HUMAN–COMPUTER INTERACTION, 2011, Volume 26, pp. 123–159

ISSN: 0737-0024 print / 1532-7051 online
DOI: 10.1080/07370024.2011.556552

# Self-Directed Learning and the Sensemaking Paradox

**Kirsten R. Butcher**[1] **and Tamara Sumner**[2]
[1]*University of Utah*
[2]*University of Colorado*

Educative sensemaking focuses on the needs of self-directed learners, a nonexpert population of thinkers who must locate relevant information sources, evaluate the applicability and accuracy of digital resources for learning, and determine how and when to use these resources to complete educational tasks. Self-directed learners face a *sensemaking paradox*: They must employ deep-level thinking skills to process information sources meaningfully, but they often lack the requisite domain knowledge needed to deeply analyze information sources and to successfully integrate incoming information with their own existing knowledge. In this article, we focus on the needs of college-aged students engaged in learning about natural sciences using web-based learning resources. We explored the impact of cognitive personalization technologies on students' sensemaking processes using a controlled study in which students' cognitive and metacognitive processes were analyzed as they completed a common educational task: writing an essay. We coded students' observable on-screen behaviors, self-reported processes, final essays, and responses to domain assessments to assess benefits of personalization technologies on students' educative sensemaking. Results show that personalization supported students' analysis of knowledge representations, helped students work with their representations in meaningful ways, and supported effective encoding of new knowledge. We discuss implications for new technologies to help students overcome the educative sensemaking paradox.

---

**Kirsten Butcher** is interested in the impact of multimedia and interactive educational technologies on students' comprehension processes and learning outcomes; she is an Assistant Professor at the University of Utah in the Department of Educational Psychology's Learning Sciences/Instructional Design and Educational Technology programs. **Tamara Sumner** is interested in how cognitive tools, computational algorithms, and interactive media can improve learning outcomes and learner engagement; she is an Associate Professor at the University of Colorado at Boulder, in the Institute of Cognitive Science and the Department of Computer Science.

**CONTENTS**



## 1. INTRODUCTION

Sensemaking has been studied for people engaged in a wide variety of information-rich tasks, from typical tasks such as writing a monthly newsletter to complex, specialized tasks such as intelligence analyses. A common aim across these studies is to better understand the activities and cognitive processes that people engage in when analyzing and "encoding" collections of information sources to solve a problem or "answer task-specific questions" (Russell, Stefik, Pirolli, & Card, 1993). A significant portion of this prior work has focused on developing high-end visualizations for use by sensemaking experts, that is, people with considerable technical skills and significant knowledge of the problem domain (Russell, Jeffries, & Irani, 2008). However, as noted by Russell and colleagues, this focus may miss the unique sensemaking needs of a much larger population of nonexperts.

In this vein, we focus on the sensemaking needs of an important, nonexpert population: self-directed learners. Specifically, we are interested in secondary and tertiary students engaged in learning about natural sciences using web-based educational resources. For these sensemakers, typical questions to be answered might include explaining why mountains often occur at plate tectonic boundaries, or explaining what happens when two air fronts of different densities meet. For students, self-directed learning episodes increasingly are online experiences. Like many other knowledge workers, students have embraced the World Wide Web with vigor, with a majority of students relying upon web materials as their primary—and often sole—source of information for learning (e.g., Graham & Metaxas, 2003). Although students may be enthusiastic, numerous studies suggest that they often are not very effective in their use of the web to support their learning and research, even when students self-rate their technical skills very highly (Graham & Metaxas, 2003; Stone & Madigan, 2007). Many students experience difficulties in locating appropriate resources, evaluating the applicability and accuracy of resources, determining which portions of resources are relevant to the task at hand, and integrating multiple sources of information with their own developing knowledge (Quintana, Zhang, & Krajcik, 2005), that is, activities we closely associate with sensemaking.

Decades of cognitive science research has highlighted that educative sensemaking activities benefit from two types of knowledge that students often lack: domain knowledge (in this case, science content knowledge) and metacognitive skills (Lin & Zabrucky, 1998; National Research Council, 2000). Metacognitive skills influence one's ability to "learn how to learn" and include monitoring one's state of understanding, identifying knowledge gaps, determining when more information is needed, and using deep and meaningful strategies to accomplish educational goals (e.g., Schraw, 1998). For online inquiry tasks, effective metacognitive skills require effective strategies for finding online information and monitoring what was learned from various sources (Quintana et al., 2005). However, domain knowledge is essential for determining the relevance, applicability, accuracy, and sufficiency of information sources; effective learners are able to activate their prior knowledge and use this knowledge during learning (e.g., Azevedo, Moos, Greene, Winters, & Cromley, 2008; Moos & Azevedo, 2008; see Kintsch, 1998, for a discussion of prior knowledge activation). This is the *sensemaking paradox* faced by 21st-century students: We routinely expect them to deploy metacognitive skills to learn complex topics, yet students need significant domain knowledge to apply these skills effectively. Lynch (2008) distinguished between piecemeal learning opportunities and deeper educational experiences, emphasizing that there is a fundamental difference between being able to retrieve information online and becoming educated from such information. There is a growing need to consider not simply whether students can learn isolated facts from web resources but whether students can make sense of digital materials in ways that result in coherent bodies of knowledge (Lynch, 2008).

We argue that the educative sensemaking needs of self-directed learners can be supported by *cognitive personalization* tools. Cognitive personalization tools support student learning by matching students with sets of educational resources that they, as individual learners with a unique profile of prior knowledge and misunderstandings, need in order to develop a more complete and coherent understanding of the topic at hand. Because these educational resources are selected based on students' existing knowledge, cognitive personalization tools also help students to shift more easily between representations, namely, the educational resources and the representation being constructed by the learner (e.g., an essay, report, or presentation). This movement is facilitated by drawing connections between the content of the students' generated knowledge representation and the customized resources retrieved to target that knowledge. As such, cognitive personalization tools help learners to focus on high-value sensemaking activities, such as effectively integrating and applying science content knowledge drawn from multiple sources, while performing other sensemaking activities for them, such as collecting, organizing, and identifying relevant sources.

In this article, we explore the ways in which cognitive personalization technologies can support effective sensemaking with web-based educational resources. First, we describe several use cases to illustrate how learners in a variety of settings might interact with and benefit from cognitive personalization tools. We then discuss a theoretical model of educative sensemaking. Next, we describe a prototype cognitive personalization service, the Customized Learning Service for Concept Knowledge: CLICK. Its personalization capabilities are realized through a combination of natural language processing algorithms and graph analytic techniques. We use the term "service" as CLICK has been designed and implemented as a web service application programming interface, enabling cognitive personalization capabilities to be flexibly embedded in a rich variety of tools, portals, and learning environments. We then describe a learning environment implemented with CLICK and discuss empirical findings from a controlled, mixed-method study that explored its impact on learners' sensemaking processes. Finally, we discuss implications of our work and future challenges for promoting personalized sensemaking with digital educational resources.

## 2. USE CASES

CLICK is designed to support cognitive personalization for students engaged in a variety of scientific explanation tasks. Asking students to develop scientific explanations is a common and widely respected educational activity. Proponents of this pedagogically rich activity argue that "writing a scientific explanation encompasses the processes, strategies, skills and values for constructing a valid argument in the scientific domain" (de la Chica, p. 9, 2008). Numerous studies document a variety of educational benefits, ranging from helping students to develop argumentation skills (P. Bell & Linn, 2000) to writing as a means of promoting deeper cognitive engagement with science content (P. D. Klein, 2004). From an educator's perspective, scientific explanations serve as an effective formative assessment; explanations provide an opportunity for learners to make their thinking visible, thus allowing instruction to be guided by a more informed understanding of what learners may or may not know.

Scientific explanations constructed by students can take many forms, from the traditional essay or report to alternative representations such as slide presentations, concept maps, or scientific journal entries. Despite these variations in form, common across these activities is the essential process of students developing an explanation with a significant textual component that lends itself to automatic analysis using CLICK's natural language processing-based algorithms.

Let us consider, in more detail, the common educational task of writing an essay on an assigned topic. Imagine a student who is asked to write a 250-word essay explaining why mountains often occur at plate tectonic boundaries. The student must consider what she knows about the topic, decide if her knowledge is lacking any critical concepts, and decide if and when to search for new information. When using the Web to support this task, the student must generate appropriate search terms, identify and select relevant and appropriate resources, and then read or interact with the resources to develop her understanding and support essay writing. As she writes and revises her essay, the student must analyze the strengths and weaknesses of her essay, decide on the revision strategies that are needed, and make meaningful changes to her written explanation. Even for skilled learners, this is a daunting task.

Cognitive personalization tools can support learners in their sensemaking by identifying specific portions of their explanation that are potentially problematic and suggesting a small suite of resources that learners can use in revising the problems. For instance, a learner may have written that "when two plates meet, the bigger plate causes the smaller plate to scrunch up and form mountains." The personalization tool can suggest three digital resources, including multimedia animations, which examine what happens when plates of different *densities* (not sizes) meet. Alternatively, the learner may have written, correctly, that mountain building takes place at subduction zones where oceanic and continental plates meet, citing the Sierra Nevada Mountains in California as an example. In this case, the personalization tool could detect that the student's knowledge may be incomplete and suggest a small number of resources about plate boundaries where two *continental* plates meet, including additional examples from the Himalayas. In both examples, the personalization tool is not giving students the "right" answer. Instead, the tool is using potential problems in the student's own scientific explanation to focus and motivate sensemaking, helping the student to consult multiple relevant information resources, to integrate this new knowledge with prior understanding, and to immediately apply the knowledge in revising the scientific explanation.

This essay example illustrates one use case guiding our efforts: an adaptive essay writing environment. This use case and two others were collaboratively developed with six middle and high school science teachers during a 1-day Future Learning Environments Workshop (Ahmad, 2008). During the workshop, teachers worked in groups to develop written scenarios and accompanying storyboards outlining how students might interact with learning environments of the future. These scenarios and storyboards guided the design of both the CLICK web service and the demonstration learning environment described later in this article. The other two use cases developed by the teachers were as follows:

- An adaptive presentation-building environment, where students create interactive slideshows to communicate scientific explanations. In this use case, personalized feedback focuses on both science content and the organization/structure of the presentation. Suggestions could include recommendations for grouping related bullets and for resources containing visuals and animations of science concepts.
- An adaptive concept map builder, where students create concept maps depicting their scientific explanations. Students' key ideas are represented as text in concept map nodes and the relationships between ideas are depicted as links between nodes. In this use case, personalization could provide feedback and resource recommendations on the contents of individual nodes (statements of their key ideas) and on the linkages between their ideas.

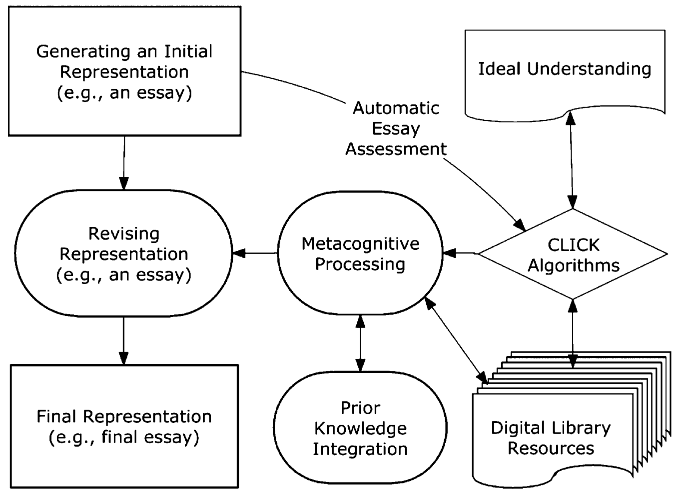
## 3. THEORETICAL PROCESSES OF EDUCATIVE SENSEMAKING

Educational tasks, such as writing essays, creating presentations, and generating concept maps, all involve finding and gathering relevant information, synthesizing information across sources that may differ widely in their content and format (e.g., textual introductions, visual representations, interactive examples), identifying key concepts and ideas, and integrating new knowledge into prior understanding and emerging work products. Although many different definitions of sensemaking have been proposed (e.g., G. Klein, Moon, & Hoffman, 2006a; Weick, Sutcliffe, & Obstfeld, 2005), it is clear that many educational tasks draw upon three major components of sensemaking models: seeking information, evaluating content, and using representations (Abraham, Petre, & Sharp, 2008).

### 3.1. A Model of Educative Sensemaking

For educative sensemaking, the paradox lies in that students are expected to self-direct the process of gathering and using resources, but they lack the prior domain knowledge and metacognitive skills that they need to be successful in doing so. Supporting successful educative sensemaking requires supporting the metacognitve processes that mediate between prior knowledge and learning materials in ways that help students gradually improve their knowledge representations. To use G. Klein and Moon's terminology (G. Klein, Moon, & Hoffman, 2006b), prior knowledge serves as a frame from which learners approach information sources and structure their responses to instructional feedback. Feedback should be more powerful when it targets learner *processes* rather than learner *outcomes*, but learners' abilities to use this feedback hinges on activating a frame (i.e., prior knowledge) that allows them to interpret and work with feedback in meaningful ways. For example, we would expect that it is more valuable to suggest that a student should elaborate on his explanation plate boundaries types—prompting the learner to analyze his current explanation,

**FIGURE 1. A model of educative sensemaking.**




identify additional relevant information, and integrate that information into his existing understanding—than to simply tell the student that he missed a plate boundary and guide him to the missing information.

Figure 1 depicts a model of educative sensemaking from the standpoint of developing a learning artifact (e.g., an essay). In this model, students progress from an initial to a final knowledge representation, using information sources and internal processes to revise their representation. In Figure 1, square nodes depict observable work products, partial squares depict information sources, and oval nodes depict learner processes. We have integrated our cognitive personalization tool, CLICK, into the model to demonstrate how educative sensemaking can be supported by this type of system. Basic components of the educative sensemaking model are discussed in more depth below, and CLICK algorithms are described in Section 4.

To demonstrate connections across sensemaking models, in the following sections we align the main components of educative sensemaking with the sensemaking loops proposed by Russell et al. (1993). Although Russell et al. developed this model by analyzing the processes of a highly knowledgeable team in the process of developing a training course on laser printing for professional technicians, the close connections of their identified loops to educational processes demonstrates the potential of sensemaking as a general theory of knowledge development.

## 3.2. Generating an Initial Representation

As seen in Figure 1, students performing an educative sensemaking task first need to generate an initial, concrete representation upon which sensemaking processes can operate. For ease of explanation, in this section we assume that learners have been assigned the task of writing an essay that explains a scientific concept. Generating the initial representation—a rough draft of the essay—requires students to *activate* their prior knowledge and *instantiate* their knowledge as a work object. That is, students first consider what they know about the topic about which they can write. Retrieving relevant knowledge corresponds to the first stage of Russell et al.'s (1993) model: a *search for representations*. In educative sensemaking, students are searching for and retrieving (from prior knowledge in long-term memory) the knowledge frames that will be used to structure further learning.

Once students retrieve their relevant knowledge, they can use this information to generate their initial representation. This initial representation is a concrete artifact in which information can be revised, added to, or selectively deleted. For an essay task, generating an initial representation involves creating an initial, rough draft of the essay that can be revised and refined during sensemaking. This aligns to the *instantiates representations* loop of Russell et al. (1993); in this case, learners are instantiating an initial representation that may be revised multiple times.

## 3.3. Revising a Representation

The second stage of educative sensemaking occurs as learners work with their instantiated representation. We define "revising a representation" as the observable behaviors that result in changes to an essay (see Figure 1). This process is iterative and corresponds to subsequent instances of the *instantiates representations* loop of Russell et al. (1993). For example, students might delete specific ideas, add new scientific content, or revise the style or wording of their essays. The actions that learners use to revise their essays are indicative of the forms of knowledge upon which they are drawing.

Contemporary theories of comprehension and learning can be used to characterize revisions to a representation as either more *deep* or more *shallow*. Comprehension theory and research has established that not all forms of knowledge are equally valuable to future learning and, hence, to sensemaking (see Kintsch, 1998, for a discussion). Construction-Integration (CI) theory (Kintsch, 1988; van Dijk & Kintsch, 1983) is a well-established cognitive theory that has demonstrated that knowledge can exist at three levels. The most basic level of representation is the surface level, which represents verbatim text or information drawn from learning materials. A learner forms a surface level representation by memorizing or borrowing exact information from a resource. The second level of representation is the textbase, which represents the content of an information resource at a propositional level. Like the surface level, the textbase representation contains basic ideas from a resource but, unlike the surface level, it does not preserve the verbatim form of the original materials. When learners paraphrase resources, they are expressing a textbase level of knowledge representation: It fails to go beyond the information contained in learning materials. The surface and textbase levels of knowledge representation characterize more shallow types of knowledge that tend to fade rather quickly and do not transfer to new situations (Kintsch, 1994, 1998).

According to the CI model, the third level of knowledge representation is the situation model, which is formed when a learner integrates information: This can

**FIGURE 2. Deep and shallow processes during essay revision.**

| Process Type | Description |
|---|---|
| Shallow Revisions | • Copy information from digital resource(s)<br>• Paraphrase information from digital resource(s)<br>• Delete a problematic statement/sentence from essay |
| Deep Revisions | • Integrate information from multiple sentences in a single digital resource and integrate into essay<br>• Integrate information from multiple sentences in 2+ digital resources and integrate into essay<br>• Construct correct inferences that are drawn from digital resources and integrated into essay<br>• Generate new information in the essay |

occur when a learner integrates multiple sources of new information (e.g., integration inferences; cf. Butcher, 2006) or when students integrate new information with prior knowledge (Kintsch, 1986, 1994). The result of this integration is a new, more flexible knowledge representation that enables (and is characterized by) logical inferences and application of knowledge to new problems or scenarios.

The CI model has been used not only to characterize the ways in which students learn from materials but also to characterize revisions to a text that support deeper learning. For example, adding content information that is more complete, better explained, and better connected to surrounding materials has been shown to support deeper learning (Britton & Gulgoz, 1991; Liederholm et al., 2000). Wiley and Voss (1999) have drawn upon this theory to categorize the use of resources during writing, noting whether information was "borrowed," "added," or "transformed." We draw upon Wiley and Voss (Wiley & Voss, 1996, 1999) and the CI model to characterize *shallow* and *deep* forms of revision (see Figure 2). *Shallow* revisions are the observable behaviors that delete ideas from a representation or borrow from existing resources without transforming or integrating the information. *Deep* revisions occur when revision behaviors change or transform information in ways that create new knowledge. Typically, this occurs when learners integrate multiple sources of information or generate inferences.

## 3.4. Metacognitive Processes

The revision of a representation is driven by internal learner processes that operate across prior knowledge and learning materials. During self-directed learning in an online environment, metacognitive skills encompass at least three major activities (Azevedo, Guthrie, & Seibert, 2004; Quintana et al., 2005; Schraw, 1998):

- Analyzing the strengths and weaknesses of the existing representation
- Seeking new or supporting information and materials
- Making use of deep, knowledge-based strategies when revising the representation

In educative sensemaking, metacognitive processes serve to drive *representational shifts* (Russell et al., 1993), whereby the sensemaker identifies that his or her instantiated representation is a poor fit to the relevant data and seeks to make necessary adjustments. Sensemakers may find that their representation is missing information, that it reflects inaccurate information, or simply that it is difficult to make meaningful connections between the representation and relevant data sources. However, research has shown that few students are able to successfully deploy effective metacognitive processes on their own, especially when engaging with online content (e.g., Azevedo et al., 2004; Quintana et al., 2005). Moreover, students who fail to monitor their work products and modify their processes also fail to learn effectively in hypermedia environments (Azevedo & Cromley, 2004; Azevedo et al., 2004; Moos & Azevedo, 2008). Research in hypermedia learning has demonstrated that highly contextualized, personalized support—such as prompting from a human tutor—can significantly increase students' success in regulating their learning strategies (Azevedo et al., 2008), including monitoring progress toward learning goals and coordinating information sources. Although personalized, human-driven support is effective, one-on-one human tutoring is not a scalable approach to supporting online learning. A major challenge for cognitive personalization tools is to demonstrate that they can support learners in using the metacognitive processes that they need to become effective, *self-directed* learners.

As discussed previously (in Section 3.3), knowledge can exist at multiple levels that correspond to more *shallow* or *deep* levels of understanding. Effective metacognitive processes are those that support learners in developing deep, situation model knowledge, whereas ineffective processes support development of surface or textbase levels of knowledge. We draw upon the CI model and the revision codes developed by Wiley and Voss (1996, 1999) to characterize metacognitive processes, namely, analysis of essay content (which we call "essay analysis") and revision strategies. These processes can be categorized as either preserving the represented content (*shallow processes*) or leading to transformations and/or integration of the represented content (*deep processes*). These analyses and strategies are shown in Figure 3.

## 3.5. Prior Knowledge Integration

The primacy of prior knowledge to act as a frame that structures and guides self-directed learning processes has been well established in the learning sciences. Students who lack relevant prior knowledge often have difficulties in managing their learning paths through free-choice learning environments (e.g., hypermedia systems: Chen & Ford, 1998; Last, O'Donnell, & Kelly, 2001); self-regulating their online learning (Moos & Azevedo, 2008); and focusing their attention on deeper, more effective learning strategies (Alexander, Jetton, & Kulikowich, 1995; Murphy & Alexander, 2002). These results highlight the important role that prior knowledge has as a mediating influence on student behaviors involved in sensemaking.

During self-directed learning tasks, supporting integration with prior knowledge requires careful mediation between a student's individual understanding and

**FIGURE 3. Shallow and deep metacognitive processes during essay analysis and revision.**

| Essay Analysis | Student-Reported Diagnosis |
|---|---|
| Shallow Analysis | Grammar, Spelling Errors<br>Poor Writing Style<br>Wordiness |
| Deep Analysis | Inaccurate Content<br>Missing Content<br>Unclear, Vague Content |
| **Revision Strategies** | **Student-Reported Process** |
| Shallow Revision Strategy | Delete or Remove Idea<br>Fix Grammar or Spelling<br>Reword or revise style |
| Deep Revision Strategy | Add New Content<br>Revise Content<br>Describe or Explain Relationships |

misconceptions and the *ideal understanding* (see Figure 1) toward which she or he should be working. Prior knowledge is not a bank of information to which learners can deposit chunks of information or wipe clean and replace with a better set of information or concepts. Instructional supports are likely to be most useful when they work with, and *gradually improve*, existing knowledge frames rather than try to replace these frames (G. Klein et al., 2006b). However, the question of how to gradually improve a knowledge frame is a complex one. Individuals must be supported in finding relevant new knowledge that becomes a part of the student's new understanding and in using new and prior knowledge to generate inferences. Thus, we are interested in identifying both how often students encode new information about a domain and how they make use of such information during sensemaking.

## 3.6. Final Representation

The end result of educative sensemaking is a final learning representation that reflects students' current understanding of the topic or domain. In an essay writing task, this consists of a final essay that explains the student's current knowledge as related to the assigned topic. This representation can be analyzed to determine the amount of knowledge represented in the essay, and the degree to which revisions have corrected omissions or other misconceptions.

## 3.7. Supporting Educative Sensemaking

Few novice learners are prepared to fully engage in effective educative sensemaking, and learners rarely can be successful without external support. Unfortunately, the complex and nonlinear path of educative sensemaking processes can prove difficult for teachers to effectively scaffold in classroom environments. Educators often seek to break down tasks into clear, manageable procedures that can undermine the value of complex sensemaking tasks like writing a scientific explanation. Previous research has found that teachers can compromise the value of scientific explanations for learning when they attempt to modify the task in order to reduce its complexity (McNeill, 2008). By structuring the task, teachers do succeed in lowering the cognitive costs involved in educative sensemaking but often wind up with the unintended consequence that they also simplify the task in ways that reduce students' deep processing of educational resources (e.g., by breaking down explanations into formulaic statements that require minimal reflection).

Thus, there is a great need for educational tools that can effectively structure sensemaking tasks without offloading the complex cognitive and metacognitive processes that allow students to engage meaningfully with information sources. Successful support must lower the cost of representational shifts during educative sensemaking at the same time that it increases student use of effective metacognitive processes that drive changes in knowledge representations. In the next section, we describe a prototype cognitive personalization service to support educative sensemaking: CLICK.

## 4. CLICK: A CUSTOMIZED LEARNING SERVICE FOR CONCEPT KNOWLEDGE

The ultimate goal of this research is to develop fully automatic, domain-independent cognitive personalization algorithms that can be flexibly embedded, using a web service protocol, in a range of learning environments and web portals to support the sensemaking activities of self-directed learners. To date, we have designed and implemented a prototype "**c**ustomized **l**earning serv**i**ce for **c**oncept **k**nowledge" (CLICK) and created a demonstration adaptive essay writing environment using this service. As described in the next section, we have evaluated the impact of CLICK on students' sensemaking processes within a specific testbed domain: high school plate tectonics.

CLICK provides personalized feedback based on specific, identified knowledge gaps or misunderstandings occurring in the learner's essay, hence our use of the term *cognitive* personalization: We assume that gaps or misunderstandings in students' essays signal gaps or misunderstandings in students' existing knowledge. The CLICK algorithms identify knowledge gaps or misunderstandings by comparing the student's essay with an idealized representation of age-appropriate understanding for the topic at hand (see Figure 1). The algorithms use information about the type and scientific content of identified knowledge gaps or misunderstandings to select educational resources that specifically target these issues. Rather than simply pointing out potential errors, CLICK feedback is designed to support learners' metacognitive processing (see Figure 1) by helping them to fluidly shift between the science content of the educational resources and their own, evolving science explanation in ways that support effective learning. CLICK changes the dynamic in how students spend their time and mental resources when engaged in self-directed learning with web-based materials, moving from activities that primarily exercise students' information-seeking skills to activities that emphasize the development of their science content knowledge using metacognitive skills. Our current prototype is designed to support more effective self-directed learning by:

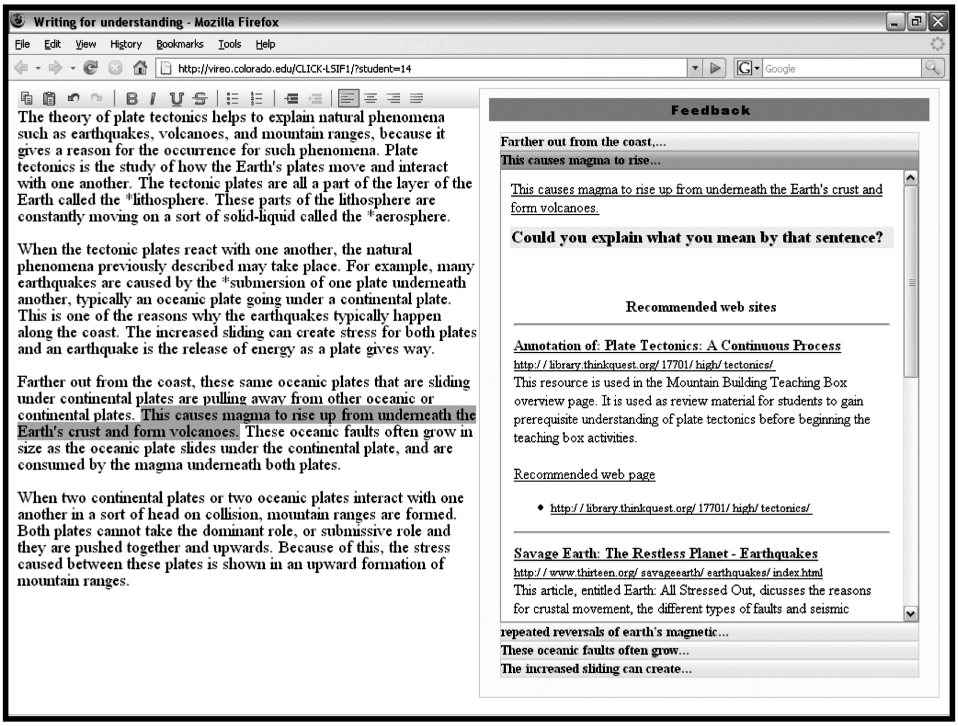
- Helping students analyze the accuracy and adequacy of their scientific explanations in capturing key science concepts and relationships between concepts. This support should help students develop their metacognitive processes and behave more like skilled, self-directed learners.
- Helping students seek new knowledge relevant to the task at hand by suggesting educational resources that are most appropriate for developing scientific knowledge in the context of their current essay. This support should help students encode knowledge at deeper levels of representation, facilitating meaningful revisions to the representation.
- Helping students to integrate new science knowledge among multiple resources and with their own prior knowledge by identifying mismatches between concepts found in educational resources and their own scientific explanations. In our case, the goal is to support students to modify their scientific explanations in ways that reflect their new conceptual understandings.

The adaptive essay writing environment shown in Figure 4 illustrates these three kinds of support.

## 4.1. Student View

Students write in the editing area on the left, and CLICK provides feedback in the scrollable area on the right, including suggestions for interactive learning resources they can explore to further their science knowledge. CLICK analyzes the student's scientific explanation to identify sentences that may be indicative of vague, incorrect, or missing conceptual knowledge. Instructional feedback is provided for each targeted sentence, including an external prompt to guide the student's thinking and motivation, and a list of recommended resources. In Figure 4, CLICK has identified the statement, "This causes magma to rise up from underneath the Earth's crust and form volcanoes" as an incomplete understanding of volcano formation and has displayed an instructional prompt aligned to this type of knowledge problem: "Could you explain what you mean by that sentence?" The recommendations include three resources identified as best fits to help the student improve on his or her understanding. Each recommendation includes the resource title, the resource web site URL, a short description, and a suggested specific page within the site. This approach is explicitly designed to encourage the students to reflect on their prior knowledge and thus to pave the way in helping them to integrate new knowledge from the suggested resources into their current understandings. The system also

**FIGURE 4. The prototype adaptive essay writing application.**




helps students to more easily apply their new knowledge by explicitly highlighting problematic sentences on which students should focus their attention.

In Figure 4, CLICK has identified a vague or incorrect sentence. The system also points out two other different types of knowledge problems: fragmented knowledge and knowledge gaps. Fragmented knowledge issues are indicated when the same concept is discussed in different parts of an essay and these parts do not appear to be well connected. Knowledge gaps occur when the student omits an important concept from his or her essay altogether. As described next, it is by comparing the student's essay with a representation of idealized understanding that enables CLICK to identify different *types* of knowledge gaps and to make personalized resource recommendations based on the type and content of the knowledge gap. The system provides an array of different instructional prompts, based on the work of Chi and her colleagues (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). Each prompt is aligned to a specific type of knowledge problem, with the aim of providing the most appropriate prompt to support learners' metacognitive processes while not repetitively showing the same prompts over and over again. For instance, an alternative prompt to the one shown in Figure 4 might have been, "What makes you think so?" An example prompt for a fragmented knowledge issue is, "How are these two concepts related?" whereas a prompt for a knowledge gap issue might be, "What's the main point of this concept?"

## 4.2. CLICK Algorithms

Here, we aim only to give the reader a flavor for how the overall system works; each of the algorithms comprising CLICK have been described in detail elsewhere (de la Chica, Ahmad, Martin, & Sumner, 2008; de la Chica, Ahmad, Sumner, Martin, & Butcher, 2008; Gu, de la Chica, Ahmad, Khan, & Sumner, 2008). CLICK consists of three major algorithms. One algorithm produces representations of current student knowledge and idealized domain knowledge. A second algorithm compares these representations to identify potential knowledge problems in the student's essay. A third algorithm uses information about the type and scientific content of identified knowledge problems to make personalized educational resource recommendations.

CLICK creates internal *knowledge map* representations of current student knowledge and idealized domain knowledge. Knowledge maps are a type of concept map, where the nodes contain knowledge propositions (phrases or sentences), as opposed to keywords; cognitive science research has highlighted the utility of knowledge maps for diagnosing student understanding and for representing the macrolevel structure of a domain (O'Donnell, Dansereau, & Hall, 2002). CLICK uses natural language processing algorithms to create two knowledge maps: One depicts what students currently understand (as represented in their scientific explanation) and one depicts what an informed person of the target age group might be expected to know about a scientific topic. We refer to this representation of idealized understanding (represented as an information source in Figure 1) as a domain knowledge map (de la Chica, Ahmad, Martin, et al., 2008). The domain knowledge map is created automatically by extracting key science concepts and their relationships from a set of web-based learning resources. Underlying this approach is the assumption that a carefully selected suite of age-appropriate, high-quality learning materials implicitly represent what educational experts believe that students should know about a particular topic. In this particular case, the domain knowledge map was constructed from 20 web-based resources selected by four earth science education experts. It describes what high school students should know about plate tectonics and related phenomena such as earthquakes, volcanoes, and mountain formation (Ahmad, de la Chica, Butcher, Sumner, & Martin, 2007).

The student knowledge map and the domain knowledge map are algorithmically compared using graph-theoretic techniques to diagnose current student understanding (de la Chica, Ahmad, Sumner, et al., 2008). CLICK can diagnose three different types of conceptual problems: incorrect statements (where concepts expressed in the student essay contradict scientifically accurate concepts in the domain map), incomplete understanding (where students provide a correct but only partial or vague description of a concept, or, students fail to mention a concept), and fragmented knowledge issues (where students provide correct descriptions of two related science concepts but fail to make explicit connections between them in their essay). In graph terms, these three different types of conceptual problems correspond to differences in the content of nodes in the student and domain knowledge maps, missing nodes in the student knowledge maps, and unconnected or poorly connected nodes in the student knowledge map, as compared to the idealized domain map. The output from this algorithm is a list of specific, identified knowledge problems from the student essay. Each identified knowledge problem is represented as (a) a sentence from the student essay, (b) the type of knowledge problem to which each sentence corresponds, (c) a fragment of the student knowledge map containing this problem and surrounding concepts, and (d) the corresponding fragment of the domain map depicting an idealized knowledge representation.

A personalized recommendation engine selects specific interactive digital library resources (Figure 1) to address each of the identified conceptual problems (Gu et al., 2008). The recommendation engine transforms the domain and learner knowledge map fragments into a single data structure—a detailed concept graph—representing both the student's current and desired knowledge. This detailed concept graph represents key terms as graph vertices, links depict related terms, and both the importance of terms and their distance from each other in the original knowledge maps are represented as weights applied to each link. This term-based data structure is designed to facilitate information retrieval, as documents are indexed by terms, while preserving the semantic structure of both the student's current and desired knowledge as represented in the knowledge map fragments. A similar term-based concept graph is produced for each educational resource that the system searches over. In the current prototype, these resources are drawn from a test bed collection created for this research, which contains 796 age and topic-appropriate learning resources drawn from the Digital Library for Earth System Education (http://www.DLESE.org). These web-based learning resources include scientific visualizations, animations, imagery, scientific data, and other interactive materials. Resources are automatically recommended to the learner based on the similarity between the student concept graph and a particular resource's concept graph.

This approach differs from conventional recommendation engines that rely on collaborative filtering approaches, which suggest resources based upon the prior actions of other users (see Herlocker, Konstan, Terveen, & Riedl, 2004, for examples). Social annotation systems are similar in spirit to collaborative filtering systems; in these systems, comments left by other users serve to guide and scaffold interactions with web resources. In a laboratory study, Nelson et al. (2009) demonstrated that carefully constructed social annotations, such as those provided by domain content experts, can support users to learn from multiple web resources. Although the goal of this research is similar to ours, the underlying computational mechanisms differ significantly. Our approach blends knowledge modeling, often associated with cognitive tutoring systems (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995; Corbett, McLaughlin, & Scarpinatto, 2000), with content-based recommender systems. Recent advances in content-based recommender systems are pursuing hybrid approaches that combine analysis of user actions with analysis of document content (Will et al., 2009) and other contextual factors such as task (White, Bailey, & Chen, 2009). Many of these systems are exploring different approaches for inferring users' intentions and goals by carefully examining the content of the web sites they visit and the usage history of their visits. In our approach, user intentions and goals are not derived from observing web site usage or individual web site content. Instead, given our focus on supporting educative sensemaking in classroom settings, the user's current knowledge state (roughly analogous to intentions) is derived from analyzing their scientific explanations (essays), and the user's desired knowledge state (roughly analogous to goals) are derived from the idealized understanding represented in the domain knowledge maps. By including a representation of idealized understanding in the process, we hope to avoid the situation where common misconceptions are reinforced by prior user actions. In science education, many students share the same misconceptions and knowledge gaps; relying only on prior user actions could reinforce these suboptimal understandings rather than helping students to overcome them.

### 4.3. Developer View

By utilizing best practices in software architectures, we are making CLICK's personalization capabilities available through a web service protocol. When fully realized, this approach should enable a broad spectrum of instructional designers, learning application developers, and educational researchers to embed this capability in their own applications and thus benefit from this research. Design requirements for the web service protocol were extracted from detailed analyses of the scenarios, mock-ups, and storyboards produced during the Future Learning Environments Workshop. An initial version of the web service protocol was developed, evaluated, and refined using iterative user-centered design methods, namely programming walkthroughs. Programming walkthroughs are a variant of cognitive walkthroughs (Wharton, Rieman, Lewis, & Polson, 1994)—a type of usability inspection technique—and are meant to assess both the ease of writing programs in the proposed language or protocol and the expressiveness of the language for writing different applications (B. Bell & Weaver, 1994). The resulting web service application programming interface exposes CLICK's problem diagnoses and knowledge map generation capabilities via several request types, including submit or remove a knowledge map, construct student knowledge map from essay, construct domain map from resource URLs, get student misconceptions and knowledge gaps, get key concepts from domain knowledge map, and get related concepts. Details of this web service protocol are described in Ahmad (2008). This protocol exposes the output of each of CLICK's three algorithms, not just the final output of the personalized recommendation system. This enables developers to construct a wide variety of interfaces and interaction models, not just the simple interface shown in Figure 4. In other applications, we have experimented with providing learners with visual representations of the domain knowledge map and their own knowledge map (constructed from their essay). The interface shown in Figure 4 was used in the learning study described in the next section. It was selected for use in the study due to its simplicity, usability, and familiarity to students.

## 5. EMPIRICAL STUDY

To assess the potential value of the CLICK personalization service on students' educative sensemaking, we designed and conducted a controlled learning study to explore the benefits of CLICK's cognitive personalization tools as implemented in an adaptive essay writing environment. This study assessed the effects of personalization on students' cognitive and metacognitive skills and on the essays that they produced as they wrote a scientific explanation with the help of digital educational resources.

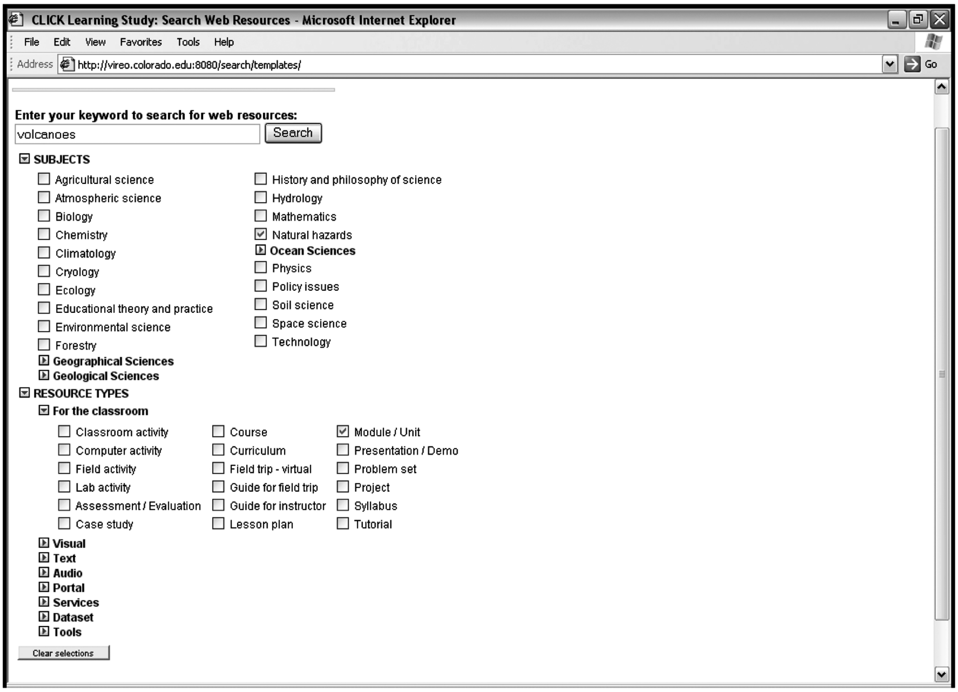
## 5.1. Method

### Design

The study took place in two sessions. In Session 1, students first completed an initial knowledge assessment to test their prior knowledge of the domain of study. Then students collected information about the topic and created a knowledge representation in the form of an initial essay. These activities align to the *search for representations* and *instantiate representations* learning loops in sensemaking (Russell et al., 1993). All of the student-generated essays were evaluated by CLICK algorithms (Ahmad, 2008; de la Chica, Ahmad, Sumner, et al., 2008) to identify potential knowledge problems for the students to address in Session 2.

Session 2 took place approximately 1 week after Session 1. In Session 2, students were randomly assigned to either the CLICK condition or the control condition. Students in the CLICK condition used the adaptive essay writing environment shown in Figure 4 as they revised their essays. Students in the control condition were provided with CLICK-generated essay feedback presented as printed text (five areas, usually sentences, from their essays identified as being problematic). However, students in the control condition were not provided with the associated list of recommended resources for each problematic sentence. Instead, they were provided with a simplified digital library search interface (see Figure 5) and asked to find resources to support essay revisions. This interface searched over the same testbed collection of resources used in the CLICK condition; all 796 resources in the collection were age- and topic-appropriate, high-quality resources from an educational digital library. This customized testbed reduced the demands of information search for students in the control condition, increasing the ease with which they could identify relevant information resources. We recorded students' on-screen interactions as they revised their original essays and used digital resources to inform their revisions. These recordings allowed us to conduct a detailed, in-depth analysis of the ways in which students analyzed their essays, made use of the feedback and digital resource recommendations (if available), and revised their essay representations. After students finalized their essays, we asked them to reflect on their analyses of essay issues and their strategic approaches to revision (i.e., their metacognitive processes). Finally, students completed the knowledge assessments for a second time. Qualitative and quantitative methods were used to assess (a) the depth of students' revisions to essay representations during sensemaking, (b) students' metacognitive processes (i.e., essay analysis and revision strategies), (c) prior knowledge integration, and (d) information encoding as evidenced by final essays.

**FIGURE 5. The digital resource search interface used by the control condition.**




## Materials

*Domain Knowledge Assessments*

***True/False Test.*** The true/false test was developed in consultation with a domain expert in geology, who reviewed the scientific accuracy of questions. The overall assessment contained 40 true/false questions (e.g., "Volcanic activity can occur in the middle of a lithospheric plate"). One point was awarded for each correct response, for a maximum of 40 points possible. The true/false test assesses factual knowledge about the domain but does not require knowledge application or transfer. Thus, the true/false test provides a good measure of basic, factual knowledge in domain but does not measure deep understanding of that information (see Kintsch, 1998, for a discussion).

***Short Answer Test.*** The short answer test consisted of five questions about plate tectonics and changes to the earth's surface (e.g., "John says that most volcanoes are located at transform plate boundaries, where friction between plate boundaries creates great heat that melts rock into molten lava. Is John right about the typical location and cause of volcanoes? Use what you know about plate boundaries to explain your answer."). Students were awarded 1 point per correct idea unit in their responses; a total of 30 points were possible overall. Correct responses to short-answer questions could not be found directly in learning materials but required students to apply, explain, and integrate domain knowledge to generate correct responses. Thus, the short-answer test measures deep understanding by assessing students' knowledge application and transfer of learned concepts to new situations (see Kintsch, 1998).

*Revision Questionnaire*

This questionnaire targeted the metacognitive processes of students, asking them to articulate their analysis of essay problems (portions of their essays highlighted by CLICK for feedback) and the revision strategies that they used as they worked with their essays in response to feedback. The revision questionnaire presented students with four of the sentences/statements from their original essays that had been targeted for revision in the feedback that they had received; for each sentence, the revision questionnaire asked students to respond to three prompts:

- Why do you think the system identified this statement for revision?
- Describe how your revised statement is different from your original one. Please be as specific as you can.
- Explain what you did to revise your original statement and why. Please be as specific as you can.

For each prompt, students were asked to generate a written response that was not constrained in length, format, or content. Students were free to respond to the prompts as they chose, providing as much or as little information as they desired.

## Participants

Thirty undergraduates from the University of Colorado at Boulder completed the study. Each participant received partial course credit in an introductory psychology class upon completion of the study. Overall, 30 students (20 female, 10 male) were randomly assigned to either the control ($n = 15$) or the CLICK ($n = 15$) condition.

## Procedure

***Session 1.*** This session was the same for all students, regardless of their assigned experimental condition. During Session 1, students' prior knowledge about the domain of study (plate tectonics) was assessed using the true/false and short-answer tests. Students had up to 25 min to complete the true/false test (10 min) and the short-answer test (15 min).

After prior knowledge testing was completed, all students were given 15 min to familiarize themselves with basic information about plate tectonics using a set of five digital resources (presented in random order) that had been handpicked by an earth sciences domain expert as exemplary resources on the theory of plate tectonics, earthquakes, volcanoes, and mountain formation. Students used this resource-exploration time to develop concepts and select ideas in preparation for writing their scientific explanations. This task was designed to be consistent with a typical educational scenario, in which a teacher or professor hands out a list of resources that students can use as background information to become familiar with a topic.

To prevent rote recitation of the digital resources' information from short-term memory, students next spent 5 min completing a learning-styles questionnaire before they moved on to generate an initial representation: an initial essay about earthquakes and plate tectonics. All students wrote their initial essays using standard word-processing software and were given 30 min for this task. In the interest of capturing students' existing knowledge representations at this point in time, students were not given access to the digital resources during writing. Essays were required to be at least 250 words and were written in response to the following prompt:

> Prior to the development of the theory of plate tectonics, geologists had difficulty understanding the origins of earthquakes and mountains. How does the theory of plate tectonics help us explain natural phenomena such as earthquakes, volcanoes, and mountain ranges? Please be as specific as you can in your explanation.

All student essays produced during Session 1 were processed by CLICK. CLICK identified the erroneous and problematic statements written by students, using the knowledge map comparison processes described earlier. As reported elsewhere (Butcher & de la Chica, in press; Butcher et al., 2008; de la Chica, Ahmad, Sumner, et al., 2008), CLICK diagnosis was highly accurate and reflected approximately 84% of the issues targeted by human experts. CLICK algorithms also were used to automatically select a set of resources that were aligned to each of the CLICK-identified misunderstandings in a student essay. Although CLICK automatically analyzed student essays and retrieved associated digital resource recommendations, at the time that the study was conducted, it had no mechanism of prioritizing the order in which misconceptions should be presented to students. Given a limited time for essay revision in this research, the study presented only five misconceptions to the student. Hence, the prioritization of problematic statements was achieved by a Wizard of Oz approach (Kelley, 1983). The Wizard of Oz approach was not used to emulate perfect performance but instead simulated a more realistic, 80% success rate. This was achieved by having two researchers use rich textual descriptions of prioritization algorithms under development to choose four CLICK-identified issues from among the total set to present to students, with a fifth issue chosen at random.

***Session 2.*** Students returned for the second session approximately 1 week after Session 1. During this session, all students were given 35 min to revise their original essays and were provided access to digital resources via the Internet. As described next, students' interactions differed according to their assigned experimental condition. We recorded all students' on-screen behaviors during essay revision using an off-the-shelf screen capture tool. After students finished revising their essays, they completed the Revision Questionnaire (20 min), the true/false test (10 min), and the short-answer test (15 min). These true/false and short-answer tests contained the same items as in Session 1, but the questions appeared in randomized order.

The experimental procedure varied by study condition only during the essay revision task; although all students had access to digital resources during revision, only the CLICK students were provided with recommended resources that had been automatically selected to align with CLICK's diagnosis of problematic statements. Thus, students in the different experimental conditions used different user interfaces during revision. Students in the CLICK condition first received a 5-min tutorial using a sample essay on how to operate the CLICK essay writing interface. Next, students in the CLICK condition received their essay feedback integrated within the CLICK essay writing application (see Figure 4). As described earlier, feedback included five sentences from their essays selected by CLICK as being potentially problematic, an associated instructional prompt, and a suite of three recommended resources for each identified sentence. Students in the CLICK condition were not given access to the search interface provided to students in the control condition; CLICK students were limited to the recommended resources identified by the system. All students were free to use the digital resources in any way that they chose (e.g., copy and pasting information directly, if they so desired).

During the essay revision task, students in the control condition first were given a 5-min tutorial on how to use the simplified digital library search interface (see Figure 5) to locate digital resources. Students then were given printed feedback on the five selected CLICK-identified issues from their essays as well as a printed copy of five recommended, general essay revision strategies that were similar to the instructional prompts used in the CLICK condition:

1. Explain or restate what you wrote in your original essay
2. Clarify what you wrote in your original essay
3. Be specific describing concepts
4. Describe concepts using your own words, and
5. Explain how concepts may be related to each other.

Control students revised their original essays using a word processor, using the simplified digital library search interface to find and link to relevant digital resources as desired. As in the CLICK condition, students in the control condition were free to use the digital resources in any way that they chose.

## 5.2. Results

Results are structured according to the educative sensemaking model discussed in Section 3 and depicted in Figure 1.

### Generating an Initial Representation

To assess students' initial knowledge representations (i.e., their prior knowledge), we analyzed students' scores on the knowledge assessments taken in Session 1 and their initial essays.

***Prior Knowledge Assessment.*** In Session 1, analyses demonstrated no differences in students' initial true/false or short-answer test scores across conditions ($F$s $< 1$). On the 50-point true/false test, students demonstrated low prior domain knowledge in both the CLICK ($M = 28.6$, $SD = 4.3$) and the control ($M = 27.9$, $SD = 5.8$) conditions. Given that chance performance is 50%, or 25 points, it is clear that the participants in this study qualify as novice learners for this topic. On the short-answer test, students also scored at a novice level in both the CLICK ($M = 2.7$, $SD = 2.2$) and the control ($M = 3.4$, $SD = 1.9$) conditions.

***Initial Essays.*** In Session 1, analyses demonstrated no differences in initial student essays in terms of holistic quality, inclusion of domain content, or clarity ($F$s $< 1$). Students' initial representations, then, were of similar quality across the experimental conditions.

**Revising Representations**

We assessed students' observable revisions to their initial essays using the screen-capture recordings generated in Session 2 of the study. First, we analyzed students' allocation of their time during revision. Data include the total length of time that students spent revising their essays (e.g., actually making changes to the text), viewing their essays, exploring links to digital library resources, and viewing the digital resources.

***Time on Task.*** Analyses demonstrated that students did not differ ($F < 1$) in the amount of time they spent revising their essays overall (CLICK: $M = 30.5$, $SE = 1.7$; Control: $M = 31.83$, $SE = 1.7$). Nor did they differ ($F$s $< 1$) in the amount of time that they spent seeking new information (CLICK: $M = 9.2$, $SE = 1.1$; Control: $M = 8.3$, $SE = 1.1$) or making actual changes to their essays (CLICK: $M = 18.0$, $SE = 1.3$; Control: $M = 18.1$, $SE = 1.3$). Thus, it was not the case that students were disproportionately searching for resources in the control condition, nor was it true that students in the control condition gave up on the task before students in the CLICK condition.

***Use of Digital Library Resources.*** We recorded how often students switched between their essays and the digital resources, as a measure of how fluidly they were able to move between representations. Fluid movement between representations should support representational shifts, as students are able to assess the conceptual adequacy and accuracy of their own essay in comparison to the body of knowledge with which they are working.

Although the experimental conditions did not differ in the overall amount of time that they devoted to data collection versus changing their representations, there were clear differences in how students moved between domain data (the digital library resources) and their developing representation (the essay). Students in the CLICK condition moved more fluidly between data and representation, switching between digital resources and their essays significantly more often, $F(1, 28) = 4.6, p = .04$, than students in the control condition. CLICK students averaged more than 57 switches ($SE = 5.8$) compared to control students' 40 switches ($SE = 5.8$) over the course of their 35-min revision period. Higher numbers of switches were correlated with less frequent preservation of original essay ideas ($r = -.42, p = .02$) reported in the revision questionnaire—that is, the more students moved fluidly between representations, the less likely they reported trying to preserve their original understanding in their revised essays. This analysis demonstrates the potential promise of promoting fluid movement across instantiated representations and learning materials.

***Shallow Versus Deep Revision Processes.*** We categorized students' observable interactions with digital resources and their essays as either *shallow* or *deep*. As noted earlier in Section 3, these categories were based on the scoring methods developed by Wiley and Voss (1996, 1999) and levels of knowledge representation drawn from text comprehension theory (e.g., Kintsch, 1998). As seen in Figure 2, shallow processes were coded when students' revision behaviors resulted in no knowledge transformation or meaningful integration of concepts, and deep processes were coded when revision behaviors transformed or integrated information.

Students in the CLICK condition demonstrated significantly *greater* use of deep revision processes, $F(1, 27) = 5.2, p = .03$, compared to control students who received essay feedback without personalized recommendations (see Figure 6 for means and standard deviations). Conversely, students in the CLICK condition tended to make fewer shallow revisions compared to students in the control condition, $F(1, 27) = 3.6$, $p < .07$. CLICK students still engaged in a fair amount of shallow processing, as would be expected by the difficult nature of this kind of educative sensemaking. Novice students face major obstacles in determining what information to include in their essays, how to describe and relate it to their other content, and how to integrate the information into the overall explanation that they are developing. Thus, we consider the balance between CLICK students' profile of deep versus shallow processes to reflect important progress in educative sensemaking for novice students. Students may frequently need to draw scientific concepts or examples directly from information resources, but these instances of borrowing should be offset by a substantial amount of knowledge transformation in which students work to make sense of the materials and to integrate them in meaningful ways.

These data suggest that personalized feedback acts both as top-down, goal-directed guidance as well as bottom-up support. Personalized recommendations scaf-

**FIGURE 6. Means (and standard deviations) for revision processes by condition.**

| Process Type | Control Condition | CLICK Personalized Feedback |
|---|---|---|
| Deep processes: % of total revisions | 27 (23) | 48 (27) |
| Shallow processes: % of total revisions | 67 (27) | 49 (24) |

fold students' attention to knowledge problems (cf. "residue" in Russell et al., 1993), providing resource links that allow students to focus fully on the analysis of these information sources and their potential impact on their developing representation.

### Metacognitive Processes

Because students spent a total of only 35 min revising their essay during Session 2, we were especially focused on assessing the impact of CLICK feedback on students' metacognitive processing during learning. Process analyses are able to capture more subtle and fine-grained changes to students' sensemaking activities than knowledge-based assessments. Standard knowledge assessments are unlikely to show large differences in student knowledge over the course of one 35-min learning session, especially when considering the inherent difficulty of changing conflicting prior knowledge in science (Chi, 2008). Process data offer a sensitive measure to assess impact by observing the cognitive and metacognitive processes that we know to be associated with deep learning over time.

To address the depth of students' metacognitive processes during revision, we analyzed student responses on the Revision Questionnaire. Verbal protocol analysis techniques (Chi, 1997) were used to identify the total set of revision responses that represented distinct types of analysis for target sentences and students' reported strategies for revising their essays. These responses were categorized into deep and shallow approaches to essay analysis and revision, based on the criteria developed and used by Wiley and Voss (1996, 1999). Shallow analyses reflected no attempt to transform or integrate learning materials, instead focusing on superficial elements of the essay separate from domain content. For example, students might report that they identified problems with grammar, spelling, or style. In response to essay feedback, students reporting shallow revision strategies might note that they revised the wording of a sentence, fixed spelling or grammatical errors, or simply removed the sentence that was highlighted by feedback in order to remove the potential problem.

In contrast, deep analyses reflected the intention to transform and integrate information in ways that support development of a flexible, situation model representation of knowledge (Kintsch, 1998). Deep analyses included the attempt to identify and integrate new information into the essays (e.g., noting missing content) and to improve the accuracy or meaning of the information that the essays contained (e.g., identifying that it was incorrect or poorly explained). Deep revision strategies articulated behaviors that would result in the transformation or integration of information, such as finding new information or better explaining essay content.

Shallow and deep categories were used to code all written statements that students had produced in response to questionnaire prompts. Figure 7 presents example statements from each shallow and deep category.

Overall, clear patterns emerged from students' self-reported metacognitive processes. Controlling for the number of codable statements produced by students, a multivariate analysis of covariance demonstrated that students in the CLICK con-

**FIGURE 7. Example deep and shallow statements from revision questionnaires.**

| Essay Analysis | |
|---|---|
| Shallow analysis | |
| Grammar, spelling | "There 'is' is not proper grammar, it should say there 'are' always" |
| Style | "[The sentence] was intended to be somewhat of a hook for the paragraph, but it is not very interesting or insightful" |
| Wordiness | "The statement seemed to have too many words and could be stated with fewer in a more concise way" |
| Deep analysis | |
| Inaccurate | "Because the fault lines that I pointed out in my original essay were incorrect. The system wanted a more accurate and specific answer." |
| Missing content | "I didn't really explain or identify the theory of plate tectonics. I needed to give a definition of it and where those plates came from." |
| Unclear, vague | "It wasn't clear what I meant by 'move horizontally or vertically'" |

| Revision Strategies | |
|---|---|
| Shallow strategies | |
| Delete idea | "I took out the sentence completely. I did not like it in my essay anymore … I deleted it" |
| Fix grammar or spelling | "In 'heating and cooling' I turned it around to say 'heated and cooled' since it was past tense" |
| Reword or revise style | "I literally just changed some words and added transitions to the sentences to start new ones" |
| Deep strategies | |
| Add new content | "I included the type of boundary that is usually associated with earthquakes and expanded on that" |
| Revise content | "I read more about the formation of mountains and then rewrote the statement so that it included accurate information" |
| Describe or explain relationships | "My revised statement explains how earthquakes are formed by a specific type of fault where the land slides against each other" |

dition reported deep revision strategies more often than students in the control condition, $F(1, 26) = 4.9, p < .04$, and, conversely, tended to report fewer shallow revision strategies, $F(1, 26) = 3.1, p = .09$ (see Figure 8).

Results for students' self-reported analysis of their essays follows the same pattern, although they do not reach the level of statistical significance ($F$s $< 1.4, p >$ .25). Students in the CLICK condition diagnosed potential essay problems by analyzing essay content in deep, compared to shallow, ways (see Figure 8). For control students, this pattern was reversed. For all students, the depth of their analysis predicted the depth of their revision approaches: Shallow analysis was correlated with shallow revision approaches and deep analysis was correlated with deep revision approaches ($r = .52, p = .003$).

**FIGURE 8. Means (and standard deviations) for shallow versus deep sensemaking approaches.**

| Students' Self-Reported Strategies | Control % | CLICK % |
|---|---|---|
| Essay analysis: Metacognitive processing | | |
| Shallow (Grammar, Wordiness, Spelling, Writing Style) | 33 (34) | 20 (27) |
| Deep (Incorrect, Missing Content, Too Vague or Broad) | 67 (34) | 80 (27) |
| Essay revision: Knowledge development | | |
| Shallow (Reword, Fix Grammar/Spelling, Delete) | 49 (28) | 38 (21) |
| Deep (Revise or Add Content, Describe Relationships) | 51 (28) | 62 (21) |

***Correlations Between Metacognitive Processes and Learning Outcomes.*** Overall, students' metacognitive skills predicted different patterns of performance across the true/false (factual) and the short-answer (application/transfer) tests. Analyses demonstrated a statistically significant relationship between students' metacognitive processes and learning outcomes for the short-answer test but not for the true/false test (see Figure 9). Deep analysis of essay feedback was positively correlated with improvement on the short-answer test (measuring transfer and knowledge application), but shallow analysis was negatively correlated with performance on the short-answer test. Thus, the more students engaged in deep metacognitive processing to analyze their representations, the more likely they were to develop a flexible understanding of domain content that could be transferred and applied to new situations. Conversely, deep processes were associated with lower levels of shallow, textbase knowledge. Students' stated revision strategies followed the same pattern. Deeper revision strategies were associated with larger increases in knowledge application (short-answer test performance) but not shallow, factual knowledge (true/false test performance; see Figure 9). These results are consistent with previous research in comprehension, which has shown that the development of a deep, situation model level of understanding can be emphasized at the expense, of shallow, textbase knowledge (e.g., Bransford & Franks, 1971; see Kintsch, 1998, for a discussion).

### Prior Knowledge Integration

***Integration With Prior Knowledge.*** An important part of sensemaking involves the ability to identify and encode new information and to integrate this

**FIGURE 9. Correlations between students' metacognitive processes and learning outcomes.**

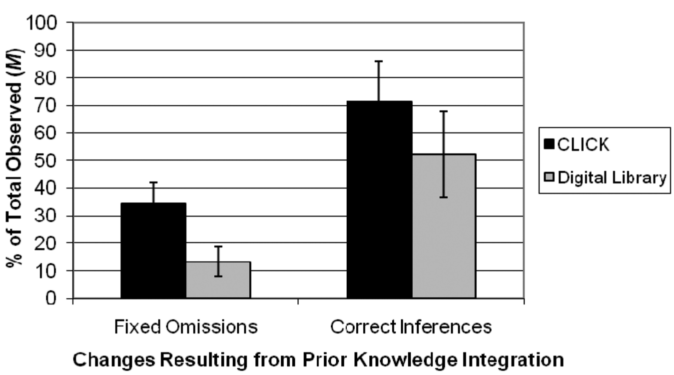
| | Shallow Essay Analysis | Deep Essay Analysis | Shallow Revision Strategies | Deep Revision Strategies |
|---|---|---|---|---|
| True/False (% of possible gain) | .10 | −.10 | −.026 | .026 |
| Short answer (% of possible gain) | −.37* | .37* | −.43* | .43* |

*$p < .05$.

information into a knowledge representation. Thus, essays were coded to assess the integration of *new*, relevant information; this information can take one of two forms. First, students can encode new information and add it to their essay explanation in response to errors of omission identified in the initial representation. Students also can encode new information and integrate it with prior knowledge, forming a deep (situation model) level of knowledge representation that can be used to add inferences about domain content to the essay. A research assistant coded essays for each of these two types of new information: (a) scientific content added to the essay that was aligned to essay feedback on errors of omission, and (b) scientific content that represented an inference drawn by the student (i.e., scientific content that was not drawn directly from a digital resource). We analyzed the percentage of total errors of omission that students fixed as well as the percentage of correct inferences (based on the total number of inferences) that students added to their essays. For an "error of omission" to be coded as "fixed," students needed to add relevant information into their initial representation in a manner that addressed the missing content and was well integrated into their initial explanation (e.g., students could not simply add a list of facts to the end of their essays).

Results demonstrated that CLICK students integrated content into their essays in order to correct errors of omission significantly more often than control students, $F(1, 26) = 4.4$, $p = .045$ (CLICK $M = 34.4$, $SE = 7.5$; control $M = 13.3$, $SE = 5.4$); in this analysis, we controlled for students' levels of prior knowledge by using students' pretest scores on true/false and short-answer tests as covariates. Although not statistically significant, $F(1, 11) = 2.2$, $p = .16$, the percentage of correct, content-based inferences that students added to their essays followed the same pattern (CLICK $M = 71.25$, $SE = 14.8$; control $M = 52.4$, $SE = 15.6$; see Figure 10). Researchers have noted that students can enrich existing conceptual understanding in science by encoding new information (e.g., Carey, 1991; Chi, 2008). These data demonstrate that CLICK support can help students enrich their existing representations via encoding of new domain content.

**FIGURE 10. Evidence for prior knowledge integration from essay representations: Percentage of fixed omissions and correct inferences by condition.**



It is important to note that these measures of prior knowledge integration were significantly associated with the percentage of deep revision strategies in which students had engaged ($r = .36$, $p = .05$) and negatively correlated to the percentage of shallow revision strategies in which they had engaged ($r = -.36$, $p = .05$), suggesting that deep revision strategies are a metacognitive process that can facilitate the integration of information during educative sensemaking.

As seen in the time on task analysis, experimental conditions differed neither in the amount of time that students spent viewing digital resources nor in the number of unique digital resources that students viewed. Thus, differences in the amount and type of prior knowledge integration cannot be explained by longer exposure to relevant information or exposure to a greater number of relevant resources. Instead, personalized resource recommendations may serve to structure the nature of cognitive interaction with the digital resources. That is, resource recommendations coupled with essay feedback may make the search for information more meaningful and may help students encode and integrate new information once it is identified.

***Student Awareness of Representation Shifts.*** One may question to what extent the data so far represent active sensemaking by students as they completed their educational task. That is, are students consciously engaged in the metacognitive processes that we associate with educative sensemaking? To address this question, we analyzed students' stated approaches to making representational shifts in their essays by coding students' responses on the reflection questionnaire for two factors. First, we coded whether students spontaneously reported trying to preserve the same idea in their revisions; when students report that they tried to preserve their original ideas during essay revisions, we can assume that they are *not* trying to make representation shifts in their essays. Second, we coded whether students spontaneously reported seeking digital resources to change or inform their intended revisions. When students report seeking new information to inform their changes, we can assume that they have identified a gap in their understanding that they are seeking to fill. As can be seen in Figure 11, the CLICK condition was significantly *less* likely to report preserving the same idea in their revisions, $F(1, 27) = 12.5$, $p = .002$, and significantly *more* likely to report pursuing digital resources to change their essay content, $F(1, 27) = 5.6$, $p = .026$.

These results demonstrate that CLICK's personalization supported students in identifying gaps or problems in their existing knowledge and, moreover, that they are less likely to avoid making representation shifts in their essays. Analyses of students'

**FIGURE 11. Means (and standard deviations) for self-reported representation shifts.**

| Students' Self-Reported Representational Shifts | Control # | CLICK # |
|---|---|---|
| (Spontaneously reported, 5 maximum) | | |
| Preserved same idea in revision | 2.1 (1.1) | .93 (.80) |
| Sought new information to change essay content | .93 (1.2) | 1.9 (1.3) |

revision processes and essays (reported in the next two sections) showed that students' stated approaches to essay revision were extremely accurate; students' descriptions matched their actual processes 92% of the time in the CLICK condition and 87% of the time in the control condition (these averages were not statistically different from each other: $F(1, 27) = 1.1, p > .30$.

## Final Representations

The analysis of learners' metacognitive processes during essay revision show clear differences between the CLICK and control conditions. However, one may question to what extent changes in learner processes lead to differences in sensemaking outcomes. That is, is there any evidence that deep metacognitive processes are actually associated with better essays if we examine students' final essay (i.e., their final representation)? Because students' stated revision strategies should be directly associated with changes to their essays, we analyze the impact of those processes here.

***Revision Strategies and the Quality of Students' Final Representations.*** Analyses show the metacognitive processes in which students engaged were a significant predictor of the final essay quality. A research assistant blind to experimental conditions scored each essay for holistic quality, taking into account the overall clarity and coherence of the essays. Correlational analyses showed that the number of deep revision strategies reported by students were correlated positively with holistic essay scores ($r = .40$, $p = .03$). Students' reported shallow revision strategies also were associated with essay improvement, showing a marginal but not statistically significant correlation to holistic quality ($r = .34$, $p = .06$). The contribution of shallow revision strategies to holistic essay scores is not surprising, as shallow strategies typically target style, language, and clarity of the essays. Overall, any revisions made by novice students in response to essay feedback are likely to improve the readability and overall quality of the essay itself. However, it is noteworthy that deep revision strategies have a demonstrable impact on the overall quality of students' final representations.

The research assistant who scored essays for holistic quality also scored each essay for the amount of relevant scientific content that students included—1 point was awarded for each unique, topic-relevant idea provided in the essay. As with holistic quality, deep revision strategies were positively correlated to content scores ($r = .36$, $p < .05$), but the correlation between shallow revision strategies and content scores was not statistically significant ($r = .32$, $p > .08$). Thus, deep processes supported better encoding of domain information. The amount of scientific essay content was strongly correlated with holistic quality scores ($r = .97$, $p < .01$), indicating that students did not randomly add information to their essays but crafted coherent explanations around scientific content.

Although these analyses point to overall links between deep processes and higher quality work products, it is difficult to identify wholesale differences between essays written by students in the CLICK versus control conditions. Our evidence for improved knowledge representations in the CLICK condition is mainly indirect. Final essays did not differ significantly by group in either holistic quality or in scientific content, likely due to the relatively short amount of time that students worked on their essays as well as wide variation in quality within each group. However, we believe that the correlations between deep metacognitive processes and essay quality, as well as the correlations between deep metacognitive processes and scores on the short-answer test, provide evidence that cognitive personalization tools have the potential to influence meaningful learning processes and, subsequently, student knowledge development.

## 6. GENERAL DISCUSSION

The current results show that personalization technologies can support educative sensemaking in three important ways. First, personalization appeared to support students' use of deep, metacognitive processes (essay analysis and strategic revision) during educative sensemaking. Results from the current study show that students who received personalization support were more likely to engage in deep, content-based analysis of their essays (focusing on the adequacy and accuracy of their knowledge) and were more likely to take deep approaches to revising their essays (focusing on the identification, encoding, and integration of new information). Conversely, students who did *not* receive personalization support were more likely to engage in shallow, stylistic analysis of their representations (focusing on grammar, spelling, etc.) and to use revision strategies that did not reflect changes to knowledge representations (removing ideas or making stylistic changes while trying to retain originally expressed ideas). Moreover, deep metacognitive processes focused on analysis of essay content were positively correlated with a knowledge assessment that required application and transfer of domain knowledge (the short-answer test) and negatively correlated with a knowledge assessment that required only increased memory for shallow facts (the true/false test). Deep metacognitive processes focused on revision strategies were positively correlated with the meaningful integration of relevant content into essay representations. Results support the theoretical importance of metacognitive processes in educative sensemaking as well as the potential for personalization technologies to support students in implementing them.

Second, personalization helped students work with their representations in meaningful ways as they engaged in educative sensemaking. Students who received personalization support were more likely to engage in revision behaviors that were consistent with the development of deep levels of domain knowledge, as informed by comprehension theory (Kintsch, 1998). Students who used CLICK were more likely to engage in revision behaviors that reflected integration of knowledge into prior understanding, whereas students who did not use CLICK were more likely to revise their essay in ways that reflected no transformation or integration of information. Students who received personalization support also were better able to move fluidly across representations (a behavior associated with the intention to enact representation shifts). These results were consistent with the metacognitive benefits found for personalization: Students who received personalization support were better able to revise a representation in meaningful ways.

Third, students who received personalization support were more likely to engage in behaviors that reflected the encoding of new information and the integration of new information with prior knowledge. Students who received personalization support were more likely to correct errors of omission in their essays, providing initial evidence that they were able to enrich their existing knowledge of the domain (Carey, 1991; Chi, 2008). Students who received personalization support also were more likely to make correct inferences when they added new content to their essays, which would be expected if new information had been successfully integrated with prior knowledge in order to develop a situation model level of representation (Kintsch, 1988, 1998). These results are consistent with our findings on metacognitive processes and essay revisions: Personalization appears to support meaningful use of online information during educative sensemaking.

The results from this study demonstrate that personalization technologies and tools may have the potential to help students overcome the educative sensemaking paradox. Despite low initial domain knowledge, the students we studied were able to make use of personalized support to deploy deep metacognitive processes, revise their essays in meaningful ways, and encode and integrate the new information that they encountered. Numerous studies have found that students have great difficulty in working with web-based resources (e.g., Graham & Metaxas, 2003; Stone & Madigan, 2007), especially when they must regulate their own learning and analyze their own understanding (e.g., Azevedo et al., 2004; Azevedo et al., 2008). In light of these difficulties, some authors have called for improved curriculum standards to support students' information literacy (Stone & Madigan, 2007), whereas others have stressed the importance of directly training students on information literacy skills (Graham & Metaxas, 2003) or self-directed learning skills (Azevedo & Cromley, 2004) before they use the Web for educational tasks. Although these techniques may help students achieve success on individual tasks, we argue there is more power and potential in developing robust tools to help students overcome the educative paradox. These tools should train skills in context, helping students to engage in deep, meaningful processes as they learn with online content.

As a flexible set of personalization algorithms that can be embedded in a variety of technologies for educational purposes, technologies like CLICK have the potential to bridge the gap between supporting isolated information literacy skills and holistic, educative sensemaking. The impact of personalization technology on students' general sensemaking processes and metacognitive skills in this study bodes well for transfer to other domains and other educational tasks. Many self-directed learning tasks require learners to be able to analyze and revise their own representations in meaningful ways; these may include concept map generation, presentation development, and creative problem solving.

However, it should be noted that the current research is not without its limitations. Our study focused on a single educational task in a single scientific domain and was conducted in a controlled setting. Although our findings show that personalization technologies can support a number of theoretically important processes, it remains to be seen whether personalization will show robust influence across domains and educational tasks. We have collected preliminary evidence showing positive results with a new domain (weather and climate), but it is too early to determine the boundaries of successful cross-domain transfer. Additional research is needed in which students have the opportunity to interact with their representations for longer periods so that we can determine if the process changes we are seeing result in clear and significant representation differences over time. We also cannot determine the extent to which students benefit from personalization tools in naturalistic settings, where they are free to discontinue tasks and abandon personalization scaffolds at any point. As we continue to explore the impact of personalization tools for sensemaking, our research will need to determine when and why students make use of—or reject—personalization tools and assess subsequent effects on sensemaking processes.

## 7. CONCLUSIONS

For novice learners, educative sensemaking is a daunting task. Students must identify relevant information, analyze its relevance to the task at hand, make connections with prior knowledge, diagnose the adequacy and accuracy of their current representations, integrate new information into their current representations, and re-analyze emerging representations. The results presented here demonstrate consistent, positive impact of personalization on students' metacognitive processes, the depth of their sensemaking behaviors, and the integration of new and prior knowledge.

Although our results were gathered using a prototype instantiation of the CLICK service, there are a wide variety of tasks that could be supported by CLICK's personalization technologies. Our work advances the capabilities of sensemaking tools and stimulates future potential research by providing a flexible set of cognitive personalization algorithms that can be embedded in a wide variety of learning technologies for a large range of educational tasks. Future technologies could support the use cases we have gathered from educators (e.g., concept map development and presentation building), or could embed the algorithms into new tasks such as developing scientific experiments (e.g., collecting and analyzing relevant prior findings) or analyzing visual representations. As we move forward, we are especially interested in studying how we can support multimedia forms of student sensemaking. We see great potential in increasing the richness of students' instantiated representations, for example, by using a combination of visual, audio, and textual information. By not tying our algorithms to a stand-alone prototype, we hope to facilitate the development of new sensemaking tools and to set the stage for future technologies that will push the envelope of educative sensemaking support and discover robust solutions to the sensemaking paradox in self-directed learning.

## NOTES

***Acknowledgments.*** We are grateful to Sebastian de la Chica and Faisal Ahmad for their contributions to this research and for running experimental sessions. We thank Shaw Ketels for his contributions to the student processes/revisions analyses. We also thank our colleagues at Digital Learning Sciences/UCAR for their help with gathering digital library resources.

***Support.*** This material is based in part on work supported by the National Science Foundation under Grant Numbers IIS-0537194, DRL-0835393, and DRL-0835454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

***Authors' Present Addresses.*** Kirsten R. Butcher, University of Utah, Department of Educational Psychology, 1705 East Campus Center Drive, MBH 327, Salt Lake City, UT 84112. E-mail: kirsten.butcher@utah.edu. Tamara Sumner, University of Colorado, Institute of Cognitive Science, Center for Innovation and Creativity, 1777 Exposition Drive, Campus Box 594 UCB, Boulder, CO 80309. E-mail: sumner@colorado.edu.

***HCI Editorial Record***. First manuscript received February 16, 2009. Revision received November 30, 2009. Accepted by Peter Pirolli. Final manuscript received November 30, 2009. — *Editor*

## REFERENCES

Abraham, A., Petre, M., & Sharp, H. (2008, April). *Information seeking: Sensemaking and interactions*. Paper presented at the CHI 2008 Workshop on Sensemaking, Florence, Italy.

Ahmad, F. (2008, December). *Generating conceptually personalized interactions for educational digital libraries using concept maps*. Unpublished doctoral dissertation, Department of Computer Science, University of Colorado at Boulder.

Ahmad, F., de la Chica, S., Butcher, K., Sumner, T., & Martin, J. (2007). Towards automatic conceptual personalization tools. *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2007)*. New York, NY: ACM.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing the model of domain learning. *Journal of Educational Psychology, 87*, 559–575.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *The Journal of the Learning Sciences, 4*, 167–207.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*, 523–535.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30*, 87–111.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research & Development, 56*, 45–72.

Bell, B., & Weaver, R. (1994). Using the programming walkthrough to aid in programming language design. *Software—Practice and Experience, 24*(1), 1–25.

Bell, P., & Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web in KIE. *International Journal of Science Education, 22*, 797–817.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331–350.

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology 83*, 329–345.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*(1), 182–197.

Butcher, K. R., & de la Chica, S. (2010). Supporting student learning with adaptive technology: Personalized conceptual assessment and remediation. In M. Banich & D. Caccamise (Eds.), *Generalization of knowledge: Multidisciplinary perspectives* (pp. 297–330). New York: Taylor & Francis.

Butcher, K. R., de la Chica, S., Ahmad, F., Gu, Q., Sumner, T., & Martin, J. H. (2008). Conceptual customization for learning with multimedia: Developing individual instructional experiences to support science understanding. In R. Zheng (Ed.), *Cognitive effects of multimedia learning* (pp. 260–287). Hershey, PA: IGI Global.

Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 257–291). Hillsdale, NJ: Erlbaum.

Chen, S. Y., & Ford, N. (1998). Modeling user navigation behaviours in a hypermedia-based learning system: An individual differences approach. *Journal of Knowledge Organization, 25*, 67–78.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences, 6*, 271–315.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale, NJ: Erlbaum.

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

Corbett, A. T., McLaughlin, M., & Scarpinatto, K. C. (2000). Modeling student knowledge: Cognitive tutors in high school and college. *User Modeling and User-Adapted Interaction, 10*, 81–108.

de la Chica, S. (2008, December). *Generating conceptual knowledge representations to support students writing scientific explanations*. Unpublished doctoral dissertation, Department of Computer Science, University of Colorado at Boulder.

de la Chica, S., Ahmad, F., Martin, J. H., & Sumner, T. (2008). Pedagogically useful extractive summaries for science education. In D. Scott & H. Uszkoreit (Eds.), *22nd International Conference on Computational Linguistics (COLING 2008)*: Proceedings of the Conference (Vol. 1, pp. 177–184). Stroudsburg, PA: Association for Computational Linguistics (ACL).

de la Chica, S., Ahmad, F., Sumner, T., Martin, J. H., & Butcher, K. R. (2008). Computational foundations for personalizing instruction with digital libraries. *International Journal on Digital Libraries, Special Issue on Digital Libraries and Education, 9*(1), 3–18.

Graham, L., & Metaxas, P. T. (2003). Of course it's true; I saw it on the Internet! *Communications of the ACM, 46*(5), 71–75.

Gu, Q., de la Chica, S., Ahmad, F., Khan, H., & Sumner, T. (2008). Personalization for using digital library resources to support intentional learning. *Proceedings of the 12th European*

*Conference on Research and Advanced Technology for Digital Libraries (ECDL 2008)*. New York, NY: ACM.

Herlocker, J. L., Konstan, J. A., Terveen, L., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *Transactions on Information Systems, 22*(1), 5–53.

Kelley, J. F. (1983, December). *An empirical methodology for writing user-friendly natural language computer applications*. Paper presented at the 1983 SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts.

Kintsch, W. (1986). Learning from text. *Cognition & Instruction, 3*, 87–108.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163–182.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*, 294–303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Klein, G., Moon, B., & Hoffman, R. R. (2006a). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems, 21*(4), 70–73.

Klein, G., Moon, B., & Hoffman, R. R. (2006b). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems, 21*(5), 88–91.

Klein, P. D. (2004). Constructing scientific explanations through writing. *Instructional Science, 32*, 191–231.

Last, D. A., O'Donnell, A. M., & Kelly, A. E. (2001). The effects of prior domain knowledge and goal strength on the use of hypermedia. *Journal of Educational Multimedia and Hypermedia, 7*(1), 51–69.

Liederholm, T., Everson, M. G., Broek, P. v. d., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition and Instruction, 18*, 525–556.

Lin, L.-M., & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*, 345–391.

Lynch, C. (2008). Digital libraries, learning communities, and open education. In T. Iiyoshi & M. S. V. Kumar (Eds.), *Opening up education* (pp. 105–118). Cambridge, MA: MIT Press.

McNeill, K. L. (2008). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education, 93*, 233–268.

Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology, 33*, 270–298.

Murphy, P. K., & Alexander, P. A. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *The Journal of Experimental Education, 70*, 197–214.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. Washington, DC: National Research Council: Commission on Behavioral and Social Sciences and Education, National Academy Press.

Nelson, L., Held, C., Pirolli, P. L., Hong, L., Schiano, D., & Chi, E. H. (2009). With a little help from my friends: Examining the impact of social annotations in sensemaking tasks. *CHI '09: Proceedings of the 27th International Conference on Human Factors in Computing Systems*. New York, NY: ACM.

O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*(1), 71–86.

Quintana, C., Zhang, M., & Krajcik, J. (2005). A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding. *Educational Psychologist, 40*, 235–244.

Russell, D. M., Jeffries, R., & Irani, L. (2008). *Sensemaking for the rest of us*. Paper presented at the CHI 2008 Sensemaking Workshop.

Russell, D. M., Stefik, M. J., Pirolli, P. L., & Card, S. K. (1993). The cost structure of sensemaking. *Proceedings of ACM InterCHI '93*. New York, NY: ACM.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*, 113–125.

Stone, J. A., & Madigan, E. (2007). Inconsistencies and disconnects. *Communications of the ACM, 50*(4), 76–79.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science, 16*, 409–421.

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (pp. 105–140). New York: Wiley and Sons.

White, R. W., Bailey, P., & Chen, L. (2009). Predicting user interests from contextual information. *SIGIR '09: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval.* New York, NY: ACM.

Wiley, J., & Voss, J. F. (1996). The effects of 'playing historian' on learning in history. *Applied Cognitive Psychology, 10*, S63–S72.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*(2), 1–11.

Will, T. C., Srinivasan, A., Bieber, M., Im, I., Oria, V., & Wu, Y.-F. B. (2009). GRE: Hybrid recommendations for NSDL collections. *JCDL '09: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*. New York, NY: ACM.