designers can choose for students to do (e.g., reading, solving problems, learning to understand charts and diagrams, etc.). We use the term "engagement" or "engagement activities" to refer to *the way* a student engages with the learning materials in the context of an instructional or learning task, reflected in the overt behavior the student exhibits while undertaking an activity, such as *summarizing* at the end of each paragraph, either orally or in written form. We refer to summarizing as an "engagement activity or behavior" that the student voluntarily undertakes while learning in the context of an instructional task. A teacher can, however, design tasks that elicit more or less engagement from students (e.g., embedding a prompt at the end of each paragraph to remind students to summarize).

Correspondence should be addressed to Michelene T. H. Chi, Mary Lou Fulton Teachers College, Arizona State University, P.O. Box 871811, Mail Code 1811, Tempe, AZ 85287-1811. E-mail: Michelene.Chi@asu.edu

In short, although “active learning” is a great idea for overcoming “passive learning,” we have identified three concrete practical challenges that teachers may face when developing lessons that promote “active learning.” First, broad recommendations such as *engage students cognitively, encourage meaningful learning,* and *get students to think about it* do not tell teachers how to create activities that overcome “passive learning.” Second, teachers have few criteria to use in deciding which are the best “active learning” activities to design and implement. Third, there are no guidelines for teachers regarding how to best modify their favorite existing assignments in order to optimize “active learning.” Although many others have also tackled the challenge of creating more “active learning” environments (e.g., Bonwell & Eison, 1991; Bonwell & Sutherland, 1996; Fink, 2013; Meyers & Jones, 1993; Rosenthal, 1995; Rowles, 2013; Rubin & Herbert, 1998) the framework we introduce in this article can remediate these practical challenges by providing specific domain-general guidelines for creating more engaging lessons as well as modifying existing activities to increase engagement, using simple overt measures of behavior to assess the level to which students are engaged. Moreover, the framework is empirically grounded and supported by evidence from a variety of learning activities, domains, and student ages.

This article has three major sections. The first section describes our ICAP framework that defines engagement in terms of overt behaviors that students can undertake and teachers can see. In brief, we propose that there are different modes or categories of “active learning,” corresponding to different overt behaviors that elicit different knowledge-change or learning processes. Moreover, we propose that learning activities and their resulting overt engagement behaviors can be differentiated into one of four modes: *interactive, constructive, active,* or *passive*. Each mode corresponds to a different set of underlying knowledge-change processes, to be elaborated later. Based on the set of knowledge-change processes, each mode predicts a different level of learning such that the *Interactive* mode of engagement achieves the greatest level of learning, greater than the *Constructive* mode, which is greater than the *Active* mode, which in turn is greater than the *Passive* mode (I>C>A>P). Thus, the ICAP hypothesis predicts different levels of learning for different modes of overt behaviors. Higher levels imply learning with deeper understanding. Although a preliminary version of this framework has been presented in the literature (Chi, 2009), this article expands the framework by explicating the underlying assumptions more concretely and fully and by clarifying the knowledge-change processes and the associated changes in knowledge for each category of activities.

In the second major section, the article describes the learning outcomes of each mode of overt activity and provides empirical evidence that supports the predictions made by the ICAP hypothesis. The evidence we provide in this article differs from our prior publication in that Chi (2009) cited evidence primarily from laboratory experiments measuring performances on a variety of tasks including, but not limited to, learning. In this article, we cite studies that focus on three explicit engagement activities (note taking, concept mapping, and self-explaining). In addition, we cite classroom studies in support of the ICAP hypothesis (that I>C>A>P), showing the relevance of the hypothesis to classroom learning.

In the third section, we consider several caveats or factors that may override or violate the ICAP hypothesis. We also consider new applications of ICAP as a tool for explaining discrepant findings, choosing control conditions, coding student responses, and guiding instructional design. We also briefly compare this theoretical framework with other frameworks in the literature.

SECTION 1: A THEORETICAL FRAMEWORK FOR DIFFERENTIATING ENGAGEMENT ACTIVITIES ACCORDING TO THE MODE OF OVERT BEHAVIORS

The ICAP framework, referred to in previous publications as DOLA for *Differentiated Overt Learning Activities* (Chi, 2009; Menekse, Stump, Krause, & Chi, 2013), consists of a taxonomy that differentiates four modes or categories of engagement, based on the overt behaviors displayed or undertaken by students. Each mode of engagement corresponds to several different types of behaviors and to differentiable knowledge-change processes. Based on the differential knowledge-change processes and the resulting changes in knowledge they produce, the ICAP framework generates the hypothesis that predicts different levels of learning outcomes (see Figure 1). In this section, we describe the four modes, the assumptions of the ICAP framework, the knowledge-change processes and the expected changes in knowledge for each mode, the cognitive outcomes, the ICAP hypothesis, and its predicted learning outcomes.

A Taxonomy of Four Modes of Overt Engagement Behaviors

Students’ engagement with learning materials can be operationalized by the overt (or lack of overt) behaviors they undertake while learning. Although far from perfect, overt behaviors are a good proxy to reflect different modes of engagement that teachers can use to ascertain whether a student is in fact engaged in a specific mode for a given activity. Specifically, students’ overt behavioral manifestations can be characterized and differentiated into four behavioral modes: *passive*, *active*, *constructive*, and *interactive.* By our definition, then, the mode at which students engage can be seen either during instruction or while students are

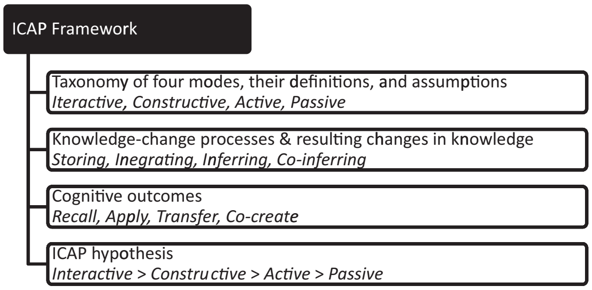

FIGURE 1 The ICAP framework consists of a taxonomy of four modes of activities, their definitions and assumptions; hypothesized knowledge-change processes and the resulting changes in knowledge; projected cognitive outcomes; and the ICAP hypothesis of learning outcomes.

undertaking specific instructor-designed learning activities (see Table 1). The top row of Table 1 shows the four modes of engagement behaviors, behaviors with their characteristic descriptors such as *receiving, manipulating, generating,* and *dialoguing*. The left column shows three different instructional contexts: listening to a lecture, reading a text, and observing a video. The cells of Table 1 list examples of engagement activities within each mode in the context of the three instructional tasks.

Table 1 shows that for any instructional task, students can engage in different modes of activities or different activities within the same mode. For example, while reading a text or watching a video, students can be engaged in a *passive*, *active*, *constructive*, or *interactive* way depending on what overt behaviors they do while reading or while watching a video. As depicted in Table 1, while reading, students can be reading silently, reading and underlining, reading and taking notes in one's own words, or reading and asking a friend questions about the materials. Similarly, while reading, students can engage in one of several types of activities that all fit into a single (e.g., *constructive*) mode, such as self-explaining, creating concept maps, and comparing and contrasting. Thus, instructional tasks are orthogonal to engagement mode.

Passive mode of engagement. In our taxonomy we define a *passive* mode of engagement as learners *being oriented toward* and *receiving* information from the instructional materials without overtly doing anything else related to learning. For example, paying attention and listening to a lecture without overtly doing anything else (i.e., listening without taking notes) is a *passive* engagement behavior. We acknowledge that it is possible for students to be covertly processing the materials deeply while listening to a lecture (or observing a video), even though overtly appearing only to be passively engaged (this point is elaborated later).

In the literature, studies that examine students' interactions with computer-based learning environments often discriminate students' engagement in a binary way, in terms of whether students are oriented toward and paying attention (i.e., zoning out or being off-task). In these studies (e.g., see D'Mello, Olney, Williams, & Hays, 2012), "being engaged" and looking at the instructional materials map to our *passive* mode, which is the lowest level of engagement in our framework.

To differentiate our use of the term *passive* (as well as the terms *active, constructive,* and *interactive*), these terms are italicized, whereas the general dichotomy in the literature of "active" and "passive learning" are in quotes. See Table 1, column 1, for other *passive* activities in the context of other instructional tasks.

Active mode of engagement. Learners' engagement with instructional materials can be operationalized as *active* if some form of overt motoric action or physical manipulation is undertaken. Using this criterion to characterize *active*, laboratory studies in the psychology and cognitive science literature provide numerous examples of *active* overt behaviors, such as *rotating objects* to inspect parts of it more carefully (James et al., 2002), *looking for objects* in an environment from a description of it (Kastens & Liben, 2007), and so forth. For various performance measures, *active* behaviors excel *passive* behaviors.

TABLE 1
Examples of Learning Activities by Mode of Engagement

	PASSIVE *Receiving*	ACTIVE *Manipulating*	CONSTRUCTIVE *Generating*	INTERACTIVE *Dialoguing*
LISTENING to a lecture	Listening without doing anything else but oriented toward instruction	Repeating or rehearsing; Copying solution steps; Taking verbatim notes	Reflecting out-loud; Drawing concept maps; Asking questions	Defending and arguing a position in dyads or small group
READING a text	Reading entire text passages silently/aloud without doing anything else	Underlining or highlighting; Summarizing by copy-and-delete	Self-explaining; Integrating across texts; Taking notes in one's own words	Asking and answering comprehension questions with a partner
OBSERVING a video	Watching the video without doing anything else	Manipulating the tape by pausing, playing, fast-forward, rewind	Explaining concepts in the video; Comparing and contrasting to prior knowledge or other materials	Debating with a peer about the justifications; Discussing similarities & differences

In the context of learning measures, undertaking *active* activities have also been shown to exceed *passive* activities, such as when students *manipulate* some parts of the learning materials, by *pointing* to or *gesturing* at what they are reading or solving (Alibali & DiRusso, 1999), *pausing and rewinding* parts of a video tape (in order to review certain selected parts of the tape; Chi, Roy, & Hausmann, 2008), *underlining* certain text sentences that they think are important (Katayama, Shambaugh, & Doctor, 2005), *copying* some of the problem solution steps (VanLehn et al., 2007), *mixing* certain chemical amounts in a hands-on laboratory (Yaron, Karabinos, Lange, Greeno, & Leinhardt, 2010), *choosing* a justification from a menu of options (Conati & VanLehn, 2000), and so forth.

By restricting *active* activities to those that require some form of motoric behaviors that cause *focused attention while manipulating*, we are distinguishing them from overt activities that are carried out mindlessly. For example, if students are asked to read passages out loud in their entirety (Oakhill, Cain, & Bryant, 2003), then this *reading out loud* activity is closer to *passive* than *active* because attention was not focused on some specific passages or sentences. However, it is not always clear-cut when focusing is or is not involved while being *active*. For example, if students read some parts of the passages with greater emphasis (such as in a louder voice), then it would appear that focused attention is involved. Sometimes it is ambiguous whether an activity should be classified as *passive* or *active,* unless one can determine whether it causes focused attention.

Constructive mode of engagement. Our taxonomy defines *constructive* behaviors as those in which learners *generate* or produce additional externalized outputs or products beyond what was provided in the learning materials. Thus, a characteristic descriptor of the *constructive* mode is *generative*. To meet the criteria for *constructive*, the outputs of generative behaviors should contain new ideas that go beyond the information given; otherwise such behaviors are merely *active/manipulative*. For example, in a *constructive* behavior such as self-explaining, learners are articulating what a text sentence or a solution step means to them (Chi, Bassok, Lewis, Reimann, & Glaser, 1989), by generating inferences that are not explicitly stated in the text sentence, or providing justification for the step. Both the inferences and the justification go beyond the provided information. If the student's self-explanation is verbatim to what was read, then the student is only repeating or self-explaining *actively* rather than *constructively*, because no new information is provided.

For another example, suppose a student is reading a worked-out example solution of a physics problem in a text. If this example solution does not have any diagrams in it and the student draws a free-body diagram, then the student would have *constructed* an output. On the other hand, if the solution example had a diagram already and the student copied it, then the student was being *active*, because the diagram was already present.

Constructive activities include activities such as *drawing* a concept map (Biswas, Leelawong, Schwartz, & Vye, 2005; Novak, 1990); *taking* notes in one's own words (Trafton & Trickett, 2001); *asking* questions (Graesser & Person, 1994); *posing* problems (Mestre, 2002); *comparing and contrasting* cases (Schwartz & Bransford, 1998); *integrating* two texts (Rouet, Britt, Mason, & Perfetti, 1996), or *integrating* text and diagrams (Butcher, 2006), or *integrating* across multimedia resources (Bodemer, Ploetzner, Feuerlein, & Spada, 2004); *making* plans (Pea & Kurland, 1984); *inducing* hypotheses and causal relations (Suthers & Hundhausen, 2003); *drawing* analogies (Chinn & Malhotra, 2002); *generating* predictions (Schauble, Glaser, Duschl, Schulze, & John, 1995); *reflecting* and *monitoring* one's understanding and other self-regulatory activities (Azevedo et al., 2006); *constructing* time lines for historical phenomena (Dawson, 2004), self-explaining (Chi et al., 1989), and so forth. These behaviors can all be classified as *constructive* because they satisfy the criterion that learners generate some sort of external output such as notes, hypotheses, justifications, questions, predictions, self-evaluations, and time lines, all of which contain additional ideas that go beyond the original learning materials.

Our use and operational definition of *constructive* or *generative* is more specific and concrete than the ways the terms "constructive" and "constructionism" are used in the broader literature. In the broader literature, "constructive" has several meanings. One meaning reflects a theoretical perspective that students should "construct" their own understanding rather than "being told" or "instructed" by a teacher (Bruner, 1961; Papert, 1980; Piaget, 1930). In this context, "constructionism" is contrasted with "instructionism" (Kafai, 2006). A second use of *construction* contrasts "being told" or "direct instruction" with "discovery learning," in which students construct the rules and relations they need. A third use of "constructivism" aligns more closely with our definition, in referring to it as "knowledge construction by the individual," or "social co-constructions of knowledge" by two or more individuals. This view does align with our view of *constructive* in assuming that "knowledge is actively constructed by the learner" (Loyens & Gijbels, 2008, p. 352). However, our definition is still more concrete because we operationalize it explicitly as generating outputs that go beyond the information presented.

Besides being more concrete and explicit, by focusing our definition from the learner's perspective, we also unconfound the role of the instructor from the role of the learners. For example, it is sometimes assumed that the "constructivist" perspective is experiencing a backlash in recent educational literature because at times learners benefit from instructors' direct scaffolding (Kirschner, Sweller, & Clark, 2006; Mayer, 2004). However, such findings are consistent with our framework because instructional scaffolding allows students to be more *generative* in that the scaffolding often requires the students to respond in a way that is usually *constructive*. For example, Chi, Siler, Jeong, Yamauchi, and Hausmann (2001) argued that tutoring is effective because many of the tactics used by tutors can be reframed as prompting or encouraging tutees to construct knowledge, either through the use of content-free prompts or scaffolding prompts such as pumping for "what else," hinting ("So, it's kind of leaving out the lungs there?"), and fill-in-the-blank kinds of requests (such as "OK, and it goes from the atrium to the [pause]?"). These prompts encourage students to respond more constructively because they require more than a yes/no response (which we consider to be an *active* response only).

Interactive mode of engagement. The term "interaction" has been used loosely in the literature to refer both to interpersonal activities and to characteristics of human–computer systems (i.e., if a learning environment or computer-based system expects a response from the user, and provides feedback to that response, then often it is labeled an interactive system). In layman's terminology, *to engage* means to interact with a device, a learning environment, or system without differentiating whether the interactions of the leaner is *active* or *constructive.* In this article, we operationalize *interactive* behaviors to dialogues that meet two criteria: (a) both partners' utterances must be primarily *constructive*, and (b) a sufficient degree of turn taking must occur. We do not restrict who the partners can be, provided that the criteria are met. Examples include a learner talking with another person who can be a peer, a teacher, a parent, or computer agent (assuming the computer agent responds in a content-relevant way).

Because being *interactive* requires that each person of the group contributes *constructively,* this means that if two people interact only in some physical or motoric way without discourse, such as two students copying each other's homework solutions, then they are interacting only *actively* and not *constructively.* The available evidence agrees with our interpretation, in that when two people work together, learning seems to occur when there is verbal discussion rather than only motoric or physical interactions (Milrad, 2002). Therefore, for now, we restrict our definition of *interactive* behaviors to discourse or *dialoguing* (Salomon & Perkins, 1998).

Are all dialogues the same, and are they all equally *interactive* and beneficial to learning? As previously indicated, according to the ICAP framework, dialogues are truly *interactive* only if each speaker's utterances generate some knowledge beyond what was presented in the original learning materials and beyond what the partner has said; thus, both partners need to be *constructive*. It is in this sense that *interactive* subsumes *constructive.* Damon (1984) and Rafal (1996) similarly defined dialogues as truly *interactive* when they consist of mutual exchanges of ideas between two individuals resulting in new ideas that neither individual knew initially nor could generate alone. Such mutual exchanges of ideas imply that both partners make substantive contributions to the topic under discussion, such as defending and arguing a position (Schwarz, Neuman, & Biezuner, 2000), criticizing each other by requesting justification (Okada & Simon, 1997), asking and answering each other's questions (Webb, 1989), explaining to each other (Roscoe & Chi, 2007), and elaborating on each other's contributions (such as clarifying, building upon, correcting, etc.; Hogan, Nastasi, & Pressley, 1999). This is an ideal conception of an *interactive* or *joint dialogue* pattern in that both partners make substantive contributions.

In addition to being *constructive*, a dialogue must have a sufficient frequency of turn taking to meet our definition of *interactive*. Two students who take turns giving minilectures to each other, even if they are being *constructive,* will likely not reap the same benefits as two students who frequently interject to ask each other questions, make clarifications, and so forth. We hypothesize that by frequently taking turns, it will be easier for students to incorporate their partners' understanding of the domain and to make adjustments to their own mental model, because such dynamic and ongoing exchanges have the dual advantages of allowing for more frequent revisions on smaller components of knowledge. We have made the same argument for why self-explaining facilitates learning, because it fosters ongoing integration of minute pieces of new knowledge (Chi, de Leeuw, Chiu, & Lavancher, 1994, p. 474), and there is some evidence in support of our claim of the importance of taking frequent turns (Chi & Kang, 2014).

Another type of dialoguing that is not very *interactive* is when one partner dominates and generates most of the substantive contributions and the other partner merely agrees or contributes with backchannel responses (e.g., "mmm," "uh-huh," etc.). In such cases, the dominant partner or the "speaker" may be learning more than his or her "listening" partner because the "speaker" is being *constructive,* whereas the "listener" is simply being *active* by agreeing and affirming to some of the "speaker's" contributions. We have referred to this pattern of dialoguing as *individual dialogue* (Chi & Menekse, in press). This may explain why some collaborative learning studies found no advantage of collaboration over individual learning (Barron, 2003; Yetter et al., 2006). Thus, *individual dialogue* pattern tend to promote more learning for the dominant speaker, whereas both partners can benefit from *joint dialogue* pattern.

The same operational definition of interactions can be applied to the level of interactivity of a computer-based system (such as an intelligent tutoring system or a learning environment). Because such systems often require a response from the student, are they all truly *interactive*? No, because our definition of interaction is learner centered and applies to the kind of responses that a learner makes. Suppose a student's response to a computer-based learning environment consists of selecting an answer from a menu of choices; selecting is only *active* in our taxonomy in that the student does not generate a product. Therefore, responses to a so-called interactive computer-based system can also be classified as *passive, active, constructive,* or *interactive,* depending on the level of cognitive engagement required of the learners.

Assumptions Underlying the ICAP Taxonomy

Our taxonomy of four modes of overt behaviors requires eight assumptions that are described next.

Content relevant. First, the classification into the four modes of activities assumes that students are doing activities that are content relevant. For example, if students are gesturing as part of an *active* task (such as when teachers ask students to point at the important equations on the whiteboard), if students point instead at random figures on the whiteboard, then their behavior would not be considered a beneficial *active* one. Similarly, when students are *constructing* or *interacting*, their outputs or dialogues must be content relevant for the activity to be considered truly *constructive* or *interactive.*

Intended versus enacted. Our taxonomy starts with classifying the intended overt behaviors as designated by an activity. However, asking students to carry out a behavior does not guarantee that the expected behavior will actually be carried out in the intended way. For example, if a student is asked to be *constructive* by generating a summary, the student can reduce the task to an *active* one by summarizing using a delete-and-copy strategy (i.e., delete the irrelevant sentences and copy the relevant ones). To know accurately what mode of behavior a student is carrying out, teachers need to judge the students' products from an ICAP perspective (i.e., evaluate the summary).

Besides a mismatch between the *intended engagement mode* and the *enacted engagement mode* (the enacted behavior), there can also be a misalignment between the overt display of behavior and the covert processes. For example, a student may appear to be only *actively* involved in learning (such as underlining key concepts in a book) and yet in reality the student is covertly *constructive* (such as thinking deeply about and self-explaining the material). Nevertheless, we rely on overt manifestations as our measure of engagement not only because it is the best noninvasive measure available but also because it provides concrete criteria for teachers to use when designing classroom and homework activities. This design feature of ICAP is presented in the third section of this article.

Analyzing the outputs. Because there may be a discrepancy between the intended engagement mode and the enacted engagement mode, a content analysis of learners' responses is necessary to confirm the actual mode of an activity. For example, in dialogues, some dyads might engage in *individual* dialogue (in which only one partner is being *constructive*) instead of in *joint* dialogue, where both partners are *constructive*. As a result, an analysis of the content of the products or the behavioral transcriptions is necessary to correctly classify an activity. From the teacher's perspective, such an analysis might inform how to revise the activity so that it aligns more with her intention.

Advantages of externalized outputs. Although we acknowledge that students can be generative covertly without externalizing any outputs, we also assume that there are practical, cognitive, learning, and epistemic advantages to externalizing outputs as static products (e.g., notes, diagrams). From a practical standpoint, externalized static outputs provide performance data for teachers and researchers to analyze and verify that students are completing the tasks and doing them in the way intended by the design of the tasks. From a cognitive perspective, one advantage of externalizing outputs is that it can help overcome cognitive load (Sweller, 1994). For example, it is difficult to generate a large or complex concept map in memory, so offloading it externally onto a tablet or paper reduces the memory requirement. From a learning perspective, the externalized outputs become new materials that a student can further examine to infer new knowledge or learn from errors. Moreover, the process of externalizing provides students the opportunity to monitor whether in fact they know the materials. From an epistemic perspective, an advantage to externalizing outputs is that it commits a student to that knowledge and gives her a sense of ownership of that knowledge (Kitchner, 1983).

Greater probability. Our taxonomy assumes that it is only *more likely* that the overt behaviors are an approximation to how the students are actually engaging with the learning materials most of the time. For example, while listening to a lecture, if students are doing nothing else besides orienting and receiving, then it is *more likely* that they are engaging *passively* while listening to the lecture. We are not claiming that they are engaged *passively* the entire time that they are listening (e.g., at times students could be covertly self-explaining while appearing to be *passive*.) Moreover, even though students may not always complete an activity in the way it was intended, we assume that students are *more likely* to engage with the learning activity at the prescribed level.

Independence. Although this point was articulated earlier in the context of defining our taxonomy, this important assumption is reiterated here briefly. Engagement activities are orthogonal to instructional tasks, in that a student can undertake various engagement activities while learning, regardless of how instruction is delivered. Thus, the mode of engagement, referring to what students are actually doing while being instructed, is independent of the instruction itself.

Hierarchy. From the way the behavior of each mode is operationally defined, there is a hierarchy in our taxonomy so that a higher mode subsumes a lower mode. That is, *interactive* behaviors subsume *constructive* behaviors, and *constructive* behaviors subsume *active*, and *active* behaviors subsume *passive.* For example, we pointed out earlier that to meet our definition of *interactive*, both partners must be *constructive.* Likewise, to be *constructive,* such as drawing a diagram, one must also be *active* (i.e., engaging in the motoric act of drawing). Finally, being *active,* such as underlining a sentence in a text, requires that the learner focuses on and orients toward the underlined sentences, thus *active* behavior includes *passive.*

Intermode boundary. The boundaries between the modes of activities are not meant to be totally rigid. Obviously some activities may be difficult to classify and fall on the boundary between two categories, because accurate classification also depends on the actual cognitive processes carried out by the student. One of the most difficult activities to classify is problem solving, because it depends on how a specific student is doing it and the context in which it was taught. For example, solving an algebra problem can be classified as an *active* behavior when the student matches the conditions of application of an equation and merely applies the formula of an equation (e.g., plug and chug). However, solving a novel or more difficult problem may be classified as *constructive* if a student has to rederive an equation, or reconceptualize certain components of a problem, such as reconceptualize two masses of a physics problem (one block on top of another block) as a single compound body. There are ways to determine more precisely the students' cognitive processes when undertaking an activity in a controlled laboratory setting, but this cannot be done easily in a classroom setting.

In summary, when these assumptions are met, we assert that the ICAP taxonomy is a useful and powerful way to classify engagement activities.

Knowledge-Change Processes Underlying Each Mode of Engagement and the Expected Changes in Knowledge

In Chi (2009), we briefly introduced the possible processes underlying each mode of engagement. In this article, we elaborate and refer to these engagement processes as "knowledge-change processes," then describe (a) what they are, (b) the assumptions we make with respect to these knowledge-change processes, (c) the resulting changes in knowledge, and (d) the expected cognitive and learning outcomes as a result of the knowledge changes.

Learning is typically conceived of as causing some changes, and these changes can occur at various grain sizes, such as systemic changes in terms of people's participation in social or cultural activities; behavioral changes in terms of what people do such as eating a healthier diet; knowledge changes in terms of what people know, causing them to be able to solve some problems and explain some phenomena; and cellular changes in terms of changes in the activation patterns of their brain's neurons. These various grain sizes of change require different measurements to detect whether changes have occurred. Because the ICAP framework is intended to understand how cognitive engagement can enhance an individual student's learning of complicated school subject matters, the appropriate grain size to talk about learning is the knowledge level; thus we use the terms "changes in knowledge" and "knowledge-change processes" to reflect our assessment of students' understanding after learning and the processes by which they acquired this knowledge.

By our definition, knowledge-change processes are dynamic processes that students engage in while learning new information. We simplify and postulate that there are four broad types of knowledge-change processes that are relatively distinct from each other, and each can be associated with a *passive, active, constructive,* or *interactive* mode of behavior. We first summarize the four types of knowledge-change processes by providing a single-word label for each.

- *Store*: new information is stored in an isolated way (*passive*).
- *Integrate*: new information activates relevant prior knowledge and while storing, new information is integrated with activated prior knowledge (*active*).
- *Infer*: new information is integrated with activated prior knowledge, and new knowledge is inferred from activated and integrated knowledge (*constructive*).
- *Co-Infer*: Each learner *infers* new knowledge from activated and integrated knowledge and iteratively infers knowledge with new inputs from conversational partner(s) (*interactive*).

Processes and changes for passive. In a *passive* mode of engagement, the knowledge change processes can be postulated to be isolated "storing" of the received information in an encapsulated manner. Storing information in an isolated way means that new knowledge is not integrated with existing or prior knowledge. The consequence of isolated storing of information is that the newly learned encapsulated knowledge is "inert" (Whitehead, 1929), meaning that access to that knowledge is possible only when a specific cue to activate it is given, such as an exam question expected from a specific chapter (Brown & Palincsar, 1989, p. 394), or the same context is provided. An example of encapsulated knowledge can be seen in a protocol snippet collected by Scardamalia (1992; cited also in Chi et al., 1994). A student had read about how blood clots. When the student was asked to explain how a cut heals, the student first responded in the following way, based on the materials she had read: "When you get a cut it bleeds. In your blood there are things called platelets. Platelets made a shield on your cut, it is called a scab, it protects your skin when it is healing." So the student obviously had stored the newly learned information. After this answer, the student was asked to use her own theory to answer the same question. Here is the student's second response to the same question: "My theory is that when you get a cut the blood vessel that got cut dies and the heart stops sending blood to that vessel until it heals." Thus, the student's old ideas were not integrated with the newly learned knowledge, suggesting that the new ideas were stored in an isolated or encapsulated way.

Processes and changes for active. When students engage in an *active* way, such as highlighting key sentences when reading on a computer screen, such manipulations can be interpreted as ways of *emphasizing* certain parts of the learning materials. Such emphases may cause learners to activate a body of knowledge (such as a skeletal or partial schema or a mental model) that is relevant to the emphasized sentences. In addition, once a relevant schema is activated, learners can assimilate or *integrate* new information into the activated schema, allowing the learner to fill gaps in the schema, thereby making it more complete (Bartlett, 1958). Gap filling is a process that is easily modeled computationally and has been explored extensively (Conati & Carenini, 2001; VanLehn, Jones, & Chi, 1992).

Processes and changes for constructive. Constructive behaviors often require the processes of "inferring" (e.g., induce, deduce, and abduce). Inferring also includes *revising, repairing, reorganizing*, and *reflecting*. Inferring can be thought of generally as a process of elaborating, such as adding more details or qualifications. *Revising* and *repairing* can be thought of as the processes of changing what was initially stored but is incorrect. *Reorganizing* can be changing hierarchical relations, or pattern of relationships. In contrast, *reflecting* may require a student to evaluate what she or he does or does not understand in order to carry out a constructive activity such as generating a question.

Finally, constructive activities also occur when learners integrate various parts of the learning materials (including comparing, combining/linking, discriminating/contrasting information from disparate sources or different paragraphs within a passage, reasoning analogically, etc). Thus, the knowledge-change processes of constructive activities can involve multiple types of inferring processes.

Processes and changes for interactive. As described previously, individuals interacting in dialogues can participate in *individual* or *joint* dialogue patterns, but only the *joint* patterns are truly interactive. In these interactive joint dialogues, each member of the dyad must be *constructive* while interacting, thereby engaging in the cognitive processes of *activating, integrating*, and *inferring.* But in addition, in *joint* dialogues, each speaker's *storing*, *activating, integrating*, and *inferring* processes further benefit from the contributions provided by a peer, such as their elaborations, feedback, suggestions, and perspectives. Next is a hypothetical example of such *co-inferring processes*.

Suppose Alice (Learner A) and Bob (Learner B) are both exposed to base instructional information that they both encode (Knowledge 1) and are collaboratively learning. Alice starts the interaction by activating her own prior knowledge relevant to Knowledge 1 (Schema A) and makes an inference from Schema A (Inference 1). Alice shares Inference 1 with Bob, which is information that goes beyond what was originally presented in Knowledge 1. Bob then activates his own schema (Schema B) relevant to Knowledge 1 plus Inference 1, and based on his Schema B, makes an inference (Inference 2) that goes beyond the original Knowledge 1 + Inference 1. Alice can now make an inference and contribution based on Inference 2, plus the original presented Knowledge 1, and her own Inference 1. So in continued interactions, partners benefit from the inferences of their partners, and each subsequent inference can incorporate the inferences made by the other partner as well as integrating it with one's own contributions and inferences. Thus, the advantage of such *co-inferring* processes is that each partner can benefit from the inferences of the other partner in a cumulative and spiraling way.

In general, the knowledge-change processes for each higher mode subsume the processes for the lower mode, in the same way that the behavior for the higher modes subsumes the behavior of the lower modes.

Assumptions Underlying Knowledge-Change Processes

Just as we made assumptions regarding the classification of overt engagement behaviors, we also made a number of assumptions regarding the possible underlying knowledge-change processes as involving *storing* information in an isolated way, *activating* prior relevant knowledge, *integrating* new information with prior knowledge, and/or *making inferences* from it. Of course, this can be done in various permutations and repetition of *storing, integrating,* and *inferring,* either doing it alone or with a peer. Our main assumption is not whether we have postulated accurately the knowledge-change processes for each mode but rather that they are relatively different for each mode. Besides this main assumption that engaging behaviors elicit these distinct fundamental knowledge-change processes, there are three other important assumptions about the nature of knowledge-change processes.

Distinct from learning processes. The first assumption is that knowledge-change processes underlying engagement are distinct from the cognitive processes of learning in the context of a specific instructional/learning task itself, such as solving a problem or understanding a diagram or reading a passage. For example, the cognitive processes involved in the learning task of problem solving includes creating and searching a problem space, whereas the processes of understanding a diagram might include perceiving the spatial relations among the components, mentally rotating the angles in the diagram, and the processes of reading include decoding and identifying word meaning. These processes, unique to the specific learning tasks of solving problems, understanding diagrams, and reading, are distinct from the knowledge-change processes of engagement activities such as storing, integrating, and inferring while engaged in self-explaining or drawing a diagram. Our assumption is that the knowledge-change processes of storing, integrating, and inferring, supplement these task-specific learning processes.

Same set of knowledge-change processes. The second assumption is that, although the types of activities within each mode may be different (e.g., in reading a text, the various *active* engagement activities could be underlining text sentences or summarizing by deleting irrelevant sentences), their underlying knowledge-change processes are the same. That is, underlining text sentences causes the learner to focus attention on the underlined sentences in the same way that deleting irrelevant sentences when producing a text summary causes the learner to focus attention on the remaining sentences of the summary. Although the activities of underlining and summarizing are different, the knowledge-change process of focusing attention is the same. As another example, for *constructive* activities such as self-explaining and drawing a concept map, the knowledge-change processes for both activities include integrating and inferring new knowledge. It is precisely the similarity in the knowledge-change processes for each mode of engagement that allows us to generalize our predictions of learning outcomes as a function of the mode of engagement across various activities, to be explained next.

Cognitive outcomes resulting from the knowledge-change processes. Finally, we assume that different knowledge-change processes cause different changes in one's knowledge, resulting in different cognitive outcomes, such as being able to *recall, apply, transfer,* and *co-create*. For example, if a student has only learned the material in an encapsulated way (i.e., the information was *stored* in an isolated way while learning), then the student can only *recall* facts. The cognitive outcome for each of the four modes is elaborated in the next section along with the learning outcomes. Thus, we attribute the outcome of learning to the way information was initially learned and the changes in knowledge resulting from it.

In Table 2, the first row (labeled Example activities) gives two examples of behaviors in each mode, the second row postulates the underlying knowledge-change processes for each mode of engagement, and the third row postulates the expected changes to the knowledge as a result of the engagement processes undertaken.

A Hypothesized Order of Learning Outcomes: Interactive > Constructive > Active > Passive (ICAP)

The preceding section described the underlying knowledge-change processes postulated for each mode of activity (summarized in the second row of Table 2), and the associated knowledge changes expected for each mode (the third row of Table 2). On the basis of these knowledge changes, different cognitive and learning outcomes will result. In this section, we describe the expected relative cognitive and learning outcomes for each engagement mode, resulting in a hypothesis about learning that is referred to as ICAP (as shown in the fourth and fifth rows of Table 2).

For *passive* activities, because new knowledge is encoded in an isolated or encapsulated way during learning, the outcome is that such inert knowledge can be retrieved and recalled, but only when relatively the same cue or context is given. Such isolated *storing* is adequate for learning that does not require integration with prior knowledge, such as learning a specific procedure of how to operate an ATM machine. Knowledge stored in an isolated way can be retrieved and reused later when the same context is provided (such as seeing another similar ATM machine). It can also be retrieved when using episodic tags. For example, a student might be able to solve a problem like the one the teacher worked out at the blackboard if the student was reminded that it is the problem demonstrated last Tuesday.

For *active* activities in which students manipulate information while learning, this manipulation causes them to activate the relevant prior knowledge corresponding to the emphasized information (such as a related schema). This then allows them to assimilate and fill gaps in their schema, making their activated schema more complete and strengthened. Students can more readily retrieve the knowledge and apply it in new context, such as when solving problems somewhat similar to what they have learned, explaining similar concepts, and so forth. In short, learning with *active* engagement can be quite substantial because significant knowledge completion has occurred. One could say that they have achieved, at minimum, a shallow understanding.

For *constructive* activities, because learners typically generate inferences and relations for conceptual knowledge, and rationales and justifications for procedural knowledge, not only does their knowledge or schema become more complete and strengthened, but in fact their schema can be more enriched and coherent because it can be substantially

TABLE 2
Example Activities, Knowledge-Change Processes, Knowledge Changes, Cognitive Outcomes, and Learning Outcome by Mode of Engagement

CATEGORY *Characteristic*	PASSIVE *Receiving*	ACTIVE *Manipulating*	CONSTRUCTIVE *Generating*	INTERACTIVE *Dialoguing*
Example activities	Listening to explanations; Watching a video	Taking verbatim notes; Highlighting sentences	Self-explaining; Comparing and contrasting	Discussing with a peer; Drawing a diagram with a partner
Knowledge-change processes	Isolated "storing" processes in which information is stored episodically in encapsulated form without embedding it in a relevant schema, no integration	"Integrating" processes in which the selected & emphasized information activates prior knowledge & schema, & new information can be assimilated into the activated schema.	"Inferring" processes include: integrating new information with prior knowledge; inferring new knowledge; connecting, comparing & contrasting different pieces of new information to infer new knowledge; analogizing, generalizing, reflecting on conditions of a procedure, explaining why something works.	"Co-inferring" processes involve both partners taking turns mutually creating. This mutuality further benefits from opportunities & processes to incorporate feedback, to entertain new ideas, alternative perspectives, new directions, etc.
Expected changes in knowledge	New knowledge is stored, but stored in an encapsulated way.	Existing schema is more complete, coherent, salient, and strengthened.	New inferences create new knowledge beyond what was encoded, thus existing schema may become more enriched; procedures may be elaborated with meaning, rationale and justifications; and mental models may be accommodated; and schema may be linked with other schemas.	New knowledge and perspectives can emerge from co-creating knowledge that neither partner knew.
Expected cognitive outcomes	Recall: knowledge can be recalled verbatim in identical context (e.g., reuse the same procedure or explanation for identical problems or concepts).	Apply: knowledge can be applied to similar but non-identical contexts (i.e., similar problems or concepts that need to be explained)	Transfer: knowledge of procedures can be applied to a novel context or distant problem; knowledge of concepts permit interpretation & explanations of new concepts.	Co-create: knowledge and perspectives can allow partners to invent new products, interpretations, procedures, and ideas.
Learning outcomes: ICAP	Minimal understanding	Shallow understanding	Deep understanding, potential for transfer	Deepest understanding, potential to innovate novel ideas

revised (accommodated), and perhaps more interconnected or linked with other schemas, facilitating *transfer*. Thus the knowledge structure changes substantially. One example of accommodation is the change from conceiving of the human circulatory system as a "single-loop" to a "double-loop" model (Chi et al., 1994). Learners can achieve such accommodation changes in their knowledge structure from *constructively* engaging with the materials (e.g., self-explaining).

In the context of conceptual domains, more enriched and interconnected understanding can provide generalizations of explanations, reasoning analogically, and so forth. In the context of procedural domains, providing justifications and rationales allows the procedures to be transferred and used in a new context. Thus, knowledge-change processes from *constructive* activities can deepen one's understanding of the materials to facilitate *transfer*. Transfer is difficult to achieve, as pointed out in a 2012 special issue on transfer (Goldstone & Day, 2012).

Finally, for *interactive* activities, the changes in knowledge as a consequence of reciprocally *constructive* or *co-inferring* interaction is that each peer's alternative perspectives, guidance, and challenges will improve and expand each other's knowledge in a cyclical dynamic way. In successful dialoguing, the contributions of each peer incorporate the inferences of their partner, potentially resulting in new knowledge that neither partner could have created alone. The resulting changes in knowledge are that novel and innovative ideas and perspectives may emerge that neither peer originally knew and could not have generated while working independently. So one could say that learners in this mode have achieved the deepest understanding and can perhaps *co-create* innovations.

As previously described, the different learning outcomes that we have postulated for each mode can be mapped or translated into differing levels of learning, such as minimal understanding, shallow understanding, deep understanding, and deepest understanding (as shown in the last row of

Table 2). Thus, an *interactive* mode of engagement (I) can enhance learning more than a *constructive* (C) mode, which is better than an *active* (A) mode, which in turn is better than a *passive* (P) mode. In short, we hypothesize that the relative learning levels will be ordered more or less in the direction of I>C>A>P. Although we could generate the ICAP hypothesis from the hierarchical nature of the behavior modes alone, postulating reasonable underlying knowledge-change processes for each mode and then seeing their hierarchical nature provide further validity to the hypothesized order of the I>C>A>P hypothesis.

SECTION 2: EMPIRICAL VALIDATION OF THE ICAP HYPOTHESIS

To recap, the ICAP framework defines and differentiates different modes of overt behaviors, postulates the underlying knowledge-change processes, and the learning outcomes for each mode. The learning outcomes can be predicted to vary from minimal, to shallow, to deep, and deepest. The ICAP hypothesis postulates that learning is enhanced to the greatest degree with *interactive* behaviors, followed next by *constructive* behaviors, followed next by *active* behaviors, and followed last by *passive* behaviors. Because the entire ICAP framework and its associated ICAP hypothesis is conceptually derived, it is essential to seek empirical validation of the hypothesis.

In our prior paper (Chi, 2009), evidence in support of the hypothesis consisted of a random assortment of mostly laboratory studies that examined a variety of tasks that participants could undertake, using a variety of outcome measures. The prior paper included studies that can be interpreted to be making pairwise comparisons between two modes of engagement. In this article, we present four types of additional evidence in support of the ICAP hypothesis. First, we present the results of a study from our lab that compared all four modes of learning, as the strongest test of the ICAP hypothesis. Second, we present two studies from the literature that compared three of the four ICAP modes. Third, we present additional studies that made pairwise comparisons between engagement modes (*passive*, *active*, *constructive*, and *interactive*) for three specific engagement activities: taking notes, creating concept maps, and self-explaining. For example, we look at the learning differences between self-explaining enacted in an *active* way compared to self-explaining enacted in a *constructive* mode. Finally, we present evidence from classroom studies. Thus, all the supporting evidence presented in this article comes from studies that were not cited in the prior paper.

A Study With Four Modes of Engagement

We have undertaken a study in our lab that included all four modes of engagement in the domain of materials science (Menekse et al., 2013). Aside from reading a short passage of background knowledge, the four modes or conditions consisted of reading a text passage (*passive*); reading and highlighting important sentences within the text *(active*); interpreting a graph that described the information contained in the text passage but without having read the text passage (*constructive*); and interpreting the same graph jointly with a peer (*interactive*), again without having read the text passage. The learning outcomes, comparing pretest and posttest for each condition, are shown in Figure 2. These results support the ICAP hypothesis and show that learning improves significantly at a rate of around 8% to 10% across each mode.

Two Studies With Three Modes of Engagement

In addition to the study we conducted that included all four modes of engagement, we found two studies the instructional conditions of which can be interpreted to fall across three modes of engagement according the ICAP framework. One lab study (Coleman, Brown, & Rivkin, 1997), within the domain of evolutionary biology, consisted of a three-person learning context in which each student took one of three different roles and the roles rotated. Specifically, the three roles were listening to another student (which we categorize as being *passive* because the overt behavior is receptive in nature), summarizing (which we categorize as *active* because they report that the protocols revealed that in many of the summaries participants did not go beyond the given text; Coleman et al., 1997, p. 360), or explaining the materials (*constructive*). Although the study did include an additional factor of whether or not students prepared their summaries and explanations with the intent to teach, we do not consider this an *interactive* activity because while teaching, their partners were instructed only to listen and not ask questions. Thus we averaged the scores across this factor for summarize-to-self and summarize-to-other to create an overall summarize score (and repeated this process for the explaining-to-self and explaining-to-other conditions). We plotted these new averages, and the results show the outcomes of students in each of these three roles, for two measures of near transfer and one measure of far transfer. Figure 3 shows that students did increasingly better from *passive* to *active* to *constructive*.

A second study that can also be reinterpreted to have manipulated three of the four modes of active learning was conducted in the domain of plate tectonics (Gobert & Clement, 1999). They compared learning gains among students who studied by drawing diagrams from text (*constructive*), writing summaries of the text (*active*), or only read the text (*passive*). As predicted by the ICAP hypothesis, the results showed that the *constructive* (diagram) group did better on both measures of spatial and causal knowledge than the *active* (summary) group did, which in turn did better than the *passive* reading group.

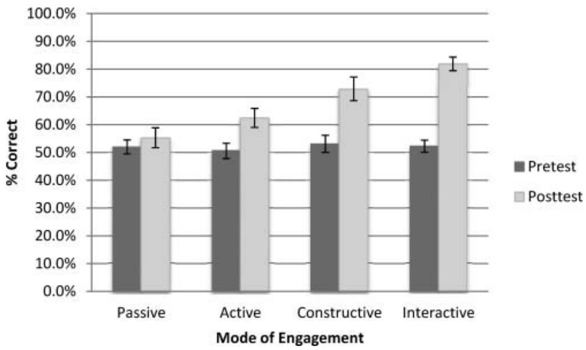

FIGURE 2 Pretest and posttest performance by mode of engagement (Menekse, Stump, Krause, & Chi, 2013).

Pairwise Comparisons Across Two Modes for Three Engagement Activities

As previously mentioned, we take a different approach to pairwise comparisons of studies. Instead of comparing two different activities undertaken in two different modes, as was done in Chi (2009), we use the ICAP taxonomy to compare how a specific activity is implemented, as was done in Fonseca and Chi (2011). Here we focus on three activities: note taking, concept mapping, and self-explaining. These three were chosen because students often engage in them while learning, and many studies have been carried out testing their effectiveness by manipulating how they are implemented. We can thus verify whether the levels of learning outcomes are in the predicted directions according to the ICAP hypothesis. The studies on the three engagement activities are summarized in Tables 3 to 5 and described next.

Note taking. In learning from text, taking notes is a common learning strategy that many students undertake on their own, even without a teacher's advice (Kobayashi, 2006). Several other activities are related to taking notes, such as underlining the most important sentences per paragraph or highlighting the key sentences using a computer interface. Taking notes is also somewhat similar to summarizing. In this section, we consider all these related activities (underlining, highlighting, and summarizing) as instances of note taking.

Previous studies have revealed a number of ways to operationalize the engagement activity of taking notes. Students may be asked to copy notes verbatim in their entirety, a *passive* activity because attention is not focused on some emphasized parts of the notes. Or students may summarize a text or lecture in a copy-and-paste way (in which case we would classify it as *active* because the copied parts can be focused on), or take notes in their own words (which is often *constructive* because the summary may include conclusions and inferences). In addition, recent technological innovations have supported students in collaborative note taking (*interactive*). As expected, the manner in which students take notes influences their learning. For example, a lab study that can be classified as *passive* versus *active*, demonstrated that *active* note taking in which students underlined the most important sentence in each paragraph was more effective than simply *passively* reading the paragraphs (Rickards & Friedman, 1978; see Peper & Mayer, 1986, for similar results with a video-watching task). In comparing *constructive* to *passive*, Trafton and Trickett (2001) saw a correlation between *constructively* taking notes using an online notepad and improved problem-solving performance. Namely, students who chose to take notes did better than those that *passively* studied the material.

Further evidence supporting the ICAP hypothesis can be seen in a study that compared *constructive* to *active* note taking. For example, taking free-form notes (*constructive*) is more effective than cutting and pasting short segments of the target text (*active*; Bauer & Koedinger, 2007). Finally, note taking can also be done interactively. Kam et al. (2005) developed a collaborating note-taking system that led students to take better notes compared to students who took notes alone. Students who worked collaboratively were more likely to include reflections and pose questions about the material than students who worked alone. However, the small sample size in their classroom study was insufficient to observe any significant learning differences

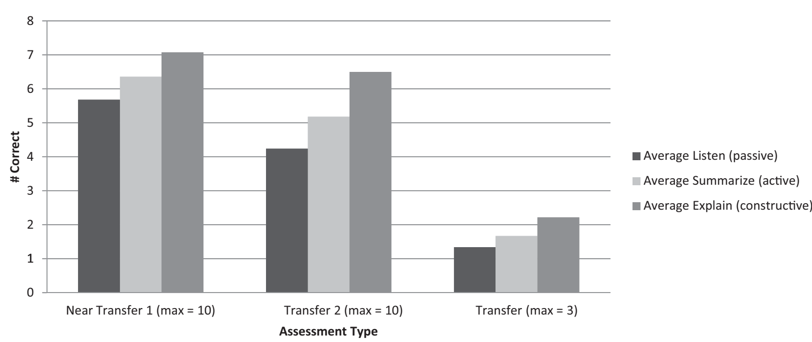

FIGURE 3 Near and far transfer posttest performance by mode of engagement (Coleman, Brown, & Rivkin, 1997).

TABLE 3
Pairwise Comparisons Between Different Modes of Note Taking

	Passive	*Active*	*Constructive*	*Interactive*
Passive	No known studies			
Active	Note taking > No note taking for far transfer (Peper & Mayer, 1986) Underlining > Reading (Rickards & Friedman, 1978)	Highlighting notes = Pasting Notes (Bauer & Koedinger, 2007) Summarize after each segment = Summarize throughout lecture (Peper & Mayer, 1986)		
Constructive	Taking notes with notepad > Not taking notes (Trafton & Trickett, 2001)	Free-form notes > Cutting/Paste Bauer & Koedinger, 2007)	No known studies	
Interactive	No known studies	No known studies	Collaboratively taking notes > Individually taking notes, in the quality of the notes, but no learning difference (Kam et al., 2005)	No known studies

Note. The light gray shading refers to the diagonal cells with activities implemented in the same mode; the dark gray shading refers to cells with redundant studies as the non-shaded ones.

between the *interactive* (collaborative) and *constructive* (individual) conditions.

The ICAP hypothesis can also explain results from studies that show no differences between different types of note-taking styles that fit the same ICAP category (e.g., both conditions are *active* or both are *constructive*). Bauer and Koedinger (2007) found no learning differences between students who used a computer interface to highlight segments of a text and those who used copy and paste to select segments of the text to create notes. This finding is expected according to the ICAP hypothesis because both are *active* activities that encourage students to physically select but not generate or integrate the information. Similarly, Peper and Mayer (1986) found no difference between students who took notes by summarizing after segments of a lecture and those who took notes throughout the lecture. Assuming that they took similar type of copy–paste notes, the ICAP framework predicts this lack of difference because both conditions can be considered *active*. However, to accurately interpret the results, we would need to examine the product of note taking. This and other limitations are discussed in Section 3. These studies are shown in the diagonal cells of Table 3.

TABLE 4
Pairwise Comparisons Between Different Modes of Concept Mapping

	Passive	*Active*	*Constructive*	*Interactive*
Passive	No known studies			
Active	Copy map > Read map (Willerman & Mac Harg, 1991)	No known studies		
Constructive	Correcting concept map > Reading text (Chang, Sung, & Chen, 2002) Concept maps + Lecture > Lecture only for higher level material and low PK students (Schmid & Telaro, 1990) Concept maps > Study + Discussion for ESL students (Chularut & DeBacker, 2004) Concept maps > Read + Discuss (Guastello, Beasley, & Sinatra, 2000)	Building a concept map from generating > Constructing a map via selection (Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005)	No known studies	
Interactive	Lecture + Collaboratively creating concept map > Lecture for science content assessment (Czerniak & Haney, 1998)	No known studies	Collaboratively building maps > individually building (Czerniak & Haney, 1998; Okebukola & Jegede, 1998)	Collaborative build concept map = Collaboratively build with 2 additional resources (van Boxtel, van Der Linden & Kanselaar, 2000)

Note. The light gray shading refers to the diagonal cells with activities implemented in the same mode; the dark gray shading refers to cells with redundant studies as the non-shaded ones.

TABLE 5
Pairwise Comparisons Between Different Modes of Self-Explaining (SE)

	Passive	Active	Constructive	Interactive
Passive	No known studies			
Active	SE worked example with menu > Study example only (Conati & Van Lehn, 2000)	No known studies		
Constructive	Explaining sentences > Repeating sentences (O'Reilly et al., 1998) Rereading with SE > Rereading only (Griffin et al., 2008) Prompted SE > Rereading twice (Chi et al., 1994) Explaining others' solution > Watching other solution (Pine & Messer, 2000) Studying incomplete examples > Studying completed examples (Atkinson et al., 2003; Stark, 1999)	No known studies	Explaining own answers = Explaining other's answers (Pillow, Mash, Aloian, & Hill, 2002)	
Interactive	No known studies	No known studies	Explain other group members contributions > Self-explain own contributions in groups (Kramarski & Dudai, 2009) Explain with partner > Explain alone (Hausmann et al., 2008)	No known studies

Note. The light gray shading refers to the diagonal cells with activities implemented in the same mode; the dark gray shading refers to cells with redundant studies as the non-shaded ones.

Concept mapping. Concept maps are graphical representations of knowledge where concepts are represented as nodes and are connected through labeled relations. For example, in a concept map of an ecosystem, two nodes may be "wolves" and "deer" and may be connected by an arrow labeled "prey upon." Creating such maps requires selecting relevant concepts, arranging them hierarchically, and determining the relationships between them. They are often used in classrooms as aids to understanding lectures and texts (Coffey et al., 2003**)**. More than 500 studies have investigated the effectiveness and applications of concept mapping (e.g., Novak, 1990; Novak & Gowin, 1984; O'Donnell, Dansereau, & Hall, 2002**)**. Two meta-analyses (Horton, McConney, Gallo, Woods, Senn, & Hamelin, 1993; Nesbit & Adesope, 2006) have found a positive effect of concept mapping on knowledge acquisition when compared to other learning activities.

Concept mapping can be implemented in a number of ways. Students can be asked to copy a concept map (an *active* way of engaging), which has been shown to be more effective for learning than reading a concept map (*passive*; Willerman & Mac Harg, 1991). Building or correcting a concept map can both be categorized as a *constructive* activity because students must generate a map based on learning materials or revise a map while correcting. Several studies have shown that generating or correcting concept maps is better than a variety of *passive* activities such as reading a text (Chang, Sung, & Chen, 2002), listening to a lecture (Schmid & Telaro, 1990), or whole-class discussions (Chularut & DeBacker, 2004; Guastello, Beasley, & Sinatra, 2000; we classify whole-class discussions as *passive* because the majority of students in class discussions are participating *passively*). In addition, there are studies that compare *constructive* to *active* version of concept mapping. For example, Yin, Vanides, Ruiz-Primo, Ayala, and Shavelson (2005) found that students who were asked to generate the labels for the links on a concept map (*constructive*) did better in total accuracy and map structure complexity than students who were asked to select the links to complete the concept map (*active*).

Likewise, collaboratively building concept maps (*interactive*) results in better learning than just listening to a lecture without building any maps (*passive*; Czerniak & Haney, 1998). Collaboratively building maps is also better than individually creating maps, supporting our prediction that *interactive* is better than *constructive* (Czerniak & Haney, 1998; Okebukola & Jegede, 1988).

There are also studies that show that when concept mapping is carried out at the same mode of engagement, the results show no differences in learning gains. For example, van Boxtel, van der Linden, and Kanselaar (2000) implemented two versions of *interactive* concept mapping. In one condition, after students received class instruction on a topic, they were asked to collaboratively create a concept map with no additional resources. In the second condition, after receiving instruction, they were asked to collaboratively create a concept map with two additional textbook chapters available. The results showed that students learned significantly in both conditions, but, as predicted by the ICAP hypothesis, there were no differences between conditions. These studies are shown in Table 4.

Self-explaining. Self-explaining is the activity of explaining an idea or concept aloud to oneself as one learns, or one can think of it as trying to make sense of the learning materials. Twenty years of research has consistently supported the finding that students learn better when they explain to themselves the materials they are studying, called the self-explanation effect (Chi, 2000). This is because self-explaining helps students clarify an idea by elaborating on it, or inferring new ideas. The self-explanation effect has been studied across age groups, domains, and instructional formats, such as reading a text passage or studying a worked-out example (Bielaczyc, Pirolli, & Brown, 1995; Chi et al., 1994; Ferguson-Hessler & de Jong, 1990; Hausmann & Chi, 2002; McEldoon, Durkin, & Rittle-Johnson, 2013; Renkl, Stark, Gruber, & Mandl, 1998; Siegler, 1995; Wong, Lawson, & Keeves, 2002).

Self-explaining can be carried out in a conceptual domain when students read a text passage, or it can be carried out in a procedural domain when students study a worked-out example. Worked-out examples are problem statements with line-by-line solution steps; they can be demonstrated by a teacher on a whiteboard or embedded in textbooks. In this section, we examine studies that involve self-explanation using texts, worked-out examples, and other related contexts.

Self-explaining is most often implemented in a *constructive* way because students are encouraged to make sense of the learning materials by generating inferences or by integrating new knowledge with prior knowledge. For conceptual domains, many studies have compared self-explaining implemented in a *constructive* way with other *passive* activities. For example, O'Reilly, Symons, and MacLatchy-Gaudet (1998) tested college students' recall and recognition ability for factual knowledge of the human circulatory system when students were asked to either explain what the sentence means to them (i.e., what new information it provides, and how it relates to what they already know, standard self-explaining *constructive* prompts) or repeat each sentence on the computer screen until the next fact appeared (physically repeating is an *active* condition because repeating causes more focused attention but does not generate new knowledge). As predicted, they found that recall and recognition were significantly higher for the self-explanation group when compared to the repetition group. Other studies comparing the same two conditions showed that self-explaining was better than reading the text twice for both meta-comprehension (Griffin, Wiley, & Thiede, 2008) and learning (Chi et al., 1994).

Self-explaining can also be used in a broader context than reading from a text. Pine and Messer (2000) for example, asked children to "explain" how the instructor was able to balance a beam (a *constructive* activity) compared with asking the children simply to "watch" (without making any comments) the instructor demonstrate how to balance the beam (*passive*). Children in the *constructive* "explain" condition improved in their ability to balance the beam significantly more than children in the *passive* "watch" condition.

For procedural domains, self-explaining has been traditionally implemented in a *constructive* way in the context of a worked example by having students explain each example step to themselves, such as justifying how a solution step follows from a prior step (Chi et al., 1989; Renkl et al., 1998). Other, more indirect *constructive* methods include providing incomplete worked-out examples (Stark, 1999), thus requiring students to infer the omitted steps (a form of self-explaining), resulting in more transfer than simply studying completed worked-out examples (a more *passive* mode of engagement). Incomplete examples can also involve fading (taking away) some of the example steps, especially as the learner gets more proficient at solving problems (Atkinson, Renkl, & Merrill, 2003). Thus, fading provides an opportunity for the learner to generate her own steps (more *constructive*) and this has been shown to be more effective than not fading (more *passive*).

Self-explaining can be implemented as an *active* activity by having students choose an explanation or elaboration among a set of options. An example of an *active* way of self-explaining is a computer-based example-studying task that asks students to select a justification for each solution step from a menu of choices (Conati & VanLehn, 2000). Such *active* ways of studying examples have been shown to be superior to a *passive* way, in which students simply studied examples without any prompts to justify or explain by choosing.

Kramarski and Dudai (2009) compared *interactive* versus *constructive* forms of self-explaining by training ninth-grade students to generate self-explanations to prompts such as, "What is my conclusion?" and "Is my explanation clear?" as they solved math problems in groups of four (*constructive*). In the *interactive* condition, students were prompted with questions that focused on building on and responding to the other group members' contributions, such as "How can I respond to my friend regarding the correctness of his/her explanation?" and "How can I modify my friend's solution and explanation?" The interactive groups scored significantly higher in their mathematical accuracy and problem transfer scores than the self-explain groups. A similar study had students studying worked-out examples in the context of solving physics problems using an intelligent tutoring system. Students either worked alone at the computer to solve the problems and were prompted to explain the solution steps presented on the screen (the *constructive* condition) or worked in dyads at the computer to solve the problems and were prompted to generate joint explanations to the solution steps (the *interactive* condition). The jointly explaining dyads outperformed the solo solvers on multiple measures: They answered faster, finished more problems in the allotted time, entered more correct entries, displayed a lower error rate, requested fewer hints when solving problems with the ITS, and so on (Hausmann, van de Sande, & VanLehn, 2008).

With respect to comparing versions of self-explaining implemented in the same mode, Pillow, Mash, Aloian, and Hill (2002) asked 4- to 5-year-old children to predict misinterpretation of ambiguous pictures. In one condition, the children were prompted to explain their own misinterpretations (explain-own condition is *constructive*), and in the other condition they were asked to explain the misinterpretations of a puppet viewing similar pictures (explain-puppet condition is also *constructive*). Both conditions are clearly *constructive*, and neither condition received feedback. As predicted by ICAP, there were no significant differences between the two groups in their final scores on the posttests. These studies are shown in Table 5. For more detailed analyses (with effect sizes) and comparisons of self-explaining implemented in different modes, see Fonseca and Chi (2011).

Classroom Studies

There are also many classroom studies whose interventions can be reinterpreted using the ICAP framework. These studies are summarized in Table 6, with the cells indicating the modes involved in the pairwise contrasts. For example, a classroom study comparing students who took notes using a partial, scaffolded outline to those who were given a handout with an almost verbatim transcript of the lecture showed that students who used the partial outline to take notes did significantly better on both immediate and retention tests than those who were given the comprehensive handout (Russell, Caris, Harris, & Hendricson, 1983). This finding is predicted by the ICAP hypothesis in that taking notes with a partial outline is more comparable to an *active* task because students must copy additional notes from the lecture to complete the handout, whereas having a complete transcript of the lecture requires no copying and reviewing it is thus *passive.* Other studies have compared and found that having students build concept maps (*constructive*) promotes better learning than having students participate in a whole-class discussion (*passive*; Chularut & DeBacker, 2004; Guastello et al., 2000). As mentioned earlier, we assume that a whole-class discussion is *constructive* only for the few students who are involved in the discussion. For the majority of students who are silently listening to the discussion, the activity is *passive.*

A classroom study by Hendricks (2001) found that discussing cause–effect relations with a peer is superior to observing the teacher identify cause–effect relations. Again, this is consistent with ICAP's prediction because discussing cause–effect relations with a peer is an *interactive* activity, whereas watching the teacher identify cause–effect relations is only *passive.* Similarly, if students are interacting in cooperative groups (*interactive*), this leads to greater learning than listening to a traditional lecture (*passive*; Ebert-May, Brewer, & Allred, 1997).

Two classroom studies have contrasted an *interactive* mode with an *active* mode. One study found that learning using the jigsaw strategy (in which each student within a group develops expertise on a subtopic and explains it to the other students within his or her group, thus being *interactive*) promotes more learning than classrooms where individual students gather information for themselves (*active*; Doymus, 2008). Another study found peer tutoring (*interactive*) to be better than filling out guided notes (*active*; Mastropieri, Scruggs, Spencer, & Fontana, 2003). Likewise, an *interactive* self-explaining group tends to performed better than a *constructive* self-explain-only group. In Kramarski and Dudai's (2009) study, students who self-explained and received feedback from each other (thus *interactive*) performed better than students who self-explained but did not receive feedback (*constructive only*).

TABLE 6
Pairwise Contrasts of Classroom Studies Involving Comparisons Between a Mix of *Passive*, *Active*, *Constructive*, and *Interactive* Engagement Activities

	Passive	*Active*	*Constructive*	*Interactive*
Passive				
Active	Taking scaffolded notes > Comprehensive handout (Russell et al., 1983)			
Constructive	Building concept maps > Whole class discussions (Chularut & DeBacker, 2004; Guastello, Beasley, & Sinatra, 2000)	Guided prompts + Solve > Solve (Whitten & Rabinowitz, 2010) Compare + Contrast + Write + Solve > Solve (Docktor et al., 2010)		
Interactive	Cooperative groups > Traditional lecture (Ebert-May, et al., 1997) Student–student discussion + small-group activities + feedback > Traditional lecture (Deslauriers, Schelew, Wieman, 2011) Discussing cause-effect with peer > Observing teacher identify cause-effect relationships (Hendricks, 2001)	Peer tutoring > Filling out guided notes (Mastropieri et al., 2003) Jigsaw groups > Individuals gathering information (Doymus, 2008) Generating with partner > Selecting with partner (Zheng & Linn, 2013)	SE + Receive Feedback > Self-explanation without feedback (Kramarski & Dudai, 2009)	

Note. The light gray shading refers to the diagonal cells with activities implemented in the same mode; the dark gray shading refers to cells with redundant studies as the non-shaded ones.

There are also a number of classroom studies comparing *constructive* learning activities with *active* activities. For example, two studies compared problem-solving alone (assuming this tends to be an *active* activity when solving in a plug-and-chug way) to solving problems with an added *constructive* activity (such as compare and contrast, Docktor, Strand, Mestre, & Ross, 2010; and responding to guided prompts, Whitten, 2011). The results showed that adding the *constructive* activity improves learning.

Classroom studies also tend to include multiple activities as one intervention. For the purpose here, we classify the intervention according to the highest activity mode. For example, in the study by Deslauriers, Schelew, and Wieman (2011), they compared traditional listening to lecture (*passive*) to an intervention consisting of predicting, reasoning, solving, critiquing, and peer discussion. We classify this intervention as an *interactive* one because it includes peer-to-peer discussion. As predicted, the intervention promoted more learning than listening to traditional lectures.

Classroom studies also illustrate a key point of the ICAP framework, namely, simply asking students to work together does not automatically make an activity *interactive.* As mentioned in the first section, to be *interactive*, students must work together on a constructive task. In a classroom study, Zhang and Linn (2013) compared the effects of two conditions on learning chemical reactions. In the first condition, students worked in pairs to draw molecules linking the atoms with lines (drawing is a generative activity). In the second condition, pairs of students physically selected static images that represented the correct molecular structure. Even though both conditions involved working with a partner, only the generate condition can be considered *interactive* according to the subsumptive nature of the ICAP framework; that is, *interactive* activities require partners to be *constructive* and go beyond the provided material. Thus, this comparison can be reinterpreted as comparing *interactive* to *active* and, as expected by the ICAP hypothesis, the *interactive* (drawing in pairs) condition showed greater learning gains than the *active* (select in pairs) condition.

Summary

This section reviewed studies in the literature to see if their results validate the predictions of the ICAP hypothesis. First, we reported results from a study conducted in our lab that compared all four modes and identified two studies in the literature that contained three conditions that can be reinterpreted using ICAP. Results from all three studies validate our predictions. Second, we showed that findings from studies that compared the same engagement activity implemented in different modes also support the ICAP hypothesis; that is, depending on how (or which mode) an activity is implemented influenced the learning outcomes. Third, we validated the prediction of the ICAP hypothesis in classroom intervention studies. In sum, the empirical evidence from both laboratory and classroom studies provide support for the ICAP hypothesis.

SECTION 3: CAVEATS, NEW TOOL, AND COMPARISONS WITH OTHER THEORIES

In this third major section, we consider several caveats and factors that need to be addressed. We also entertain ways that ICAP can serve as a tool for (a) explaining discrepant findings, (b) determining proper control conditions for experimental studies, (c) developing rubrics to code learning data, and (d) guiding instructional design. Finally, we briefly compare ICAP with other learning theories.

Factors and Considerations That May Override the ICAP Hypothesis

The ICAP taxonomy of a learner's overt behaviors is a gross first cut at predicting learning outcomes as a function of what students are asked to do. However, our analyses and predictions are based on only one aspect of a learning situation, the aspect of what students do to engage with the materials. Clearly there are alternative factors that influence how well students learn. In this section, we consider three factors that can override the predictions of the ICAP hypothesis.

Assessment measures. An important factor that may lead to unexpected results according to the ICAP hypothesis is the type of measures used to assess learning. In general, performance on shallow or easy assessment questions typically is not sensitive to various intervention efforts, and this is true for ICAP activities as well. This is because shallow questions that require only recall of information can be answered with *passive* receiving type of activities. For example, Aleven and Koedinger (2002) compared students who were prompted to explain in addition to solving geometry problems to those who solved the same problems without prompts to explain. They found no advantage for explanation on easy-to-guess items but significant benefits of explanation for hard-to-guess and transfer problems. Similarly, for the study (shown in Figure 2) in which we manipulated all four conditions within the same study (Menekse et al., 2013), we found that performance on easy multiple-choice questions was the same for all levels of engagement activities, but condition differences emerged on more challenging assessment items.

Thus, when the learning results are nonsignificant, one problem may be that the assessment measures are too shallow; thus, they are not sensitive enough to detect differences in learning outcomes as a function of the level of student engagement.

Domain. A second factor that can lead to violations of the ICAP hypothesis is the topic domain. Although rare, not all topics or skills necessarily show increased learning benefits with increased cognitive engagement. We next provide two examples and an explanation for why this might be the case.

One type of topic for which increased cognitive engagement may not be helpful is simple procedural domains for which the rules are arbitrary and cannot be logically deduced. For example, the English article system (teaching students when to use *a, an, the,* or no article) is governed by a complicated set of human-made rules and exceptions exist. In a computer-based tutoring system developed to teach students how to select the correct article, students were asked to select either the correct article or the article and the rule (explanation) for why that article was correct (e.g., "The noun is made definite by a prepositional phrase"). There were no learning differences between the two conditions (Wylie, 2011).

There are two ways to interpret this result. The simple one is that both are *active* selection tasks, so one would not expect a learning difference. A second interpretation is also possible, if we assume that selecting an explanation is closer to being *constructive* than selecting the answer alone, thus making the result counterintuitive. This second interpretation suggests that self-explaining article usage may not be a useful way to learn about the appropriate article to use because there is no way to rationalize or work through the reasons for when a particular article is used. In contrast, in the many examples of effective self-explanation discussed earlier, it makes sense to ask a *why* or a *how* question, because presumably with more elaborations and inferences, one can generate a reason. The same may be true for topics that are so challenging that students cannot bootstrap themselves by self-explaining into deeper or more correct understand because they lack the relevant schema to interpret the new information correctly. One such example in the science domain is the concept of emergence, because emergence is a novel idea that many students do not have relevant or related knowledge to understand. Therefore, *constructive* activities, such as answering prompted questions, are able to help students learn only the basic detailed information about the microlevel behavior of the molecules and not the interlevel causal relationships of how behaviors of the molecules at the microlevel collectively explain the behavior at the macrolevel pattern (Chi, Roscoe, Slotta, Roy & Chase, 2012). This is because this collective interlevel relationship is an emergent one, which is foreign to students who are used to sequential causal chain relationships. Thus, for domains or topics in which students cannot bootstrap themselves into deeper understanding due to a lack of relevant schemas, or for topics for which no deeper rationales exist, the predictions made by ICAP may not hold.

Task differences within a mode. The four engagement modes are based on a student's overt behavior. Typically the activities within each mode are distinct, so that pairwise comparisons of two modes usually compare two different activities (as was done in Chi, 2009) or the same activity implemented in two different ways (resulting in different modes), as shown in Tables 3, 4, and 5. In our analyses of the studies in Tables 3 to 5, we have shown that the same task, implemented with variations that do not change its mode, produces equivalent learning outcomes.

Although it is often the case that when two activities fall within a single cell (i.e., use the same mode of engagement), nonsignificant learning differences are obtained; however, ICAP cannot always make accurate predictions about learning outcomes from undertaking two different engagement activities that can be classified into the same mode. This is because a number of factors can affect the complexity of different tasks or activities within a mode (e.g., the goal or knowledge to-be-learned of a task, the time it takes to carry out the task, the procedure or process by which a task is carried out, etc.). We illustrate next with an example.

The task of interacting with a peer to invent and the task of comparing and contrasting two cases while working in pairs are both *interactive* activities because they require students to produce new knowledge: either inventing a rule or method in the invent case or come up with similarities and differences in the compare-and-contrast case. These two conditions were manipulated by Chi, Dohmen, Shemwell, Chase, and Schwartz (2012) when they asked sixth-grade students to learn about horizontal projectile motion. In the invent condition, students were asked to work in pairs to invent a single method to figure out where to put the final destination as a result of a set of projectile trajectories that were shot at different speeds and from different heights. In the compare–contrast condition, students were asked to work in pairs to compare and contrast sets of the same projectile trajectories. Students in the invent condition showed greater learning gains compared to students in the compare–contrast condition, even though both conditions are *interactive*.

A finer grained analysis of the task, however, suggests why the two *interactive* activities result in different learning gains. An analysis of the worksheets shows that the two engagement activities cause different ways of processing the cases. In the compare–contrast condition, students paid attention to single factor or feature (e.g., speed, height, etc.). This is revealed in comments such as, *They go in different speeds* or *They start at different heights.* As coded by the authors, a double feature comment would involve two features, such as *more speed = further distance*, and a triple statement would relate three features. The worksheet analyses showed that only 10% of the compare–contrast group produced triple-feature statements, whereas 100% of the invent group did. Conversely, the compare–contrast group produced on average 3.0 single-feature statements, whereas the invention groups produced no single-feature statements. In short, the activity of making comparisons across two cases biased the students to compare features individually, whereas the invent activity encouraged students to make more holistic statements and connect features with each other, which is necessary in order to invent the correct method.

This study demonstrates that although both learning activities are *interactive*, their differing task structures biased the way students performed them, resulting in different qualitative ways of learning, which leads to different amounts of learning. Notice, however, that the ICAP hypothesis focuses its predictions across modalities, which may not be violated in this case, in that both inventing and comparing and contrasting (both *interactive*) should produce greater learning outcomes than an *active* or *passive* activity.

In summary, it is possible that different levels of learning outcomes are achieved for activities within the same mode of engagement. One explanation is that some activities are more cognitively demanding than other activities, even if they both belong to the same mode. This is because overt behavior is a gross analysis of engagement behavior and not a cognitive analysis of a given task activity. However, even so, the prediction may still hold across modes of activities.

ICAP Can Further Serve as a Tool to Explain, Dictate, Evaluate, and Design

In this section, we describe the contribution of the ICAP framework and hypothesis as a tool for both research purposes and instructional design.

Using ICAP to explain discrepant findings about learning activities. In a preceding section of this article, we stated that one way of validating the ICAP hypothesis was to examine studies using three specific learning activities (note taking, concept mapping, and self-explaining) when each activity was implemented according to different modes of engagement. Under its lens, we can provide an interpretation for seemingly discrepant findings. By discrepant findings, we mean that for a given learning activity, such as note taking, some research suggests that it is helpful for learning, whereas others report that it is not. To illustrate, consider the following two studies. In one study (Mastropieri et al., 2001), the results showed that note taking by answering who–what questions was beneficial, whereas Coleman et al. (1997) concluded that note taking by summarizing was not as beneficial. Using the ICAP framework, we can resolve these discrepant findings by first categorizing the learning activity and then comparing it to the alternative condition used in each study for comparison. In the Mastropieri et al. publication, answering who–what questions is a *constructive* way of summarizing, and they compared that condition to a *passive* task of reading out loud. Thus, it is not surprising that learning is better in the *constructive* summarizing case (Mastropieri et al., 2001). On the other hand, Coleman et al. compared summarizing by selecting only the most important sentences, an *active* task, to summarizing by explaining, which is a *constructive* task. Therefore, it is not surprising that summarizing through selection (*active*) is worse for learning than summarizing through explanation (*constructive*; Coleman et al., 1997). Thus, an activity such as summarizing cannot be claimed as beneficial in an absolute way; its utility for learning depends on how it is implemented and to what appears activity it is compared. In short, what appears to be discrepant results in the literature is completely systematic from ICAP's point of view.

As the preceding sections show, in scrutinizing such studies, it is apparent that whether note taking is an effective learning activity depends on three factors: (a) how it is implemented, that is, what are the instructions/directions given to the student, thus making it an *active, constructive*, or *interactive* activity; (b) what the alternative (or often the control) condition it is contrasted with (i.e., if it was implemented as a *constructive* activity, then was it compared with another *passive, active,* or *constructive* activity); and (c) how students actually carried out the activity. In short, ICAP can highlight the systematicity that underlies apparent contradictions between studies.

Using ICAP to dictate the choice of a control condition in research design. Besides explaining discrepant results in the literature, ICAP can also be used as a tool for determining areas that are understudied and, most importantly, for deciding control conditions. As Tables 3 to 5 show, the majority of the studies on interventions compared a *constructive* mode with a *passive* mode. Far fewer studies have compared the *constructive* mode with *active*, *interactive*, or other *constructive* modes.

With respect to choosing meaningful control conditions, the most stringent comparison of an intervention is to compare it with another intervention that employs an engagement activity of the same mode. For example, to claim that inserting a metacognitive prompt is a useful intervention (Beal & Stevens, 2010), one needs to compare it with another type of prompt that also elicits generative responses so that both activities are *constructive*, rather than comparing it with reading academic advice that does not elicit a response. Only by comparing two interventions that fit in the same mode of engagement can one draw legitimate conclusions about the effectiveness of a specific type of intervention.

That is, suppose one finds that Intervention A is better than Intervention B. Before claiming that it is the specific processes behind Intervention A that caused the learning gains, it is important to examine the engagement mode of the two interventions. For example, if Intervention A is *constructive* and Intervention B is *active* than the learning benefits may not be due to the specific nature of Intervention A but may only lead to a more general conclusion that Intervention A is more cognitively engaging than Intervention B.

Using ICAP to evaluate students' outputs. So far, we have discussed only the mode of an activity, as intended by the designer of the activity, such as a teacher. For example, if a teacher asks students to draw concept maps, then presumably this is a *constructive* task. However, in many cases, this output (the concept maps) needs to be evaluated to determine whether in fact the students treated it as a *constructive* task. This is the distinction mentioned earlier between intended versus enacted. For example, suppose in such an activity, the teacher populated or provided a set of six concepts as starting concepts to be used as "nodes" in the concept map. Upon evaluating or coding of the students' maps, it becomes apparent that some students created concept maps that only contained the six concepts that were provided, whereas other students created maps that had many other concept nodes. This suggests that the students who used only the six provided nodes did not create any new outputs beyond what was provided (thus they were only *active*), whereas the other students who produced additional concept nodes were more *constructive.* Thus, the ICAP framework can be used to design a rubric to code and evaluate students' outputs. And often an accurate classification of an activity's mode can only be done after such evaluations.

Using ICAP as a guide for instructional design. In addition to the research contributions, the ICAP framework has strong practical implications as teachers and other instructional designers can use it to choose, modify, or design tasks for students to perform. Thus, even though engagement behaviors are strictly defined from the learners' perspective, clearly students can be encouraged to engage in certain ways through the careful design of learning activities.

In the current depictions in the literature, even though "active learning" (along with the related concept "constructive learning"; Cobb, 1994; Jonassen, 1991; Pelech, 2010; Wilson, 1997) may be ambiguous, "active learning" does have instructional implications. It suggests that teachers can "encourage the learner to engage in making sense of the material" (Mayer, 2008, p. 17). But are there concrete methods recommended for teachers to adopt and adapt? King (1993) explicitly offered 12 *getting involved* and *really thinking about it* activities for "active learning" (e.g., posing problems, thinking analogically, developing critiques, etc.). However, these 12 examples do not specify what qualifies exactly as an "active learning" activity. And if a teacher chooses not to use one of the 12 examples, how would she know what alternative activity constitutes an "active learning" activity? Moreover, how do teachers decide which activity is relatively better for fostering student learning? For example, is think-pair-share better than developing critiques? Finally, research provides few guidelines for teachers to tell them how to modify their current homework and seatwork assignments to encourage students to learn more effectively by being active.

Our framework, on the other hand, can provide specific guidelines for how to create lessons that incorporate overt behaviors that are associated with higher levels of engagement and their associated knowledge-change processes. We have evaluated whether the ICAP framework and ICAP hypothesis are beneficial for professional development by conducting several small-scale training workshops in which we introduced teachers to the framework and evaluated whether they could successfully use the framework when designing lessons. Initial results are promising and show that teachers are able to both understand the framework and correctly apply it when designing new lessons.

We have also developed an online module containing information about ICAP. Specifically, the module can scaffold teachers during the design process and encourage them to consider not only the specific task but how to design better assessments that measure deep knowledge as well. For example, a teacher who normally delivers instruction via a *passive* mode such as lecturing will learn from the module to provide students with a guided notes worksheet thus creating an *active* activity, or the teacher may learn to create a *constructive* activity by prompting students to build a concept map during the lecture. Similarly, a teacher who previously considered an *interactive* activity to be any task a student did not complete alone could learn from the module to ensure that *interactive* activities not only involve working in groups but also require students to be *constructive* or generative while completing the tasks.

With the explosion of computer-based instruction, this same set of challenges of how to design student activities also confronts designers of computer-based learning environments. Because some forms of activities (such as selecting an answer from a menu of choices) are easier to implement in a computer-based learning environment than others (such as generating a free-form response that needs to be verified), designers of learning environments also need clear-cut recommendations about the trade-offs between the cost of implementation and effectiveness of students' activities for learning. In this article, even though we focus primarily on teachers' practices, our work has direct implications for designers of all learning environments.

Comparison With Other Theories

To summarize, the ICAP framework consists of four modes of overt engagement activities, *passive, active, constructive,* and *interactive*, the associated knowledge-change processes for each mode, the associated resulting changes in knowledge, with the concomitant improvement in learning as a function of the knowledge changes and cognitive outcomes. In other words, the ICAP framework lays out a possible causal chain, leading from engagement activities to the knowledge-change processes underlying the engagement activities, resulting in changes in knowledge itself and the use of that changed knowledge. We further clarified the ICAP hypothesis and described the circumstances under which it can be violated. In addition, we showed how it can explain discrepant findings, help determine appropriate control conditions, and improve teachers' design of learning activities.

In general, ICAP is a theory of cognitive engagement with a behavioral metric. However, two caveats must be stated. First, the knowledge-change processes associated with each mode of engagement are hypothetical. We have not carried out studies to verify that these knowledge-change processes are in fact taking place when students engage in one mode over another. The second caveat is that there may be other theoretical interpretations of ICAP, viewing cognition that relies much less on representations and memory. However, we cannot derive an ICAP hypothesis from such an alternative view, other than the behavioral view, based on the hierarchical nature of the overt activities of each mode. Thus, we welcome any theoretical lens that can derive the ICAP hypothesis. In this final section, we highlight how ICAP differs from existing learning theories (constructivism, cognitive load theory, and Bloom's revised taxonomy), in terms of its theoretical stance and interpretation.

Constructivism. Constructivism is a broad framework that assumes that the responsibility of learning should reside increasingly with the learner (Von Glasersfeld, 1989). Moreover, constructivism emphasizes the importance of the learner being actively involved in the learning process, unlike previous educational viewpoints where the responsibility rested with the instructor to teach and where the learner played a passive, receptive role. Thus ICAP is very similar to constructivism in this sense of focusing more on the actions of the learner than the instructor. Constructivism translates to instruction by encouraging a variety of learner-controlled or learner-centered activities such as discovery, hands-on, experiential, collaborative, project-based, and task-based learning. These activities have analogs in the *active, constructive,* and *interactive* modes of learning. So the basic difference between constructivism and the ICAP framework is that ICAP differentiates in a more concrete and fine-grained way (both in terms of the behaviors and in terms of the learning outcomes) the activities that constructivism has promoted. Furthermore, perhaps constructivism is misinterpreted, but the notion of "constructing an understanding," let's say of the sentence "The heart has a double loop in circulation," is interpreted in ICAP as the way (if doing it *constructively*) to integrate that sentence with one's prior knowledge, making additional inferences, and so on, in order to understand what having a double loop in circulation might mean. *Constructing* in ICAP does not mean that the learner discovers the fact that "The heart has a double loop in circulation." *Constructing* in ICAP is a means to achieve or learn with understanding.

Cognitive load theory. In cognitive psychology, cognitive load refers to the load imposed on working memory from an information-processing framework. Because humans have a limited working memory capacity, the amount of information that can be processed concurrently during complex learning activities can overload the finite amount of working memory capacity. Cognitive load theory provides explicit empirically based guidelines for the best ways to design instructional materials, with the goal of decreasing extraneous cognitive load during learning. For example, cognitive load theory suggests integrating diagrams and text information rather than presenting them separately (i.e., split attention effect). In general, cognitive load theory suggests that instruction should be designed to focus the learner's attention toward the germane materials (Sweller, 1994).

The obvious difference between cognitive load theory and the ICAP framework is that ICAP is concerned with eliciting higher modes of activities from learners, instead of changing and reducing the load imposed by an activity. So the difference is one of focusing on the learners' activities versus the load of instructional activities. Moreover, ICAP's predictions seem to be the opposite of predictions from the load theory. For example, cognitive load theory states that the greater the load in the to-be-processed presented materials, the more difficult a task becomes, and therefore the less resulting learning. However, ICAP makes the opposite prediction. That is, as the mode of engagement goes from *passive* to *interactive,* the activity becomes more effortful (i.e., imposes more load). Although we can infer that this is true by the hypothetical knowledge-change processes involved, direct evidence can be gathered from asking students in a self-report. McEldoon (2014) asked students to rate their perceived cognitive load (or working memory demand) after completing a worked example that included instructional explanations of calculating an analysis of variance. One condition completed it *actively* by determining final values within the worked example and copying or paraphrasing instructional explanations. The other condition completed the worked example *constructively* by generating intermediate values and generating self-explanations prior to receiving instructional explanations. She found that students in the *constructive* condition reported significantly higher levels of cognitive load or working memory demand than those in the *active* condition. For one of the lessons, the *constructive* condition also had significantly higher learning gains. Thus, with respect to cognitive engagement, ICAP may make the opposite prediction from cognitive load theory, in that the more load or effort is used to process learning materials, the more learning is achieved. However, at times ICAP and cognitive load theory predict congruent results; namely, if increased engagement corresponds to increased germane cognitive load (e.g., when students are asked to generate self-explanations, they are being *constructive*, according to ICAP, and by explaining they are also increasing germane load).

Bloom's taxonomy. It is a daunting task to compare and contrast Bloom's taxonomy with ICAP. We are basing our comparison with Bloom's taxonomy as described in the 2001 revised version (Anderson & Krathwohl, 2001). The goal of Bloom's taxonomy is to first classify learning objectives and then to design instructional activities and assessment that align with the objectives. A teacher's learning objectives must first be classified into one of six categories of cognitive processes: *remember, understand, apply, analyze, evaluate,* and *create*. Suppose an instructional objective is to "differentiate between rational numbers and irrational numbers." "Differentiate" or "distinguish between" are assumed to require the cognitive processes of *analyze,* according to Anderson and Krathwohl (2001), because *analyze* "involves breaking material into its constituent parts and determining how the parts are related to one another and to an overall structure" (p. 79). Once teachers have classified the learning objective as *analyze*, they must design instruction that requires students to *analyze*, such as use examples and non-examples to help students form the proper categories of rational and irrational numbers. In addition, assessment items must be designed to align with the objective of *analyze*, such as designing a test question that asks students to classify each number on a list of real numbers as either a rational or an irrational number.

The major characteristic difference between Bloom's taxonomy and the ICAP taxonomy is that Bloom's taxonomy focuses its users on their instructional goals and how to measure whether the goal has been achieved, whereas ICAP focuses its users on the means for achieving the instructional goals. Because one framework focuses on *ends* and the other on *means*, the two frameworks are complementary. Minor differences also exist. First, ICAP is more parsimonious in that the framework applies to learning of various forms of knowledge, whether it is factual, conceptual, or procedural; and the categories of learner engagement activities can be easily distinguished by comparing information generated by the learner(s) vis-à-vis the information provided by the learning environment. Second, Bloom's cognitive processes refer to the processes involved in carrying out the assessment task, such as the task of *analyzing* that requires decomposing into constituent units, similar to various task-specific cognitive processes underlying other problem-solving tasks and activities mentioned in this article. ICAP's cognitive processes, however, refer to the processes of learning or what is referred to here as knowledge-change processes. Third, ICAP proposes that co-construction through interaction can achieve the greatest learning outcomes, whereas Bloom's taxonomy does not distinguish between intrapersonal and interpersonal cognition and, hence, his learning outcomes can be achieved at all levels without interaction with others.

CONCLUSION

The ICAP framework and hypothesis provide specific, operationalized definitions of engagement activities that can easily be applied to a number of learning environments. The ICAP hypothesis predicts that as activities move from *passive* to *active* to *constructive* to *interactive*, students undergo different knowledge-change processes and, as a result, learning will increase. This hypothesis has been validated through a number of classroom and laboratory studies, and in this article, we examined its validity in three specific tasks: note taking, self-explaining, and creating concept maps. In addition, we use the framework to clarify discrepant findings in the literature, propose a number of useful applications of the framework, and compare and contrast ICAP to existing cognitive and learning theories. Essentially, ICAP is a hypothesis about the relative level of learning associated with each of these four modes of student engagement, with the advantage being that it can detect regularity across a large corpus of data in the literature and can be used to inform both classroom and laboratory studies.

ACKNOWLEDGMENTS

Comments and edits from the reviewers and Clark Chinn are greatly appreciated. Special thanks also go to Bryan Henderson, Katherine McEldoon, Glenda Stump, and Kurt VanLehn for their comments.

FUNDING

The authors are grateful for support from the Institute of Education Sciences (Award #R305A110090) for the project Developing Guidelines for Optimizing Levels of Students' Overt Engagement Activities.

REFERENCES

Aleven, V., & Koedinger, K. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*, 147–179.

Alibali, M., & DiRusso, A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*, 37–56.

Anderson, L. W., & Krathwohl, D. R. (2005). *A taxonomy for learning, teaching, and assessing*. London, England: Longman.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*, 774–783.

Azevedo, R., Greene, J. A., Moos, D. C., Winters, F. I., Cromley, J. G., & Godbole-Chaudhuri, P. (2006). Is externally-regulated learning by a human tutor always effective in facilitating learning with hypermedia? In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the 7th International Conference on Learning Sciences* (pp. 16–22). Mahwah, NJ: Erlbaum.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*, 307–359.

Bartlett, F. C. (1958). *Thinking: An experimental and social study*. London, England: Allen & Unwin.

Bauer, A., & Koedinger, K. R. (2007). Selection-based note-taking applications. *CHI 07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY: ACM.

Beal, C. R., & Stevens, R. H. (2010). Improving students' problem solving in a web-based chemistry simulation through embedded metacognitive messages. *Technology, Instruction, Cognition & Learning*, *8*, 255–271.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, *13*, 221–252.

Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching a new agent paradigm for education software. *Applied Artificial Intelligence*, *19*, 363–392.

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2004). Motivation and cognitive engagement in learning environments. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 475–488). New York, NY: Cambridge University Press.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualisations. *Learning and Instruction*, *14*, 325–341.

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom*. Washington, DC: School of Education and Human Development, George Washington University.

Bonwell, C. C., & Sutherland, T. E. (1996). The active learning continuum: Choosing activities to engage students in the classroom. *New Directions for Teaching and Learning*, *1996*(67), 3–16.

Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393–451). Hillsdale, NJ: Erlbaum.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182–197.

Chang, K.-E., Sung, Y.-T., & Chen, I.-D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education*, *71*, 5–23.

Chi, M., Dohmen, I., Shemwell, J. T., Chase, C. C., & Schwartz, D. L. (2012, April). *Seeing the forest from the trees: A comparison of two instructional models using contrasting cases*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, BC, Canada.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in Instructional Psychology*, *5*, 161–238.

Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self–explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H. & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Chi, M. T. H., & Kang, S. (2014). *The learning benefits of dialogue videos over monologue videos: Implications for designing online videos*. Manuscript submitted for publication.

Chi, M. T. H., & Menekse, M. (in press). Dialogues as interactions. In C. Asterhan & L. Resnick (Eds.), *Socializing intelligence through academic talk and dialogue*. Washington, DC: AERA.

Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*, 1–61.

Chi, M. T. H., Roy, M., & Hausmann, R. G. M. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, *32*, 301–341.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*, 471–533.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, *86*, 175–218.

Chularut, P., & DeBacker, T. K. (2004). The influence of concept mapping on achievement, self-regulation, and selfefficacy in students of English as a second language. *Contemporary Educational Psychology*, *29*, 248–263.

Cobb, P. (Ed.). (1994). *Learning mathematics: Constructivist and interactionist theories of mathematical development*. Dordrecht, The Netherlands: Kluwer Academic.

Coffey, J. W., Carnot, M. J., Feltovich, P. J., Feltovich, J., Hoffman, R. R., & Cañas, A. J. (2003). *A summary of literature pertaining to the use of concept mapping techniques and technologies for education and performance support* (Tech. rep. submitted to the US Navy Chief of Naval Education and Training). Pensacola, FL: Institute for Human and Machine Cognition.

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences*, *6*, 347–365.

Conati, C., & Carenini, G. (2001). Knowledge encapsulation: Studies on the development of medical expertise. In *17th International Joint Conference on artificial intelligence* (pp. 1301–1306). San Francisco, CA: Morgan Kaufmann.

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education*, *11*, 389–415.

Corno, L., & Mandinach, E. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, 18, 88–108.

Czerniak, C. M., & Haney, J. J. (1998). The effect of collaborative concept mapping on elementary preservice teachers' anxiety, efficacy, and achievement in physical science. *Journal of Science Teacher Education*, *9*, 303–320.

Damon, W. (1984). Peer education: The untapped potential. *Journal of Applied Developmental Psychology*, *5*, 331–343.

Dawson, I. (2004). Time for chronology? Ideas for developing chronological understanding. *Teaching History*, *117*, 14–24.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, *332*, 862–864.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies*, *70*, 377–398.

Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2010, October). A conceptual approach to physics problem solving. In *AIP Conference Proceedings* (Vol. 1289, p. 137). Portland, OR: American Institute of Physics.

Doymus, K. (2008). Teaching chemical equilibrium with the jigsaw technique. *Research in Science Education*, *38*, 249–260.

Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures: Teaching for active learning. *Bioscience*, *47*, 601–607.

Ferguson-Hessler, M., & de Jong, T. (1990). Studying physics texts: Differences in study processes between good and poor performers. *Cognition and Instruction*, *7*, 41–54.

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. New York, NY: Wiley & Sons.

Fonseca, B. & Chi, M. T. H. (2011). The self-explanation effect: A constructive learning activity. In R. E. Mayer & P. A. Alexander (Eds.), *The handbook of research on learning and instruction* (pp. 296–321). New York, NY: Routledge/Taylor and Francis.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59–109.

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, *36*, 39–53.

Goldstone, R. L., & Day, S. (Eds.) (2012). New conceptualizations of transfer of learning [Special issue]. *Educational Psychologist*, *47*(3).

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*, 104–137.

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*, 93–103.

Guastello, E. F., Beasley, T. M., & Sinatra, R. C. (2000). Concept mapping effects on science content comprehension of low-achieving inner-city seventh graders. *Remedial and Special Education*, *21*, 356–364.

Hausmann, R. G. M., & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology*, *7*, 4–14.

Hausmann, R. G. M., van de Sande, B., & VanLehn, K. (2008). Shall we explain? Augmenting learning from intelligent tutoring systems and peer collaboration. In B. P. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems* (pp. 636–645). Amsterdam, The Netherlands: IOS.

Hendricks, C. C. (2001). Teaching causal reasoning through cognitive apprenticeship: What are results from situated learning? *Journal of Educational Research*, *94*, 302–311.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction*, *17*, 379–432.

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, *77*, 95–111.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers*, *34*, 383–390.

Jonassen, D. H. (1991). Evaluating constructivistic learning. *Educational Technology*, 31, 28–33.

Kafai, Y. B. (2006). Playing and making games. *Games and Culture*, *1*, 36–40.

Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., . . ., Canny, J. (2005, April). Livenotes: A system for cooperative and augmented note-taking in lectures. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI '05)* (pp. 531–540). Portland, OR.

Kastens, K. A., & Liben, L. S. (2007). Eliciting self-explanations improves children's performance on a field-based map skills task. *Cognition and Instruction*, *25*, 45–74.

Katayama, A. D., Shambaugh, R. N., & Doctor, T. (2005). Promoting knowledge transfer with electronic note taking. *Teaching of Psychology*, *32*, 129–131.

King, A. (1993). From sage on the stage to guide on the side. *College Teaching*, *41*, 30–35.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Kitchner, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development*, *26*, 222–232.

Kobayashi, K. (2006). Combined effects of note-taking/reviewing on learning and the enhancement through interventions: A meta-analytic review. *Educational Psychology*, *26*, 459–477.

Kramarski, B., & Dudai, V. (2009). Group-metacognitive support for online inquiry in mathematics with differential self-questioning. *Journal of Educational Computing Research*, *40*, 377–404.

Loyens, S. M. M., & Gijbels, D. (2008). Understanding the effects of constructivist learning environments: Introducing a multi-directional approach. *Instructional Science*, *36*, 351–357.

Mastropieri, M. A., Scruggs, T., Mohler, L., Beranek, M., Spencer, V., Boon, R. T., & Talbott, E. (2001). Can middle school students with serious reading difficulties help each other and learn anything? *Learning Disabilities Research & Practice*, *16*, 18–27.

Mastropieri, M. A., Scruggs, T. E., Spencer, V., & Fontana, J. (2003). Promoting success in high school world history: Peer tutoring versus guided notes. *Learning Disabilities Research & Practice*, *18*, 52–65.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, *59*, 14–19.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

McEldoon, K. L. (2014). *Supporting novice learning through focused processing of worked examples and explanations*. (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.

McEldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional-practice. *British Journal of Educational Psychology*, *83*, 615–632.

Menekse, M., Stump, G., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, *102*, 346–374.

Mestre, J. P. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Journal of Applied Developmental Psychology*, *23*, 9–50.

Meyers, C., & Jones, T. B. (1993). *Promoting Active Learning. Strategies for the College Classroom*. San Francisco, CA: Jossey-Bass.

Milrad, M. (2002). Using construction kits, modeling tools and system dynamics simulations to support collaborative discover learning. *Educational Technology and Society*, *5*, 76–87.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A metaanalysis. *Review of Educational Research*, *76*, 413–448.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, *27*, 937–949.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York, NY: Cambridge University Press.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, *18*, 443–468.

O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review*, *14*, 71–86.

Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science*, *21*, 109–146.

Okebukola, P. A., & Jegede, O. J. (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education*, *72*, 489–500.

O'Reilly, T., Symons, S., & MacLatchy-Gaudet, H. (1998). Brief research report: A comparison of self-explanation and elaborative interrogation. *Contemporary Educational Psychology*, *23*, 434–445.

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York, NY: Basic Books.

Pea, R., & Kurland, M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology*, *2*, 137–168.

Pelech, J. (2010). *The comprehensive handbook of constructivist teaching: From theory to practice*. Charlotte, NC: Information Age.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, *78*, 34–38.

Piaget, J. (1930). *The child's conception of physical causality*. Oxford, England: Harcourt Brace.

Pillow, B. H., Mash, C., Aloian, S., & Hill, V. (2002). Facilitating children's understanding of misinterpretation: Explanatory efforts and improvements in perspective taking. *The Journal of Genetic Psychology*, *163*, 133–148.

Pine, K. J., & Messer, D. J. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction*, *18*, 35–52.

Pintrich, P. R., & De Groot, E. (1990). Motivation and self regulation components of academic performance. *Journal of Educational Psychology*, *82*, 33–40.

Rafal, C. T. (1996). From co-construction to takeovers: Science talk in a group of four girls. *Journal of Learning Sciences*, *5*, 279–293.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, *23*, 90–108.

Rickards, J. P., & Friedman, F. (1978). The encoding versus the external storage hypothesis in note taking. *Contemporary Educational Psychology*, *3*, 136–143.

Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*, 534–574.

Rosenthal, J. S. (1995). Active learning strategies in advanced mathematics classes. *Studies in Higher Education*, *20*, 223–228.

Rouet, J.-F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. *Journal of Educational Psychology*, *88*, 478–493.

Rowles, C. J. (2013). Strategies to promote critical thinking and active learning. *Teaching in nursing: A guide for faculty*, 258–279.

Rubin, J., & Herbert, M. (1998). Peer teaching-model for active learning. *College Teaching*, *48*, 26–30.

Russell, I. J., Caris, T. N., Harris, G. D., & Hendricson, W. D. (1983). Effects of three types of lecture notes on medical student achievement. *Academic Medicine*, *58*, 627–636.

Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. *Review of Research in Education*, *23*, 1–24.

Scardamalia, M. (1992, July). *The role of adaptation and understanding in knowledge-building communities*. Paper presented at conference of NATO Advanced Science Institute, Kolymbari Crete, Greece.

Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Journal of the learning students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences*, *4*, 131–166.

Schmid, R. F. & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *Journal of Educational Research*, *84*, 78–85.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling a time for telling. *Cognition and Instruction*, *16*, 475–522.

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right . . . if they argue together! *Cognition and Instruction*, *18*, 461–494.

Siegler, R. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, *28*, 225–273.

Stark, R. (1999). Lernen mit Lösungsbeisplielen: Einfluss unvollständiger Lösungsbeispiele auf Beispielelaboration, Motivation and Lernerfolg [Learning by worked-out examples: The impact of completion tasks on example elaboration, motivation and learning outcomes]. Bern, Switzerland: Huber.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences*, *12*, 183–218.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312.

Trafton, J. G., & Trickett, S. B. (2001). Note-taking for self-explanation and problem solving. *Human–Computer Interaction*, *16*, 1–38.

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction*, *10*, 311–330.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*, 3–62.

VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, *2*, 1–59.

Von Glasersfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, *80*, 121–140.

Webb, N. M. (1989). Peer interaction, problem solving, and cognition: Multidisciplinary perspectives. *International Journal of Educational Research*, *13*, 1–119.

Whitehead, A. N. (1929). *The aims of education*. New York, NY: Macmillan.

Whitten, W. B., II. (2011). Learning from and for tests. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 217–234). New York, NY: Psychology Press.

Willerman, M., & Mac Harg, R. A. (1991). The concept map as an advance organizer. *Journal of Research in Science Teaching*, *28*, 705–711.

Wilson, B. G. (1997, March). *Reflections on constructivism and instructional design, instructional development paradigms*. Englewood Cliffs, NJ: Educational Technology.

Wong, R. M. F., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction*, *12*, 233–262.

Wylie, R. (2011). *Examining the generality of self-explanation* (Unpublished doctoral dissertation). Carnegie-Mellon University, Pittsburgh, PA.

Yaron, D., Karabinos, M., Lange, D., Greeno, J. G., & Leinhardt, G. (2010). The ChemCollective—virtual labs for introductory chemistry courses. *Science*, *328*, 584–585.

Yetter, G., Gutkin, T., Saunders, A., Galloway, A., Sobansky, R., & Song, S. (2006). Individual practice for complex problem solving: A cautionary tale. *Journal of Experimental Education*, *74*, 137–159.

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*, *42*, 166–184.

Zhang, Z. H., & Linn, M. C. (2013). Learning from chemical visualizations: Comparing generation and selection. *International Journal of Science Education*, *35*, 2174–2197.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *21*, 3–17.