

Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments

James R. Segedy, John S. Kinnebrew, & Gautam Biswas

Institute for Software Integrated Systems

Department of Electrical Engineering and Computer Science

Vanderbilt University

Nashville, Tennessee, USA

james.r.segedy@vanderbilt.edu

ABSTRACT: Researchers have long recognized the potential benefits of open-ended computer-based learning environments (OELEs) to help students develop self-regulated learning (SRL) behaviours. However, measuring self-regulation in these environments is a difficult task. In this paper, we present our work in developing and evaluating *coherence analysis* (CA), a novel approach to interpreting students' learning behaviours in OELEs. CA focuses on the learner's ability to seek out, interpret, and apply information encountered while working in the OELE. By characterizing behaviours in this manner, CA provides insight into students' open-ended problem-solving strategies as well as the extent to which they understand the nuances of their current learning task. To validate our approach, we applied CA to data from a recent classroom study with *Betty's Brain*. Results demonstrated relationships between CA-derived metrics, prior skill levels, task performance, and learning. Taken together, these results provide insight into students' SRL processes and suggest targets for adaptive scaffolds to support students' development of science understanding and open-ended problem-solving skills.

KEYWORDS: open-ended learning environments, self-regulated learning, learning analytics, coherence analysis, metacognition, computer-based learning environments

1 INTRODUCTION

For several years, researchers have sought to leverage the power of computer-based learning environments (CBLEs) to study aspects of students' self-regulated learning (SRL) behaviours (Sabourin, Shores, Mott, & Lester, 2013). SRL is an active theory of learning that describes how learners are able to set goals, create plans for achieving those goals, continually monitor their progress, and revise their plans when necessary. SRL is a multi-faceted construct: it involves emotional and behavioural control, management of one's learning environment and cognitive resources, perseverance in the face of difficulties, and social interactions to promote effective learning (Zimmerman & Schunk, 2011). For decades, researchers have recognized academic advantages for learners who are self-regulated (e.g., Bransford, Brown, & Cocking, 2000; Butler & Winne, 1995; Zimmerman, 1990), and devising techniques for automatically detecting and supporting students' development of self-regulation in CBLEs is an active area of research (Winters, Greene, & Costich, 2008).

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

SRL is particularly important for students working in *open-ended computer-based learning environments* (OELEs; Clarebout & Elen, 2008; Land, Hannafin, & Oliver, 2012), which provide a learning context and a set of tools for exploring, hypothesizing, and building solutions to authentic and complex problems. Such environments are demanding; they require students to wrestle simultaneously with their emerging understanding of a complex topic (e.g., ecosystems or macroeconomics), develop and utilize problem-solving skills to support their learning, and employ SRL processes for managing the complexity and open-ended nature of the learning task. As such, OELEs can *prepare students for future learning* (Bransford & Schwartz, 1999) by developing their ability to investigate and solve open-ended problems independently.

However, research with OELEs has produced mixed results. While some students with higher levels of prior knowledge and SRL skills show large learning gains as a result of using OELEs, many of their less capable counterparts experience significant confusion and frustration (Azevedo & Witherspoon, 2009; Hacker, Dunlosky, & Graesser, 2009; Kinnebrew, Loretz, & Biswas, 2013). Research examining the activity patterns of those students indicates that they typically make ineffective, suboptimal learning choices when they work independently toward completing open-ended tasks (Kinnebrew et al., 2013; Land, 2000; Mayer, 2004; Sabourin, Mott, & Lester, 2013).

Thus, an important goal of learning analytics research is to develop techniques for studying aspects of SRL and their manifestations in OELEs. These environments can provide a wealth of fine-grained process data, and the inferences made from such data necessarily depend on the analytic lens applied. In this paper, we present our work in developing and evaluating *coherence analysis* (CA), a novel approach to analyzing and interpreting student behaviour in OELEs. CA, an extension of our model-based approach to analyzing learner behaviour in OELEs (Segedy, Biswas, & Sulcer, 2014), focuses on the learner's ability to seek out, interpret, and apply information encountered while working in the OELE. By characterizing behaviours in this manner, CA provides insight into students' open-ended problem-solving strategies, as well as the extent to which they understand the nuances of the learning task they are currently completing.

To validate our approach, we applied CA to data from a recent classroom study with the *Betty's Brain* OELE (Kinnebrew, Segedy, & Biswas, 2014; Leelawong & Biswas, 2008). Results demonstrate the effectiveness of CA in 1) predicting students' task performance and learning gains and 2) identifying common problem-solving approaches among the students in the study. Further, the results demonstrate relationships between CA-derived metrics and students' prior skill levels, offering a potential explanation for students' problem-solving behaviours. Taken together, these results provide insight into students' SRL processes and suggest targets for adaptive scaffolds to support students' development of open-ended problem-solving skills.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

2 OPEN-ENDED LEARNING ENVIRONMENTS AND SELF-REGULATED LEARNING

OELs focus on supporting learners' development of strategies for independently completing open-ended problem-solving tasks. They are typically designed "to support thinking-intensive interactions with limited external direction" (Land, 2000, p. 62) by providing a learning context and a set of tools for learning and problem solving. Some OELs provide explicit goals, while others allow learners to define their own goals. Examples include hypermedia environments (e.g., Bouchet, Harley, Trevors, & Azevedo, 2013), modelling and simulation environments (e.g., Barab, Hay, Barnett, & Keating, 2000; van Joolingen, de Jong, Lazonder, Savelsbergh, & Manlove, 2005; Sengupta, Kinnebrew, Basu, Biswas, & Clark, 2013), and immersive narrative-centred environments (e.g., Clark et al., 2011; Spires, Rowe, Mott, & Lester, 2011). While OELs may vary in the particular sets of tools they provide, they often include tools for 1) *seeking and acquiring* information, 2) *applying* that information to a problem-solving task, and 3) *assessing* the quality of the constructed solution. For example, students may be given the following task:

Use the provided simulation software to investigate which properties relate to the distance that a ball will travel when allowed to roll down a ramp, and then use what you learn to design a ramp suitable for wheelchairs at a local community centre. To test a solution, enter the details of your ramp into the system and press "Test."

OELs place students in a *self-regulatory context* in which they must utilize both cognitive and metacognitive processes to achieve success (Kinnebrew, Segedy, & Biswas, 2014; Segedy et al., 2014). To accomplish this wheelchair task, for example, students must manage their own learning processes in order to 1) use the system's resources to learn about factors important to the design of ramps; 2) apply their knowledge to a problem-solving context by designing a wheelchair ramp; and 3) assess their developing understanding by testing their designs. As part of managing their learning processes, students need to plan their interactions with the system, monitor their progress toward completing their goals, and, when necessary, modify their problem-solving strategies.

2.1 Metacognition and Self-Regulated Learning

Metacognition (Brown, 1975; Flavell, 1976), when applied to learning, is a key component of SRL that describes the ability to reason about, manage, and redirect one's own approach to learning (Whitebread & Cárdenas, 2012). It is often broken down into two sub-components: knowledge and regulation (Schraw, Crippen, & Hartley, 2006; Young & Fry, 2008). Metacognitive knowledge refers to an individual's understanding of her own cognition and strategies for managing that cognition. Metacognitive regulation refers to how metacognitive knowledge is used for creating plans, monitoring and managing the effectiveness of those plans, and then reflecting on the outcome of plan execution in order to refine metacognitive knowledge (Veenman, 2011).

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Metacognitive regulation is often considered a subset of SRL that deals directly with *cognition* without explicitly considering its interactions with emotional or motivational constructs (Whitebread & Cárdenas, 2012). Despite this, models of self-regulation are valuable in depicting key metacognitive processes. For example, Roscoe, Segedy, Sulcer, Jeong, & Biswas describe SRL as containing “multiple and recursive stages incorporating cognitive and metacognitive strategies” (2013, p. 286). Their description of SRL is summarized in Figure 1; it presents SRL as involving phases of orientation and planning, enactment and learning, and reflection and self-assessment.

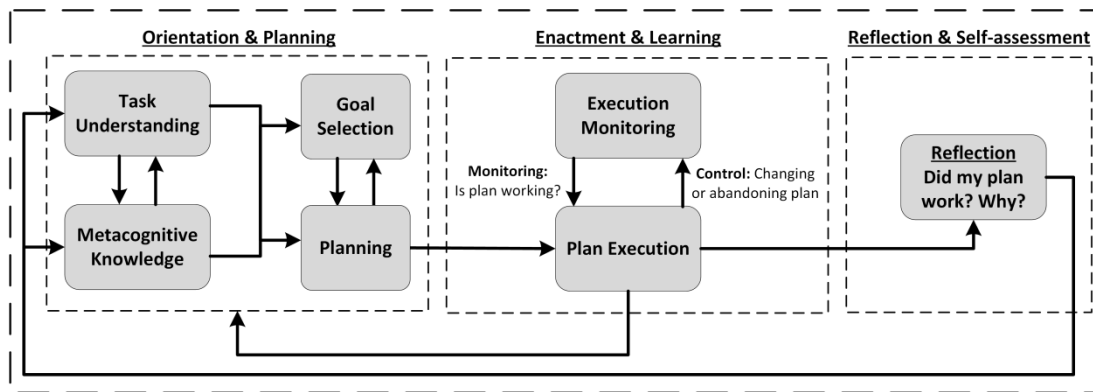


Figure 1. A model of SRL according to the description in Roscoe et al. (2013)

Students may start by orienting themselves to the task and formulating *task understanding* (i.e., an understanding of what the task is). A student’s task understanding is necessarily influenced by her metacognitive knowledge about her own abilities and available strategies for completing the task (Veenman, 2013). Together, these two sources of information, task understanding and metacognitive knowledge, provide a foundation that, in conjunction with other student attributes such as self-efficacy, governs students’ subsequent goal-setting and planning processes. Once a plan has been formulated, students begin executing it. As they carry out the activities specified in their plans, students may exercise metacognitive monitoring as they consciously evaluate the effectiveness of their plans and the success of the activities they are engaging in. The result of these monitoring processes may lead students to exercise metacognitive control by modifying or abandoning their plan as they execute it. Once a plan has been completed or abandoned, students may engage in reflection as they analyze the effectiveness of their plans and their planning processes. Such reflection may lead students to revise their metacognitive knowledge and task understanding.

2.2 Real-Time Measurement of Metacognition and Self-Regulated Learning

Measuring students’ self-regulation and metacognitive behaviour in real time is a difficult task; it requires developing systematic analysis techniques for detecting aspects of goal setting, planning, monitoring, and reflection in the context of the learning environment. In OELEs, such diagnoses involve identifying and assessing learners’ cognitive skill proficiency, interpreting their action sequences in

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

terms of their goals and the learning strategies they apply to achieve those goals, and evaluating their success in accomplishing their current tasks. The open-ended nature of OELEs further exacerbates the measurement problem; since the environments are learner-centred, they typically do not restrict the approaches that learners take to solving their problems. Thus, interpreting and assessing students' learning behaviours is inherently complex; they may simultaneously pursue, modify, and abandon any of a large number of both short-term and long-term approaches to completing their tasks.

Despite this complexity, researchers have developed several approaches to measuring aspects of self-regulation and metacognition in OELEs. For example, *MetaTutor* (Bouchet et al., 2013) adopts a very direct approach; it provides interface features through which students can externalize their SRL processes. By selecting an option from the *SRL Palette*, students indicate that they would like to, for example, judge their learning or activate their prior knowledge. To ensure that these features are used regularly, the system includes pedagogical agents that prompt students to engage in SRL processes through these features. This allows the system to capture students' SRL processes directly without having to make inferences based solely on their activities in the system.

Another approach employed in several OELEs involves developing a predictive, data-driven model for diagnosing constructs related to SRL in real time (e.g., engagement, frustration, or confusion). In some OELEs, such as *Crystal Island* (Sabourin, Shores, Mott, & Lester [2013]) and *EcoMUVE* (Baker, Ocumpaugh, Gowda, Kamarainen, & Metcalf, 2014), models have been created by first employing human coding to label students' log data with aspects of SRL and then using that labelled data to construct predictive models. For example, Sabourin, Mott, & Lester (2013) asked students to author "status updates" at regular intervals while using *Crystal Island*. These updates were later coded according to whether or not they included evaluations of the student's progress toward a goal, and this coded data, along with other features related to the student and her behaviours in the system, was used to build a predictive model of good versus poor self-regulation.

In other OELEs, researchers have developed theory-driven models of SRL and embedded those models into learning environments. For example, *EcoLab* (Luckin & du Boulay, 1999; Luckin & Hammerton, 2002) measures students' metacognitive awareness of their own ability by comparing the system's assessment of students' ability levels with the difficulty of the activities they choose to pursue. Should students choose activities that are too easy or too difficult, the system decreases its confidence in the student's metacognitive awareness and then prompts them to reconsider their choice (e.g., "*You should try a more difficult activity*"). As another example, Snow, Jackson, & McNamara (2014) measured the order and stability of students' behaviour patterns as they used iSTART-ME, a computer-based learning environment for helping students improve their science comprehension. In their model, lower levels of the information theoretic measure, *Shannon Entropy* (Coifman & Wickerhauser, 1992) were interpreted as indicative of ordered and self-regulated behaviours.

The approach presented in this paper is similar to this second set of environments: we have developed a

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

novel theory-driven approach to modelling learning behaviours in OELEs, called *coherence analysis* (CA), and applied that approach to the interpretation of log data from *Betty's Brain*. However, rather than focusing on a specific aspect of SRL (e.g., awareness of one's own ability), CA focuses on students' ability to seek for, interpret, and apply information encountered while working in the OELE. In doing so, CA models aspects of students' problem-solving skills and metacognitive abilities simultaneously. The approach is designed to be general, and should apply to OELEs beyond *Betty's Brain*, allowing researchers to study the coherence aspect of SRL in multiple contexts. The next section presents a brief overview of *Betty's Brain*, and the one following presents our coherence analysis approach.

3 BETTY'S BRAIN

Betty's Brain (Kinnebrew, Segedy, & Biswas, 2014; Leelawong & Biswas, 2008), shown in Figure 2, presents the task of teaching a virtual agent, Betty, about a science phenomenon (e.g., climate change) by constructing a causal map that represents that phenomenon as a set of entities connected by directed links representing causal relationships. Once taught, Betty can use the map to answer causal questions and explain those answers by reasoning through chains of links (Leelawong & Biswas, 2008). The goal for students using *Betty's Brain* is to construct a causal map that matches an expert model of the domain.

As an OELE, *Betty's Brain* includes tools for acquiring information, applying that information to a problem-solving context, and assessing the quality of the constructed solution. Students acquire domain knowledge by reading hypertext resources that include descriptions of scientific processes (e.g., shivering) and information pertaining to each concept that appears in the expert map (e.g., friction). As students read, they need to identify causal relations, such as "*skeletal muscle contractions create friction in the body.*" Students can then apply the learned information by adding the two entities to the causal map and creating the causal link between them (which "teaches" the information to Betty). In *Betty's Brain*, learners are provided with the list of concepts, and link definitions are limited to the qualitative options of increase (+) and decrease (-). Students can also add textual descriptions to each link.

Learners can assess their causal map by asking Betty to answer questions (using a causal question template) and explain her answers. To answer questions, Betty applies qualitative reasoning methods to the causal map (e.g., *the question said that the hypothalamus response increases. This causes skin contraction to increase. The increase in skin contraction causes...*). After Betty answers a question, learners can ask Mr. Davis, another pedagogical agent that serves as the student's mentor, to evaluate her answer. If Betty's answer and explanation match the expert model (i.e., in answering the question, both maps utilize the same causal links), then Betty's answer is correct. Note that a link's textual description is not considered during this comparison.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

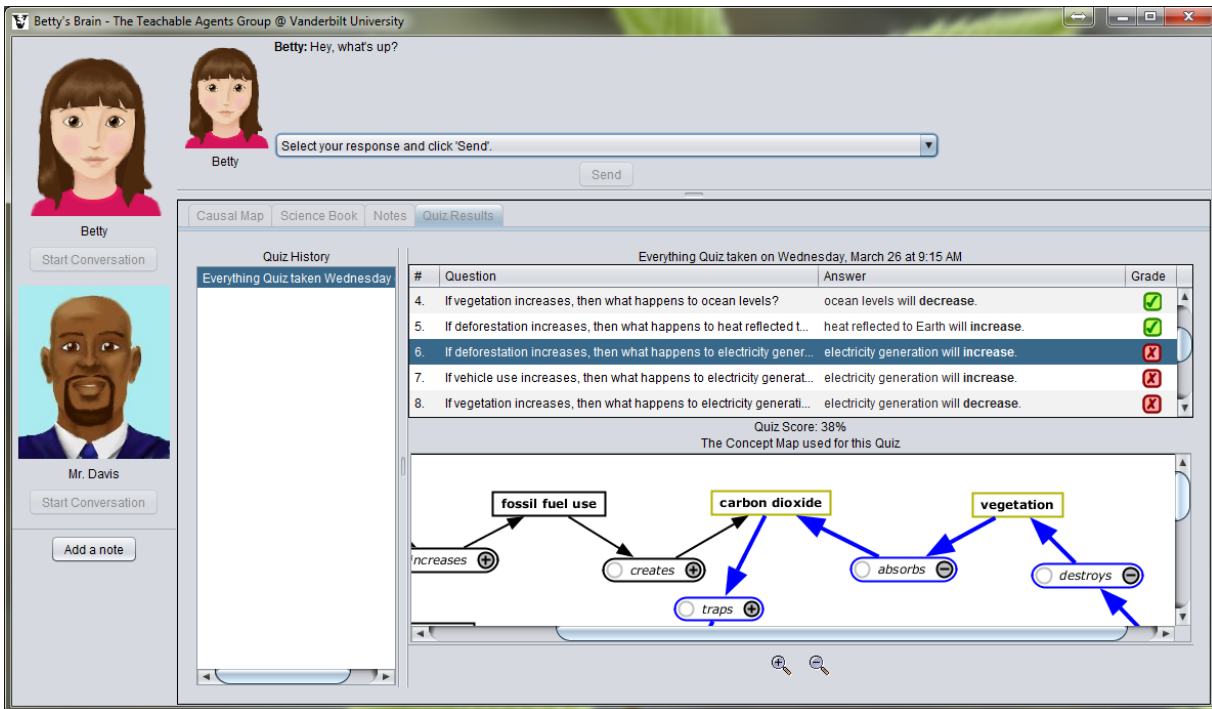


Figure 2. Betty's Brain showing the quiz interface

Learners can also have Betty take *quizzes*, which are a set of questions that can be answered by chaining together causal links in the map. An example quiz question is “If vehicle use increases what happens to heat reflected back to earth?” The question can be answered by following a chain of links from the concept *vehicle use* to the concept *heat reflected back to earth*, to derive the answer “heat reflected back to earth will increase.” Quiz questions are selected dynamically by comparing Betty’s current causal map to the expert map such that a portion of the chosen questions, in proportion to the completeness of the current map, will be answered correctly by Betty. The rest of her quiz answers will be incorrect or incomplete, helping the student identify areas for correction or further exploration. When Betty answers a quiz question correctly, students know that the links she used to answer that question are correct. Similarly, when Betty answers a question incorrectly, students know that at least one of the links she used to answer the question is incorrect. To help students keep track of correct links, the system allows students to annotate causal links as being correct.

4 COHERENCE ANALYSIS

This section describes our novel *Coherence Analysis* (CA) approach to learner modelling in OELEs. To develop this approach, we first performed a task-driven analysis of *Betty’s Brain* (similar to cognitive task analysis; Chipman, Schraagen, & Shalin, 2000) to derive 1) the primary tasks that students should be able to complete to succeed in an OELE, and 2) the processes students must execute to complete those tasks. The result of this analysis is presented in the following section on “Cognitive and Metacognitive Problem-solving Processes in OELEs”; the CA approach is presented in “Modelling Learner Behaviours

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

with Coherence Analysis” below.

4.1 Cognitive and Metacognitive Problem-solving Processes in OELEs

As discussed above, learners working in OELEs need to access and interpret information, apply that information to constructing their problem solutions, and assess the quality of the constructed solutions using assessments provided in the system (Clarebout & Elen, 2008; Land et al., 2012; Land, 2000). These tasks and their more specific implementations in *Betty’s Brain* have been incorporated into a task model (Figure 3) that specifies the tasks important for achieving success in *Betty’s Brain*. The highest level of the model identifies the three broad classes of OELE tasks related to 1) *information seeking and acquisition*, 2) *solution construction and refinement*, and 3) *solution assessment*. Each of these task categories is further broken down into three levels that represent 1) general task descriptions common across all OELEs (according to the definition of OELE discussed above); 2) *Betty’s Brain* specific instantiations of these tasks; and 3) interface features in *Betty’s Brain* through which students can accomplish their tasks.

The directed links in the task model represent dependency relations. Information seeking and acquisition depends on one’s ability to identify, evaluate the relevance of, and interpret information in the context of the overall task. Solution construction and refinement tasks depend on one’s ability to apply information gained both by conducting information seeking tasks and by analyzing the solution assessment results. Finally, solution assessment tasks depend on one’s ability to interpret the results of solution assessments as actionable information that can be used to refine the solution in progress. In order to accomplish these general tasks in *Betty’s Brain*, students must understand how to perform the related *Betty’s Brain* specific tasks by utilizing the system’s interface features.

Identifying and evaluating the relevance of information describes the processes students employ as they observe, operate on, and make sense of the information presented in an OELE’s information acquisition tools (Land, 2000; Quintana et al., 2004). Productively employing these processes requires an understanding of how to identify critical information and interpret it correctly. While learning in *Betty’s Brain*, students need to identify sections of the hypertext resources that describe causal relations between entities in the problem domain. They must then correctly interpret those relations in order to create an accurate mental model of the science phenomena involved. Such processes can be difficult for learners; they may not have a firm grasp of causal reasoning mechanisms and the corresponding representational structures, or they may have difficulty in extracting the correct causal relations from the nuanced, technical writing style typical of science texts (McNamara, 2004). Further complications exist when the information contained in the resources conflicts with or challenges learners’ prior inaccurate understandings of the problem domain. Land (2000) explains that in such situations, learners are resistant to restructuring their knowledge; instead, they often misinterpret the information in a way that supports their original conceptions.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

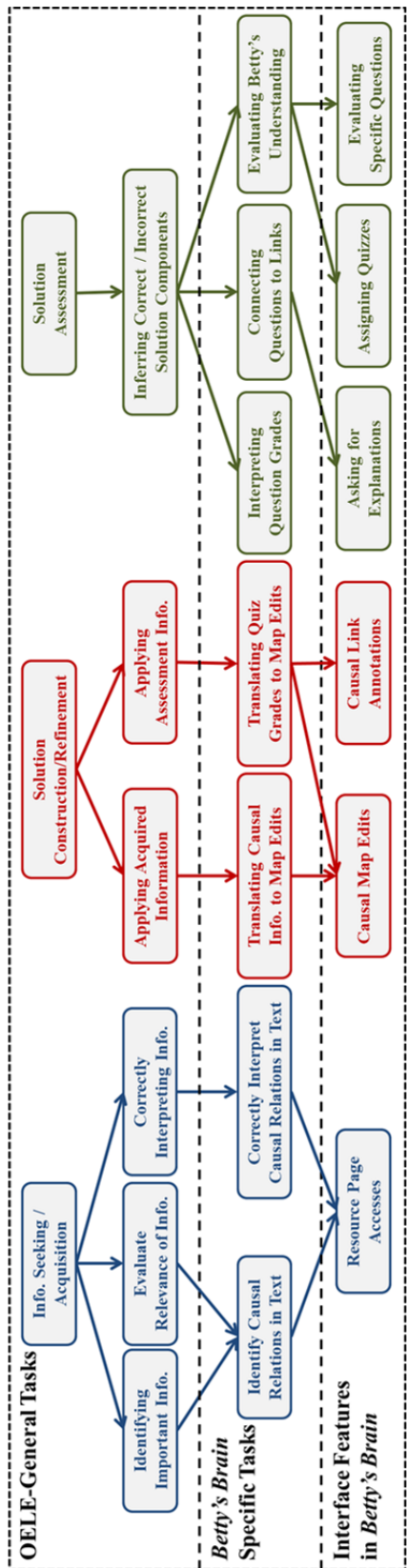


Figure 3. Task model for Betty's Brain OELE

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

When constructing problem solutions, learners utilize their developing understanding of the problem domain to make decisions about how to construct solutions. Productively employing these processes requires an understanding of 1) the structure of problem solutions; 2) the tools available for constructing solutions; and 3) methods for translating one's own understanding of the problem domain and solution requirements into explicit plans for solution construction in the OELE. In *Betty's Brain*, solutions take the form of a visual causal map, and accurately constructing such a map requires *representational fluency* (Suh & Moyer, 2007). Students must be able to convert causal information between and among the system's hypertext resources, their internal knowledge structures, and the causal map representation. Students unfamiliar with causal structures or how to represent knowledge using them will most likely struggle to succeed in completing the *Betty's Brain* learning task (Segedy, Kinnebrew, & Biswas, 2013; Roscoe et al., 2013).

Assessing the quality of constructed solutions describes the processes students employ as they submit their solutions to automated tests within the system and interpret the resulting feedback. In *Betty's Brain*, learners receive feedback in the form of Betty's quiz results: a list of questions that are either addressed appropriately by the model (i.e., Betty can answer these questions correctly), not addressed by the model (i.e., Betty cannot answer these questions), or addressed incorrectly by the model (i.e., Betty generates an incorrect answer to these questions). Learners are expected to use this information to determine which of their causal links are correct, which are incorrect, and what information is missing. This requires understanding how to interpret question grades, identify the causal links used to generate an answer, and evaluate the assessment information obtained via quizzes and question explanations. If students do not understand the relationship between a question, its quiz grade, and the links used to answer it, then they will most likely experience difficulty in obtaining meaningful information from quizzes.

The task model, along with the model of SRL presented in Figure 1, identifies and draws connections among the cognitive and metacognitive processes critical for learning in OELEs. Students need to leverage their metacognitive knowledge and task understanding in order to select intermediate goals for completing their tasks and then create plans for coordinating their use of skills and strategies in service of achieving those goals. Creating these plans requires understanding the purposes of, and relationships among, the tasks identified in the task model. Effective plans will utilize information gained from both information acquisition and solution assessment activities in order to build and refine a causal map that more closely approximates the expected solution. Because students are likely to make mistakes in constructing their solutions, they need to understand how to utilize the results of solution assessments to direct their thinking as they reflect on the sources of their errors.

4.2 Modelling Learner Behaviours with Coherence Analysis

The Coherence Analysis (CA) approach analyzes learners' behaviours by combining information from sequences of student actions to produce measures of *action coherence*. CA interprets students'

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

behaviours in terms of the information they encounter in the system and whether or not this information is utilized during subsequent actions. When students take actions that put them into contact with information that can help them improve their current solution, they have *generated potential* that should *motivate future actions*. The assumption is that if students can recognize relevant information in the resources and quiz results, then they should act on that information. If they do not act on information that they encountered previously, CA assumes that they did not recognize or understand the relevance of that information. This may stem from incomplete or incorrect understanding of the science topic, the learning task, and/or strategies for completing the learning task. Additionally, when students edit their map when they have not encountered any information that could motivate that edit, CA assumes that they are guessing¹. These two notions come together in the definition of action coherence:

*Two ordered actions ($x \rightarrow y$) taken by a student in an OELE are **action coherent** if the second action, y , is based on information generated by the first action, x . In this case, x provides **support** for y , and y is **supported** by x . Should a learner execute x without subsequently executing y , the learner has created **unused potential** in relation to y . Note that actions x and y need not be consecutive.*

The task model (Figure 3) implies two critical coherence relations: 1) applying information acquired from the hypertext resources to editing the map; and 2) applying inferred link correctness information (as obtained via quizzes) to editing the map. More specifically, an information seeking action (e.g., reading about a causal relationship) may generate support for a future solution construction action (e.g., adding that causal relationship to the map). Similarly, a solution construction action can be supported by a solution assessment action. An example of the latter situation occurs in *Betty's Brain* when a student deletes a causal link from their map after observing that Betty used that link to generate an incorrect answer to a quiz question.

CA assumes that learners with higher levels of action coherence possess stronger metacognitive knowledge and task understanding. Thus, these learners will perform a larger proportion of supported actions and take advantage of a larger proportion of the potential that their actions generate. In the analyses presented in this paper, we incorporated the following coherence relations:

- Accessing a resource page that discusses two concepts *provides support for* adding, removing, or editing a causal link that connects those concepts.
- Viewing assessment information (usually quiz results) that proves that a specific causal link is correct *provides support for* adding that causal link to the map (if not present) and annotating it

¹ In reality, students may be applying their prior knowledge; however, CA assumes that since students are typically wrestling with their emerging understanding of the domain, they should verify their prior knowledge before attempting a solution.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

- as being correct (if not annotated)².
- Viewing assessment information (usually quiz results) that proves that a specific causal link is incorrect *provides support* for deleting it from the map (if present).³

Action coherence metrics measure whether or not learners' actions take advantage of previously encountered information. To measure whether or not a learner's actions *contradict* the information generated during previous activities, CA also incorporates measures of *action incoherence*:

*Two ordered actions ($x \rightarrow y$) taken by a student in an OELE are **action incoherent** if the second action, y , is action coherent with the negation of information generated by the first action, x . In this case, x provides **negative support** for y , and y is **contradicted** by x .*

CA assumes that learners with higher levels of incoherence among their actions possess a weaker understanding of the science domain and the relations between different concepts in the domain. For example, when students have a misconception, they may add an incorrect link to their map due to their incorrect prior knowledge (Segedy, Kinnebrew, & Biswas, 2011). During solution assessment, they may obtain evidence that the link is incorrect and then delete it. However, in deleting the link, they may not restructure their own understanding of the problem domain, and, as a consequence, their established misconception may lead them to add the same incorrect link to the map at a later point in time. It is important to note that while incoherence is the natural complement to coherence in our analysis framework, space limitations compel us to focus on the primary (coherence-based) CA metrics in this paper, leaving a detailed analysis of students' action incoherence for future work.

Low levels of action coherence (and high levels of action incoherence) may indicate that learners do not possess sufficient task understanding or all of the metacognitive knowledge necessary for generating coherent plans. However, these CA-derived metrics are general measures of performance, and learners may exhibit low levels of action coherence for a variety of reasons. They may be struggling with 1) the task understanding and metacognitive knowledge underlying the coherence relations, 2) the related cognitive processes, and/or 3) their understanding of the domain content.

Our hypotheses in developing CA were as follows:

1. Students' CA-derived metrics would predict their learning and success in teaching Betty;
2. Students' prior levels of task understanding would predict their CA-derived metrics while using *Betty's Brain*.

² A quiz can only prove that a link is correct when it is already on Betty's map; however, a student can view an old quiz after deleting a link proven correct by that old quiz. In this case, viewing the old quiz would provide support for adding that link back to the map.

³ A student can view an old quiz that proves a link is incorrect even if that link is not currently on their map.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

To explore these hypotheses, we applied CA to a recent classroom study of students using *Betty's Brain*. This study is presented next.

5 EXPERIMENTAL STUDY WITH BETTY'S BRAIN

The goals of this experimental study were to test the two hypotheses listed above. In addition, we sought to investigate whether or not CA-derived metrics would reveal common problem-solving approaches as a set of distinct behaviour profiles from the study data. The data presented in this paper comes from a larger experiment with *Betty's Brain* in which students completed two instructional units: climate change and human thermoregulation. During the first unit on climate change, students received different types of support from Mr. Davis. While analyses on this data have revealed statistically significant learning gains overall, they failed to reveal any significant effects of the type of support from Mr. Davis on students' learning and performance. Therefore, this paper focuses on the behaviour of all students, irrespective of the type of support received in the first unit, as they worked on the second unit, human thermoregulation, where students did not receive any support from the agents.

5.1 Participants

Ninety-nine 6th grade students from four mid-Tennessee science classrooms participated in the study. The participating school was an academic magnet school with competitive admission requirements. To enrol in this school, students need to pass all of their classes and achieve an average grade of B+ during the previous academic year. Demographic data for individual students were not released; however, the participants were typical of the school environment. Of the school's 701 students, 7.8% identified as Asian, 26.2% as Black, 4.0% as Hispanic, and 61.8% as White. None of the students was eligible for English as a Second Language programs, 1.4% of the students participated in special education programs, and 26.8% of the students were served by the Free and Reduced Price Lunch program. None of our study participants was enrolled in special education programs. One student was excused from the study due to an unrelated injury, therefore, the sample included data from 98 students.

5.2 Topic Unit and Text Resources

Students used *Betty's Brain* to learn about human thermoregulation when exposed to cold temperatures. The expert map, shown in Figure 4, contained 13 concepts and 15 links representing cold detection (cold temperatures, heat loss, body temperature, cold detection, hypothalamus response) and three bodily responses to cold: goose bumps (skin contraction, raised skin hairs, warm air near skin, heat loss), vasoconstriction (blood vessel constriction, blood flow to the skin, heat loss), and shivering (skeletal muscle contractions, friction in the body, heat in the body). The resources were organized into two introductory pages discussing the nervous system and homeostasis, one page discussing cold detection, and three pages discussing the three bodily responses to cold temperatures, one response per page. Additionally, a dictionary section discussed some of the concepts, one per page. The text was 15 pages (1,974 words) with a Flesch-Kincaid reading grade level of 9.0.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

5.3 Betty's Brain Interface and Features

The version of *Betty's Brain* used in this study was similar to the version presented above and illustrated in Figure 2. Students had access to hypertext resources, causal map editing tools, and the quiz feature. They could also ask Betty to answer questions and explain her answers, and they could ask Mr. Davis to grade Betty's answer to a specific question. However, Mr. Davis avoided grading answers where Betty used a single link to generate the answer. This was done to prevent students from gaming the system (Baker et al., 2006) by repeatedly adding a link to Betty's map and asking Mr. Davis if Betty's answer to a question using only that link was correct. If students were unsure of what to do, they could ask Mr. Davis to explain concepts important for success in *Betty's Brain* (e.g., *what are cause and effect relationships, and how do I find them while reading?*).

In addition, all students had access to a *Teacher's Guide* and a second set of hypertext resources that explained skills and strategies for seeking information, constructing the causal map, and assessing the causal map. For information seeking, the guide discussed how to identify causal links in text passages that use different presentation formats. For example, some passages present a causal link by describing what happens when the source concept *decreases* (e.g., "When cold detection decreases, the hypothalamus response also decreases"). For constructing the causal map, the guide explained how to use the causal map interface to add, edit, and remove concepts and links. It also explained the mechanics of causal reasoning (e.g., how to use a causal map to answer questions). For assessing the causal map, the guide discussed strategies for using quizzes, explanations, and Mr. Davis's answer evaluations to identify correct and incorrect links on Betty's map. In total, the guide contained 31 pages (6,247 words) with a Flesch-Kincaid reading grade level of 6.6.

5.4 Learning Assessments

Learning was assessed using a pre-test–post-test design with two parts: a set of computer-based exercises and a set of paper-and-pencil questions. The computer-based test consisted of 20 causal reasoning items, 10 causal link extraction items, and 14 quiz evaluation items designed to test students' understanding of the skills discussed in Section 4.1. The written test consisted of six multiple-choice science content items and four short answer science content questions. Further details on these assessments are available in Segedy (2014), Appendix C.

5.4.1 Causal Reasoning Items

Causal reasoning items (n=20) presented students with an abstract causal map (i.e., concepts were named *A*, *B*, etc...) and asked students to reason with the map to answer questions (e.g., "If concept *A* increases, what will happen to concept *B*?"). Each problem presented students with four possible choices: *B* will increase; *B* will decrease; *B* will not be affected; and it depends on which causal relations are stronger. Students were awarded one point for each question they answered correctly. An abstract causal map from this assessment is shown in Figure 5. Two causal reasoning items associated with this

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

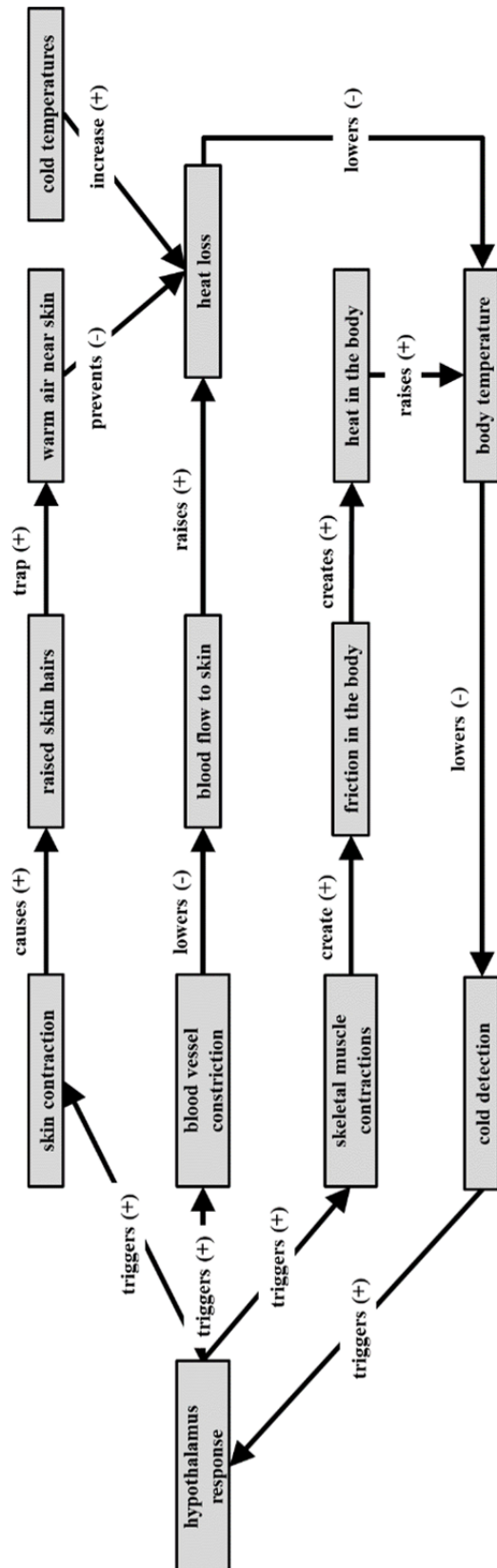


Figure 4: Thermoregulation expert map

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

map were as follows: 1) *If N increases, then what will happen to P?*; and 2) *If N decreases, then what will happen to P?* The causal reasoning items were found to be reliable on both the pre-test ($\alpha = 0.74$) and the post-test ($\alpha = 0.81$).

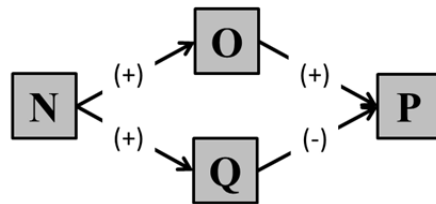


Figure 5. An example causal reasoning item.

5.4.2 Causal Link Extraction Items

Causal link extraction items (n=10) presented students with a text passage discussing the relationship between two abstract entities “Ticks” and “Tacks” (e.g., “*Tacks increase when Ticks decrease*”), and they were asked to choose the corresponding causal link described by that passage from the following choices: *Tacks increase Ticks*, *Tacks decrease Ticks*, *Ticks increase Tacks*, and *Ticks decrease Tacks*. Students were awarded one point for each correct answer. The ten causal link extraction items and their correct answers are included in Table 1. These items were found to be reliable on the pre-test ($\alpha = 0.71$) and the post-test ($\alpha = 0.76$).

Table 1. Causal link extraction items and their correct answers

| | Text Passage | Correct Causal Link |
|-----|---|-----------------------|
| 1. | Tacks increase Ticks. | Tacks increase Ticks. |
| 2. | A decrease in Ticks decreases Tacks. | Ticks increase Tacks. |
| 3. | Tacks are decreased by Ticks. | Ticks decrease Tacks. |
| 4. | Ticks are decreased by a decrease in Tacks. | Tacks increase Ticks. |
| 5. | When Ticks increase, Tacks increase too. | Ticks increase Tacks. |
| 6. | When Tacks decrease, Ticks increase. | Tacks decrease Ticks. |
| 7. | When Tacks increase, Ticks decrease. | Tacks decrease Ticks. |
| 8. | Ticks decrease when Tacks increase. | Tacks decrease Ticks. |
| 9. | Tacks decrease when Ticks decrease. | Ticks increase Tacks. |
| 10. | Ticks are increased when Tacks increase. | Tacks increase Ticks. |

5.4.3 Quiz Evaluation Items

Quiz evaluation items (n=14) presented students with a quiz whose questions, answers, and grades were linked to an abstract causal map (see Figure 6). Students received one point for every problem in which they correctly annotated links as correct or incorrect according to the information in the quiz⁴. These

⁴ See the “Betty’s Brain” section for more information about how Betty’s answers can be used to infer correct and incorrect causal links.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

items were found to be somewhat reliable on both the pre-test ($\alpha = 0.62$) and the post-test ($\alpha = 0.68$).

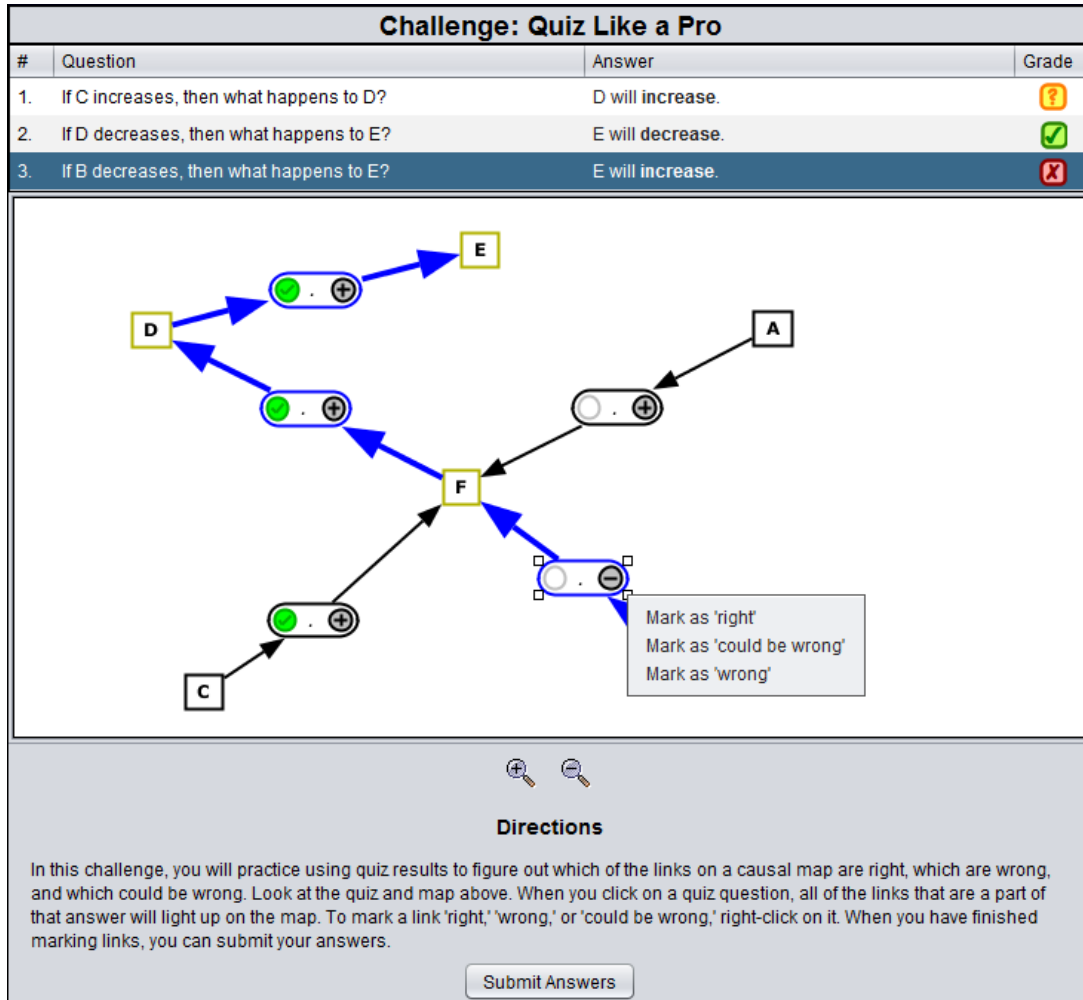


Figure 6. The quiz evaluation problem interface

5.4.4 Science Content Multiple-Choice Items

Science content multiple-choice items (n=6), each with four choices, tested students’ knowledge of concepts, processes, and causal relations among concepts in the thermoregulation domain. These items are shown in Table 2.

5.4.5 Science Content Short Answer Items

Short answer items asked students to combine the causal relations among concepts to explain how the human body detects and responds to cold temperatures. The items are listed in Table 3. These items were coded by identifying the chain of causal relationships in learners’ answers, and these chains were then scored by comparing them to the chain of causal relationships used to derive the answer from the expert map. One point was awarded for each causal relationship in the student’s answer that was the same as or closely related to a relation specified in the expert map. For example, to answer question 1

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

correctly, students needed to note that skin contraction raises hairs near the skin (1 point), that these raised hairs trap warm air and keep it near the skin (1 point), and that this warm air near the skin reduces the rate at which heat is lost from the body (1 point). The maximum combined score for these questions was 11. Two coders independently scored five of the pre- and post-tests with over 85% agreement, at which point one of the coders individually coded the remaining answers and computed the scores.

Table 2. Science content multiple-choice items.

| | Item |
|----|---|
| 1. | What is thermoregulation? |
| 2. | How does the hypothalamus regulate body temperature when the body gets too cold? |
| 3. | How does shivering help regulate body temperature in cold temperatures? |
| 4. | How do blood vessels change when the body is exposed to cold temperatures? |
| 5. | How do raised skin hairs affect body heat? |
| 6. | When a person drinks alcohol, their blood vessels become wider. How would drinking alcohol affect a person outside on a cold day? |

Table 3. Science content short answer items.

| | Items |
|----|--|
| 1. | Explain, step-by-step, how skin contraction reduces heat loss from the body. |
| 2. | Explain, step-by-step, how skeletal muscle contractions increase body temperature. |
| 3. | Explain, step-by-step, how blood vessel constriction decreases heat loss from the body. |
| 4. | Explain, step-by-step, how cold temperatures cause a hypothalamus response in the brain. |

5.5 Log File Analysis

This version of *Betty’s Brain* generated *event logs* that captured every *action* taken by the student, Betty, and Mr. Davis. A logged action corresponds to an atomic *expression of intent*, such as deleting a causal link or asking Betty to take a quiz. In addition, the logs contain information on every *view* that was displayed when the system was running. A logged view captures the information visible to a user during a specific time interval. For example, a view is created each time a page of the hypertext resources is visible. Unlike actions, which are distinct and orderable, views can overlap each other and span across multiple actions.

The log files provided the information required to calculate a measure of task performance for each student. By tracking the evolution of a student’s causal map, we could compute how the student’s causal map score changed over time. The *map score* at any point in time is calculated as the number of correct links (i.e., links that appear in the expert map) minus the number of incorrect links in the student’s map. A student’s *best map score* was computed as the highest map score they attained during

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

the intervention⁵. The log files also served as input to CA, which automatically calculated the following metrics for each student:

1. *Edit Frequency*: the number of causal link edits and annotations made by the student divided by number of minutes that the student was logged onto the system.
2. *Unsupported edit percentage*: the percentage of causal link edits and annotations not supported by any of the previous views that occurred within a five-minute window of the edit.
3. *Information viewing time*: the amount of time spent viewing either the science resource pages or Betty's graded answers. *Information viewing percentage* is the percentage of the student's time on the system classified as *information viewing time*.
4. *Potential generation time*: the amount of *information viewing time* spent viewing information that could support causal map edits that would improve the map score. To calculate this, we annotated each hypertext resource page with information about the concepts and links discussed on that page. *Potential generation percentage* is the percentage of *information viewing time* classified as *potential generation time*.
5. *Used potential time*: the amount of *potential generation time* associated with views that both occur within a prior five-minute window and also support an ensuing causal map edit. *Used potential percentage* is the percentage of *potential generation time* classified as *used potential time*.

Metrics one and two capture the quantity and quality of a student's causal link edits and annotations, where supported edits and annotations are considered to be of higher quality. Metrics three, four, and five capture the quantity and quality of the student's time viewing either the resources or Betty's graded answers. These metrics speak to the student's ability to *seek* and *identify* information that may help her build or refine her map (potential generation percentage) and then *utilize* information from those pages in future map-editing activities (used potential percentage). In these analyses, a page view generated potential and supported edits only if it lasted at least 10 seconds. Similarly, students had to view quiz results for at least 2 seconds. These cut-offs helped filter out irrelevant actions (e.g., rapidly flipping through the resource pages without reading them).

We also calculated a measure of *disengaged time*, which is defined as the sum of all periods of time, at least 5 minutes long, during which the student neither 1) viewed a source of information (i.e., science resources and quiz results) for at least 30 seconds; nor 2) added, changed, deleted, or annotated concepts or links. This metric represents periods of time during which the learner is *not measurably engaged* with the system. *Disengaged percentage* is the percentage of the student's time on the system classified as disengaged time.

⁵ Not every student's final map was the best map she had created. For example, a student might decide to delete her entire map and start over near the end of the intervention.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

As a complement to the CA-derived metrics, we employed an information-theoretic differential sequence mining approach (Kinnebrew, Mack, Biswas, & Chang, 2014) to analyze students' action sequences further. This allowed us to identify sequential action patterns that best differentiated groups of students defined by the CA-derived metrics. The analysis considered the following student actions: 1) reading resource pages; 2) adding, removing, or editing causal links in the map (further distinguished by whether or not the edit improved the map score); 3) asking Betty to answer causal questions; 4) having Betty take quizzes; 5) asking Betty to explain her answer to a question; 6) creating, editing, and viewing notes; and 7) annotating links to keep track of their correctness. The derivation of action definitions from raw activity logs is discussed further in (Kinnebrew et al., 2013).

5.6 Procedure

The study was conducted over a period of approximately 6 weeks. At the beginning of the study, the first author spent 20 minutes introducing students to the causal reasoning methods used in the system. In particular, this lesson focused on understanding how to interpret and reason with both individual links and chains of links (i.e., a sequence of one or more links). Students then spent two weeks completing the climate change unit, which are not reported in this paper. At the beginning of the climate change unit, students were introduced to the software by Mr. Davis, who explained the task goal (i.e., teach Betty the correct causal map) and each of the *Betty's Brain* system features. As Mr. Davis explained a feature, he required students to use the feature in a specific way. For example, Mr. Davis asked students to add the concept "wolves" to their maps and he did not let them proceed until they had followed his instructions. Students practiced adding and deleting concepts and links, annotating links, asking Betty to take a quiz, and viewing Betty's quiz results. Mr. Davis also explained the importance of these features in successfully completing the *Betty's Brain* task. For example, he noted that students needed to identify relevant causal relations as they read the science resources and then teach these relations to Betty.

A two-week break separated the climate change and thermoregulation units. At the start of the thermoregulation unit, students spent two days completing the thermoregulation pre-tests. Students then spent four class periods using *Betty's Brain* to learn about thermoregulation. They completed the thermoregulation post-test approximately 1.5 weeks after the pre-test.

6 RESULTS

6.1 Learning and Performance Results

Table 4 summarizes the means (and standard deviations) of the students' pre-test and post-test scores, significant tests for gains, and a measure of effect size (Cohen's d). Overall, students exhibited strong gains on science multiple choice ($d = 1.04$) and short answer items ($d = 1.55$), suggesting that *Betty's Brain* facilitated students' ability to recognize and reason with relationships and definitions important for understanding thermoregulation. Conversely, students did not show statistically significant gains on

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

the remaining three measures. However, during the first unit, students did exhibit statistically significant gains on causal reasoning ($p < 0.01, d = 0.27$), causal link extraction ($p < 0.01, d = 0.72$), and quiz evaluation items ($p < 0.01, d = 0.75$), suggesting that *Betty’s Brain* facilitated students’ ability to reason with causal maps, identify links in text passages, and interpret Betty’s quiz results during the first unit.

Table 4. Means (and standard deviations) of assessment test scores

| Measure | Maximum | Pre-test | Post-test | <i>t</i> | <i>p</i> | Cohen’s <i>d</i> |
|-------------------------|---------|--------------|--------------|----------|----------|------------------|
| Science Multiple-Choice | 6 | 2.46 (1.07) | 3.90 (1.63) | 7.87 | 0.001 | 1.04 |
| Science Short Answer | 11 | 1.09 (1.14) | 4.63 (2.55) | 13.83 | 0.001 | 1.55 |
| Causal Reasoning | 20 | 11.44 (3.78) | 11.61 (4.05) | 0.72 | 0.474 | 0.07 |
| Causal Link Extraction | 10 | 6.06 (1.98) | 6.09 (2.17) | 0.22 | 0.824 | 0.02 |
| Quiz Evaluation | 14 | 5.27 (2.35) | 5.63 (2.51) | 1.79 | 0.076 | 0.18 |

Figure 7 displays the distribution of best map scores achieved by students ($\mu = 6.87, \sigma = 5.24$). As in previous studies with *Betty’s Brain*, student performance on the task varied widely (Kinnebrew et al., 2013), with 37 students scoring below 5, 28 students scoring between 5 and 10, and 33 students scoring higher than 10. The maximum score students could obtain was 15, and 13 of the 98 students (13.3%) attained the maximum score.

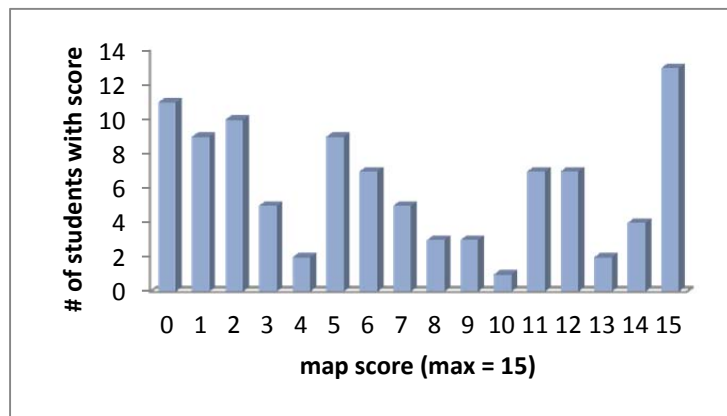


Figure 7. Map score distribution

6.2 Relationships between CA-Derived Metrics, Learning, and Performance

To investigate our first hypothesis, that students’ CA-derived metrics would predict their learning and success in teaching Betty, we first analyzed correlations between the CA metrics, learning gains, and students’ best map scores. The bottom row of Table 5 shows the overall descriptive statistics for the CA metrics. Students edited their maps fairly often (0.60 times per minute), and on average, 55.7% of these edits were supported. Students spent roughly one-third of their total time viewing information, but only 65.3% of this viewing time was spent on information that could support causal map edits. Students used, on average, a majority of the potential that they generated (62.3%), and they were mostly engaged in their learning (88.8%).

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Table 5. CA metrics, learning gain, and performance correlations

| Measure | Best Map | Sci. Content | Short Answer | Edit Freq. | Unsup. Edit % | Info. Viewing % | Potential Gen. % | Used Potential % | Disengaged % |
|------------------|-------------|--------------|--------------|-------------|---------------|-----------------|------------------|------------------|---------------|
| Sci. Content | 0.30** | 1 | | | | | | | |
| Short Answer | 0.58** | 0.29** | 1 | | | | | | |
| Edit Freq. | 0.56** | 0.16 | 0.40** | 1 | | | | | |
| Unsup. Edit % | -0.49** | -0.12 | -0.15 | -0.14 | 1 | | | | |
| Info. Viewing % | 0.11 | -0.02 | -0.07 | 0.35* | -0.27** | 1 | | | |
| Potential Gen. % | 0.41** | 0.10 | 0.26** | 0.14 | -0.48** | -0.01 | 1 | | |
| Used Potential % | 0.60** | 0.16 | 0.26* | 0.57* | -0.43** | -0.18 | 0.34** | 1 | |
| Disengaged % | -0.46** | 0.05 | -0.17 | 0.49* | 0.28** | -0.39** | -0.15 | -0.38** | 1 |
| Mean (SD) | 6.87 (5.21) | 1.42 (1.89) | 3.54 (2.52) | 0.60 (0.36) | 44.3% (21.9%) | 37.0% (11.2%) | 65.3% (17.1%) | 62.3% (21.0%) | 11.2% (10.5%) |

Note. * $p \leq 0.05$. ** $p \leq 0.01$.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

To test for the relationships between students' CA-derived metrics and their learning and performance, we calculated the pairwise correlations between students' CA metrics, their short answer learning gains, and their best map scores (Table 5). The results show that several of the CA metrics were significantly and moderately-to-strongly correlated with students' best map scores. Best map scores were positively correlated with edit frequency ($r = 0.56$), potential generation percentage ($r = 0.41$), and used potential percentage ($r = 0.60$). Best map scores were negatively correlated with unsupported edit percentage ($r = -0.49$) and disengaged percentage ($r = -0.46$). Students' CA metrics were also correlated with their short answer learning gains. More specifically, short answer learning gains were positively correlated with edit frequency ($r = 0.40$), potential generation percentage ($r = 0.26$), and used potential percentage ($r = 0.26$).

Interestingly, several CA-derived metrics were significantly correlated with each other. Students with higher levels of disengagement performed fewer edits per minute ($r = -0.49$), a higher proportion of which were unsupported ($r = 0.28$). They also spent a smaller percentage of their time viewing sources of information ($r = -0.39$) and took advantage of proportionally less of the information they encountered ($r = -0.38$).

To investigate these relationships further, we conducted multiple regression analyses to predict each of the learning gain measures and the best map score with the six CA metrics. The CA metrics predicted best map scores ($F = 22.87$, $p < 0.001$, $R^2 = 0.601$) and gains on short answer items ($F = 4.544$, $p < 0.001$, $R^2 = 0.231$) with statistical significance. With respect to map scores, the CA metrics of edit frequency ($Beta = 0.56$, $t = 5.10$, $p < 0.01$), information viewing percentage ($Beta = 0.36$, $t = 3.44$, $p < 0.01$), potential generation percentage ($Beta = 0.20$, $t = 2.52$, $p = 0.01$), and used potential percentage ($Beta = 0.26$, $t = 2.80$, $p < 0.01$) each added significantly to the prediction. With respect to short answer items, edit frequency ($Beta = 0.56$, $t = 3.73$, $p < 0.01$) and potential generation percentage ($Beta = 0.25$, $t = 2.29$, $p = 0.02$) each added significantly to the prediction. Conversely, the CA metrics did not predict gains on multiple choice ($F = 1.60$, $p = 0.156$, $R^2 = 0.095$), causal reasoning ($F = 0.53$, $p = 0.782$, $R^2 = 0.034$), causal link extraction ($F = 1.02$, $p = 0.417$, $R^2 = 0.063$), or quiz evaluation items ($F = 1.20$, $p = 0.316$, $R^2 = 0.073$) with statistical significance.

Together, these analyses provide potential insight into why particular students experienced more or less success. Negative correlations between unsupported edit percentage and information viewing percentage, potential generation percentage, and used potential percentage along with the positive correlation between unsupported edit percentage and disengaged percentage may suggest a behaviour profile characterized by disengagement, effort avoidance, and/or a difficulty in identifying causal links in the resources.

In summary, these results provide support for our first hypothesis. Students' CA-derived metrics were collectively predictive of their short answer learning gains and collectively and individually predictive of

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

their success in teaching Betty. Students who edited their maps more often, spent more time viewing information, viewed proportionally more relevant sources of information, and attempted to apply that information (via supported edits) achieved higher map scores. To gain further insight into this and other possible behaviour profiles, we performed a more comprehensive behaviour analysis in the next section.

To test our second hypothesis — that students’ prior levels of task understanding would predict their CA-derived metrics while using *Betty’s Brain* — we calculated correlations between CA metrics and students’ pre-test skill levels (6). Most correlations are weak. However, some specific pre-test skill levels were weakly and significantly correlated with the CA metrics. Students with higher causal reasoning scores edited their maps more often ($r = 0.23$) and used proportionally more of the potential they generated ($r = 0.27$). Those with higher causal link extraction scores edited their maps more often ($r = 0.34$), had higher potential generation percentage ($r = 0.26$), and higher levels of used potential percentage ($r = 0.21$). Finally, students with higher quiz evaluation pre-test scores had higher levels of used potential percentages ($r = 0.24$).

To investigate this further, we conducted regression analyses to predict each of the six behaviour metrics from students’ pre-test skill levels. The pre-test skill levels predicted edit frequency ($F = 4.91, p = 0.003, R^2 = 0.136$), potential generation percentage ($F = 2.76, p = 0.047, R^2 = 0.284$), and used potential percentage ($F = 4.09, p = 0.009, R^2 = 0.115$) with statistical significance. In these tests, only students’ causal link extraction score added to the prediction of edit frequency with statistical significance ($Beta = 0.29, t = 2.50, p = 0.01$). Conversely, the pre-test skill levels did not predict unsupported edit percentage ($F = 2.40, p = 0.072, R^2 = 0.071$), information viewing percentage ($F = 0.28, p = 0.843, R^2 = 0.009$), or disengaged percentage ($F = 1.93, p = 0.129, R^2 = 0.058$) with statistical significance.

Table 6. Correlations between skill level at pre-test and behaviour metrics

| | Causal Reasoning | Causal Link Extraction | Quiz Evaluation |
|------------------------|------------------|------------------------|-----------------|
| Causal Link Extraction | 0.53** | 1 | |
| Quiz Evaluation | 0.17 | 0.20* | 1 |
| Edit Frequency | 0.23* | 0.34** | 0.19 |
| Unsupported Edit % | -0.14 | -0.24* | -0.16 |
| Info Viewing % | 0.07 | -0.01 | -0.02 |
| Potential Generation % | 0.17 | 0.26** | 0.16 |
| Used Potential % | 0.27** | 0.21* | 0.24* |
| Disengaged % | -0.14 | -0.23* | -0.13 |

Note. * $p \leq 0.05$. ** $p \leq 0.01$.

To summarize, we found only limited support for our second hypothesis. Students who were better at interpreting causal relations in text passages during the pre-test edited their maps somewhat more frequently and exhibited slightly higher levels of coherence by generating and using proportionally more

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

potential.

6.3 Exploratory Clustering Analysis

In addition to testing our hypotheses, we performed an exploratory analysis to identify and characterize common behaviour profiles exhibited by students. These profiles may provide insight into students' SRL strategies as they worked toward completing the *Betty's Brain* task. For this analysis, we clustered students with a complete-link hierarchical clustering algorithm (Jain & Dubes, 1988; Murtagh, 1983), where each student was described by their set of CA metrics (listed in the Log File Analysis section above). The Euclidean distance between students' normalized CA metrics was used as the measure of dissimilarity among pairs of students. Clustering was performed using version 2.7 of the Orange data mining toolbox (Demšar, Curk, & Erjavec, 2013).

Figure 8 illustrates the resulting dendrogram. The analysis revealed five relatively distinct clusters containing 24, 39, 5, 6, and 24 students. Table 7 displays the means (and standard deviations) of the CA metrics for each cluster. The clustering results show distinct behaviour profiles among the 98 students in the study. Cluster 1 students (n=24) may be characterized as *frequent researchers and careful editors*; these students spent large proportions of their time (42.4%) viewing sources of information and did not edit their maps very often. When they did edit their maps, the edit was usually supported by recent activities (unsupported edit percentage = 29.4%). Most of the information they viewed was useful for improving their causal maps (potential generation percentage = 71.4%), but they often did not take advantage of this information (used potential percentage = 58.9%). Cluster 2 students (n=39) may be characterized as *strategic experimenters*. These students spent a fair proportion of their time (33.5%) viewing sources of information, and, like Cluster 1 students, often did not take advantage of this information (used potential percentage = 62.6%). Unlike Cluster 1 students, they performed more map edits, a higher proportion of which were unsupported, as they tried to construct the correct causal model.

Cluster 3 students (n=5) may be characterized as *confused guessers*. These students edited their maps fairly infrequently and usually without support. They spent an average of 58.9% of their time viewing sources of information, but most of their time viewing information did not generate potential (potential generation percentage = 45.8%). One possibility is that these students struggled to differentiate between more and less helpful sources of information. Unfortunately, when they did view useful information, they often did not take advantage of it (used potential percentage = 23.1%), indicating that they may have struggled to understand the relevance of the information they encountered. Students in Cluster 4 (n=6) may be characterized as *disengaged from the task*. On average, these students spent more than 30% of their time on the system (more than 45 minutes of class time) in a state of disengagement. Like confused guessers, disengaged students had a very high proportion of unsupported edits, low potential generation percentage, and low used potential percentage. In addition, their information viewing percentage was much lower, though their edits per minute were slightly higher

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

than the confused guessers.

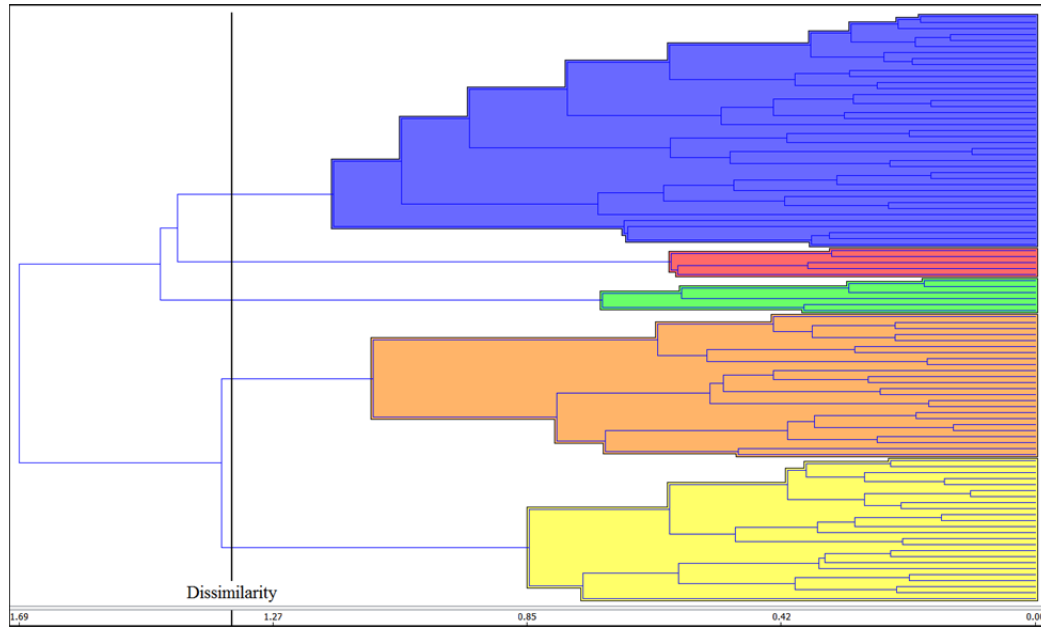


Figure 8. Dendrogram of students’ thermoregulation behaviour profiles

Table 7. Means (and standard deviations) of CA-derived metrics by cluster

| Cluster | Edit Freq. | Unsup. Edit % | Info. View % | Potential Gen. % | Used Potential % | Disengaged % |
|--------------------------------|-------------|---------------|---------------|------------------|------------------|--------------|
| 1. Res./Careful Editors (n=24) | 0.30 (0.11) | 29.4% (16.1%) | 42.4% (11.0%) | 71.4% (10.6%) | 58.9% (15.4%) | 15.7% (9.9%) |
| 2. Strat. Experimenters (n=39) | 0.60 (0.23) | 54.4% (14.8%) | 33.5% (8.3%) | 58.7% (18.9%) | 62.6% (16.2%) | 10.9% (7.4%) |
| 3. Confused Guessers (n=5) | 0.21 (0.06) | 73.5% (13.5%) | 58.9% (7.7%) | 45.8% (19.4%) | 23.1% (12.6%) | 4.8% (5.4%) |
| 4. Disengaged (n=6) | 0.33 (0.11) | 74.7% (17.4%) | 27.0% (9.6%) | 54.9% (9.3%) | 28.0% (8.7%) | 33.6% (8.4%) |
| 5. Engaged/Efficient (n=24) | 1.04 (0.32) | 29.1% (15.2%) | 35.4% (8.6%) | 76.8% (9.5%) | 82.0% (9.0%) | 3.1% (5.0%) |

Cluster 5 students (n=24) are characterized by a high edit frequency (just over 1 edit per minute), and most of these students’ edits (70.9%) were supported. Additionally, they spent just over one-third of their time viewing information, and over three-fourths of this time viewing information that generated potential. These students are distinct from students in the other four clusters in that they used a large majority of the potential they generated (82.0%) and were rarely in a state of disengagement (3.1%). In other words, these students appeared to be *engaged and efficient*. Their behaviour is indicative of students who knew how to succeed in *Betty’s Brain* and were willing to exert the necessary effort.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Table 8 shows the pre-test and post-test scores broken down by cluster. Cohen's d calculations were computed using equation 8 from Morris & DeShon (2002), which corrects for dependence among the means in within-subjects t-tests. Note that the small cluster sizes of Clusters 3 and 4 ($n=5$ and 6 , respectively) necessitates caution when interpreting statistical tests performed on the data from these two clusters. Nevertheless, a 5×2 repeated-measures ANOVA run on the data revealed a main effect of cluster on short answer questions ($F = 5.09$, $p = 0.001$), causal extraction problems ($F = 2.82$, $p = 0.029$), and quiz evaluation problems ($F = 3.96$, $p = 0.005$). Tukey HSD-adjusted pairwise comparisons between the clusters showed that 1) Cluster 5's short answer scores were significantly higher than the scores of Cluster 1 ($p^{adjusted} = 0.001$) and higher, but not significantly higher, than the scores of Cluster 4 ($p^{adjusted} = 0.064$); 2) Cluster 5's causal extraction scores were significantly higher than the scores of Cluster 4 ($p^{adjusted} = 0.027$); and 3) Cluster 3's quiz evaluation scores were significantly lower than the scores of Clusters 1 ($p^{adjusted} = 0.036$), 2 ($p^{adjusted} = 0.030$), and 5 ($p^{adjusted} = 0.003$).

The analysis also revealed an interaction between time and cluster for short answer questions ($F = 4.86$, $p = 0.001$). Follow-up ANOVAs on the pre-test and post-test short answer scores did not reveal a significant effect of cluster on short answer pre-test scores ($F = 1.92$, $p = 0.102$), but they did find a significant effect of cluster on short answer post-test scores ($F = 5.70$, $p < 0.001$). Tukey HSD-adjusted pairwise comparisons between the clusters showed that Cluster 5's short answer post-test scores were significantly higher than the scores of Clusters 1 ($p^{adjusted} < 0.001$, $d = 1.20$), 2 ($p^{adjusted} = 0.018$, $d = 0.83$), and 4 ($p^{adjusted} = 0.043$, $d = 1.34$). These results show that Cluster 5 students, characterized as engaged and efficient, exhibited significantly higher short answer item gains when compared to most of the other student clusters.

Table 9 displays the means (and standard deviations) of the best map scores achieved by students in each cluster. Because the map scores exhibited a non-normal distribution, we tested for differences among clusters using a Kruskal-Wallis H test. The test identified a statistically significant difference in map scores between the clusters ($\chi^2(4) = 35.70$, $p < 0.001$), with a mean rank score of 41.58 for Cluster 1, 47.91 for Cluster 2, 24.20 for Cluster 3, 14.00 for Cluster 4, and 74.25 for Cluster 5. Follow-up Mann-Whitney tests between the groups showed that Cluster 5 students achieved higher map scores than students in Clusters 1 ($p < 0.001$, $d = 1.50$), 2 ($p = 0.001$, $d = 1.49$), 3 ($p = 0.001$, $d = 3.70$), and 4 ($p < 0.001$, $d = 4.05$). As with the learning results, engaged and efficient students performed significantly better than most other clusters.

To explore more detailed behaviour differences between the identified clusters, we employed the information-theoretic differential sequence mining approach described in Kinnebrew, Mack, Biswas, & Chang (2104) and the Log File Analysis section above. This approach identified the action patterns that best differentiate the five clusters, the top seven of which are presented in Table 10⁶. In previous work,

⁶ The top differential activity patterns presented in Table 10 are those that included multiple distinct actions, leaving out trivial patterns of the same action simply repeated multiple times.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

we have argued that these patterns, when interpreted in the context of the task model, may be indicative of strategies students employ in building and refining their maps (Kinnebrew, Segedy, & Biswas, 2014). The pattern that most effectively differentiated clusters involved adding an incorrect link and then annotating an incorrect link as being correct (usually this was the same link just added). This pattern, most frequently used by strategic experimenters and disengaged students, suggests a potential misunderstanding of the use of link annotation functionality. According to the task model, students should only mark links as being correct once those links have been used in Betty’s correct quiz answers. On average, the pattern was performed 5 to 6 times by strategic experimenters and disengaged students and less than once by students in the other clusters.

Table 8. Means (and standard deviations) of assessment test scores by cluster

| Measure | Max | Cluster | Pre-test | Post-test | <i>t</i> | <i>p</i> | Cohen’s <i>d</i> |
|-------------------|-----|--------------------------|--------------|--------------|----------|----------|------------------|
| Science Content | 6 | 1 (Res./Careful Editors) | 2.17 (1.40) | 3.88 (1.51) | 3.48 | 0.02 | 0.71 |
| | | 2 (Strat. Experiment.) | 2.59 (1.21) | 3.72 (1.73) | 4.52 | 0.01 | 0.76 |
| | | 3 (Confused Guessers) | 2.20 (1.79) | 3.60 (1.95) | 1.25 | 0.28 | 0.56 |
| | | 4 (Disengaged) | 2.83 (1.47) | 3.33 (1.21) | 0.75 | 0.49 | 0.31 |
| | | 5 (Engaged/Efficient) | 2.58 (1.32) | 4.42 (1.59) | 5.10 | 0.01 | 1.05 |
| Short Answer | 11 | 1 (Res./Careful Editors) | 0.73 (1.13) | 3.46 (2.50) | 5.17 | 0.01 | 1.15 |
| | | 2 (Strat. Experiment.) | 1.44 (1.27) | 4.53 (2.14) | 8.75 | 0.01 | 1.47 |
| | | 3 (Confused Guessers) | 0.70 (0.45) | 3.80 (2.68) | 2.44 | 0.07 | 1.24 |
| | | 4 (Disengaged) | 0.67 (0.61) | 3.42 (2.06) | 3.51 | 0.02 | 1.84 |
| | | 5 (Engaged/Efficient) | 1.08 (1.00) | 6.44 (2.46) | 11.29 | 0.01 | 2.69 |
| Causal Reasoning | 20 | 1 (Res./Careful Editors) | 10.67 (3.71) | 11.17 (4.40) | 0.89 | 0.38 | 0.19 |
| | | 2 (Strat. Experiment.) | 11.72 (3.75) | 11.62 (3.77) | 0.25 | 0.81 | 0.04 |
| | | 3 (Confused Guessers) | 9.60 (3.36) | 9.80 (3.70) | 0.22 | 0.84 | 0.10 |
| | | 4 (Disengaged) | 8.83 (1.33) | 9.83 (1.72) | 2.24 | 0.08 | 0.97 |
| | | 5 (Engaged/Efficient) | 12.79 (3.97) | 12.88 (4.46) | 0.20 | 0.84 | 0.05 |
| Causal Extraction | 10 | 1 (Res./Careful Editors) | 5.79 (1.77) | 5.88 (1.87) | 0.31 | 0.76 | 0.07 |
| | | 2 (Strat. Experiment.) | 5.85 (1.71) | 5.97 (2.13) | 0.51 | 0.61 | 0.08 |
| | | 3 (Confused Guessers) | 5.80 (3.42) | 6.20 (3.27) | 1.00 | 0.37 | 0.45 |
| | | 4 (Disengaged) | 4.83 (1.33) | 4.00 (1.55) | 1.27 | 0.26 | 0.52 |
| | | 5 (Engaged/Efficient) | 7.04 (2.16) | 7.00 (2.09) | 0.20 | 0.84 | 0.04 |
| Quiz Evaluation | 14 | 1 (Res./Careful Editors) | 5.21 (2.11) | 5.75 (3.19) | 1.25 | 0.23 | 0.29 |
| | | 2 (Strat. Experiment.) | 5.26 (1.94) | 5.64 (1.86) | 1.42 | 0.17 | 0.22 |
| | | 3 (Confused Guessers) | 2.00 (2.00) | 3.00 (2.83) | 1.41 | 0.23 | 0.73 |
| | | 4 (Disengaged) | 5.00 (2.76) | 3.67 (2.58) | 1.75 | 0.14 | 0.72 |
| | | 5 (Engaged/Efficient) | 6.08 (2.69) | 6.54 (2.11) | 0.92 | 0.37 | 0.19 |

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Table 9. Means (and standard deviations) of map score metrics by cluster

| Cluster | Best Map – Correct Links | Best Map – Incorrect Links | Best Map Score (max = 15) |
|--------------------------|--------------------------|----------------------------|---------------------------|
| 1 (Res./Careful Editors) | 6.42 (5.69) | 0.96 (1.65) | 5.46 (5.27) |
| 2 (Strat. Experiment.) | 7.54 (4.76) | 1.41 (1.58) | 6.13 (4.40) |
| 3 (Confused Guessers) | 2.80 (2.39) | 0.80 (1.30) | 2.00 (2.00) |
| 4 (Disengaged) | 1.17 (1.94) | 0.00 (0.00) | 1.17 (1.94) |
| 5 (Engaged/Efficient) | 12.67 (2.85) | 0.75 (1.15) | 11.92 (3.37) |

Table 10. Means (and standard deviations) of activity pattern frequency by cluster

| Pattern | Res./Careful Editors | Strat. Exper. | Confused Guessers | Disengaged | Engaged/Efficient |
|--|----------------------|---------------|-------------------|-------------|-------------------|
| 1. [Add Incorrect Link] → [Mark Incorrect Link as Being Correct] | 0.04 (0.20) | 5.28 (8.81) | 0.60 (0.89) | 5.83 (6.52) | 0.83 (3.29) |
| 2. [Quiz] → [Remove Incorrect Link] | 2.00 (4.24) | 3.67 (6.18) | 0.20 (0.45) | 0.33 (0.52) | 12.29 (12.33) |
| 3. [Remove Incorrect Link] → [Quiz] | 1.25 (1.33) | 2.97 (5.73) | 0.00 (0.00) | 0.17 (0.41) | 9.33 (8.65) |
| 4. [Add Incorrect Link] → [Quiz] → [Remove Incorrect Link] | 1.08 (4.27) | 2.56 (5.53) | 0.00 (0.00) | 0.17 (0.41) | 8.79 (10.47) |
| 5. [Add Incorrect Link] → [Quiz] | 4.54 (4.62) | 7.72 (7.63) | 2.20 (1.30) | 2.50 (1.38) | 17.54 (13.06) |
| 6. [Check Quiz Answer Explanation] → [Remove Incorrect Link] | 1.79 (1.61) | 2.54 (2.87) | 0.20 (0.45) | 0.33 (0.82) | 9.08 (6.40) |
| 7. [Remove Incorrect Link] → [Add Incorrect Link] → [Quiz] | 1.08 (1.06) | 1.59 (2.58) | 0.00 (0.00) | 0.00 (0.00) | 6.46 (6.11) |

All of the other top differential activity patterns described in Table 10 involve a combination of map editing (specifically with respect to incorrect links) and quizzing, patterns that have been associated with successful performance in *Betty’s Brain* (Kinnebrew et al., 2013). For example, patterns 2 and 6 are characteristic of *supported map edits based on quiz results*. In particular, pattern 6 is characteristic of exploring the quiz results more deeply by viewing Betty’s answer explanation. Patterns 3, 5, and 7 are characteristic of *using quizzes to monitor progress*. After the student edits the map, they have Betty take a quiz in order to evaluate the effect of that edit on her quiz performance. Pattern 4 combines these two pattern types into a pattern characteristic of an *edit-and-check* strategy, in which students add a link to their map, use a quiz to monitor the effect of that link on Betty’s performance, and then, upon discovering the link is incorrect, remove it from their maps.

All six of these activity patterns display similar relative use across the clusters, having the highest average frequency in engaged and efficient students, moderate frequencies in researchers/careful editors and strategic experimenters, and low frequencies in confused guessers and disengaged students.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Interestingly, these relative frequencies follow the same pattern as overall map scores and number of correct links in comparing performance across clusters, as illustrated in Table 9. In other words, clusters that used these behaviour patterns more often were ones that also achieved greater success in teaching Betty. Moreover, they also had lower average unsupported edit percentages and higher used potential percentages. This suggests, but does not prove, that when these patterns were employed by students, the quizzes provided support for subsequent edits. Altogether, these results indicate that students in more successful clusters were more likely to employ behaviours illustrating productive uses of the quiz for solution assessment.

7 DISCUSSION AND CONCLUSIONS

This paper presented *Coherence Analysis (CA)*, a novel approach to measuring aspects of students' self-regulated learning (SRL) behaviours in open-ended learning environments (OELEs). CA focuses on the learner's ability to seek, interpret, and apply information encountered while working in an OELE. By characterizing behaviours in this manner, CA provides insight into students' open-ended problem-solving strategies, as well as the extent to which they understand the nuances of the learning task they are currently completing. We applied CA to data from a recent classroom study with *Betty's Brain* to test two hypotheses: 1) students' CA-derived metrics would predict their learning and success in teaching Betty; and 2) students' prior levels of task understanding would predict their CA-derived metrics while using *Betty's Brain*. Results showed some support for both hypotheses: CA-derived metrics were predictive of students' task performance and learning gains, and students' prior skill levels were (weakly) predictive of some of the CA metrics, suggesting a link between task understanding and effective open-ended problem-solving behaviours. In addition to testing these hypotheses, we applied a clustering analysis to characterize students based on their CA metrics, and this provided insights into common problem-solving approaches used by students in this study.

One important limitation of this work is the fact that we directly assessed students' task understanding (via their skills) during the pre- and post-tests without similarly assessing aspects of their metacognitive knowledge and regulation. Students with high task understanding may still exhibit difficulty in employing metacognitive processes, such as goal setting, planning, monitoring, and reflection. Future work should investigate the relationships between metacognitive knowledge, task understanding, and CA-derived metrics in OELEs. Another limitation is that the CA metrics were based on action coherence metrics without considering action incoherence. In future work, we will examine the relationships between students' learning, performance, action coherence, and action incoherence.

7.1 Coherence Analysis and SRL in OELEs

CA, as distinct from analyses of students' learning and task performance, provides insight into aspects of students' SRL behaviours, particularly in OELEs. Several of the behaviour profiles, identified using cluster analysis with the CA metrics as features, exhibited similar levels of prior knowledge, prior skill levels,

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

success in teaching Betty, and learning while using the system. CA helps us understand how different behaviours can result in the same level of performance and learning. In fact, one of the more interesting findings that emerged from this study is that CA metrics were able to distinguish groups of students, based on their behaviours, beyond what was possible when only focusing on learning gains and map scores. Certainly, students in the engaged and efficient cluster had higher prior skill levels, better map scores, and higher learning gains than students in most other clusters. However, it is harder to distinguish the remaining four clusters in terms of learning gains and performance. Confused guessers scored lower on quiz evaluation items when compared to researchers/careful editors and strategic experimenters. However, there were no other measurable differences in performance and learning gains. Despite this, these groups of students adopted distinct approaches to completing the Betty's Brain learning task, as measured by CA. Future work should investigate these profiles further with more detailed analyses of strategic behaviours. For example, further analysis of observed behaviour patterns and interviewing students with different behaviour profiles may help us understand the intentions that drove students' problem-solving strategy selections. Additional work should also look for interactions between CA-derived metrics and other aspects of SRL, such as affect and self-efficacy.

7.2 Implications for the Design of OELEs

One interesting set of findings from this study involves the predictive relationships between students' task understanding (as measured by their skill levels) and their behaviours (as measured by CA). The results seem to validate, at least to some extent, the *metacognitive knowledge dilemma* presented by Land (2000). This dilemma states that success in OELEs depends not only on students' metacognitive skills, but also on their understanding of the overall task and its components. In this study, students' prior skill levels were predictive of some CA metrics, which, in turn, predicted their success in teaching Betty. Specifically, students with higher task understanding had higher levels of coherence, and students with higher levels of coherence were more successful in their map-building tasks and demonstrated larger learning gains on short answer questions. Therefore, building coherence detectors into OELEs can provide a mechanism for first identifying low levels of coherence and then performing more targeted diagnosis of students' task understanding (perhaps via a method similar to rapid dynamic assessments; Kalyuga & Sweller, 2005). This mechanism, then, could identify and scaffold causes of poor SRL and problem-solving behaviours. For example, the system could provide students with opportunities to practice and develop their skills (Segedy et al., 2013) while explaining relevant problem-solving strategies.

Further, mining of behaviour sequences across identified clusters showed that students with more successful behaviour profiles were more likely to use the quiz productively (patterns 2 to 7 in Table 10). This is especially interesting given the fact that students' quiz evaluation skills were far less predictive of their behaviour than were their causal extraction skills. It makes sense to hypothesize that students cannot take advantage of the quiz functionality unless they can identify causal relations in text passages. Given this, scaffolding agents may support students by first helping them develop their information-

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

seeking skills, and then helping them to develop their solution assessment skills. Future work should investigate these relationships in more detail and how to provide the right scaffolding at the right time.

Another potentially powerful application of CA in OELEs, and one that we are particularly excited about, is presenting CA metrics to classroom teachers for evaluation and formative assessment. Ideally, teachers could use these reports to quickly and easily 1) understand learners' problem-solving approaches, 2) infer potential reasons for the levels of success achieved by students, and 3) make predictions about students' learning and task understanding. Moreover, teachers could use these reports to assign performance and effort grades and implement classroom and homework activities that target the aspects of SRL and problem solving with which students are struggling. However, additional research is required to understand how best to present and use this data with classroom teachers.

In future work, we plan to investigate the predictive power of additional CA-derived metrics via feature engineering and selection (Peng, Long, & Ding, 2005). In this study, we chose six metrics that successfully differentiated students and predicted aspects of their learning and performance. However, other CA-derived metrics may be better predictors. For example, it may be valuable to represent actions based on the *amount of support* they have from previous actions, rather than a binary measure of whether they do/do not have *any* support. As another example, it may be valuable to investigate CA metrics that incorporate more fine-grained aspects of how student behaviour *changes over time*. Ideally, this will allow us to study the development of SRL as students' task understanding and problem-solving skills improve in the *Betty's Brain* environment. Further, by studying aspects of coherence across multiple OELEs, we could also gain insight into how students generalize aspects of SRL and open-ended problem-solving strategies and skills over a more extended period.

ACKNOWLEDGEMENTS

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120186 to Vanderbilt University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

REFERENCES

- Azevedo, R., & Witherspoon, A. (2009). Self-regulated learning with hypermedia. In D. Hacker, J. Donlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 319–339). New York, NY: Taylor & Francis.
- Baker, R. S. J. d., Corbett, A., Koedinger, K., Evenson, S., Roll, I., Wagner, A., Naim, M., Raspat, J., Baker, D., & Beck, J. (2006). Adapting to when students game an intelligent tutoring system. In M. Ikeda, K. D. Ashley, & T. Chan (Eds.), *Lecture Notes in Computer Science: Intelligent Tutoring Systems, 4053* (pp. 392–401). Berlin: Springer-Verlag.
- Baker, R. S. J. d., Ocumpaugh, J., Gowda, S. M., Kamarainen, A. M., & Metcalfe, S. J. (2014). Extending log-

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

- based affect detection to a multi-user virtual environment for science. In V. Dimitrova, T. Kuflik, D. Chin, F. Ricci, P. Dolog, & G. Houben (Eds.), *Lecture Notes in Computer Science: User Modeling, Adaptation, and Personalization*, 8538 (pp. 290–300). Cham, Switzerland: Springer International Publishing AG.
- Barab, S., Hay, K., Barnett, M., & Keating, T. (2000). Virtual solar system project: Building understanding through model building. *Journal of Research in Science Teaching*, 37(7), 719–756.
- Bouchet, F., Harley, J., Trevors, G., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining*, 5(1), 104–146.
- Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn*. Washington, DC: National Academy Press.
- Bransford, J., & Schwartz, D. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24(1), 61–101.
- Brown, A. (1975). The development of memory: Knowing about knowing, and knowing how to know. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 103–152). New York, NY: Academic Press.
- Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281.
- Chipman, S., Schraagen, J., & Shalin, V. (2000). Introduction to cognitive task analysis. In J. Schraagen, S. Chipman, & V. Shalin (Eds.), *Cognitive task analysis* (pp. 3–23). Abingdon, UK: Psychology Press.
- Clarebout, G., & Elen, J. (2008). Advice on tool use in open learning environments. *Journal of Educational Multimedia and Hypermedia*, 17(1), 81–97.
- Clark, D., Nelson, B., Chang, H., Martinez-Garza, M., Slack, K., & D'Angelo, C. (2011). Exploring Newtonian mechanics in a conceptually-integrated digital game: Comparison and learning and affective outcomes for students in Taiwan and the United States. *Computers & Education*, 57(3), 2178–2195.
- Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *Information Theory, IEEE Transactions on*, 38(2), 713–718.
- Demšar, J., Curk, T., & Erjavec, A. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, 14, 2349–2353.
- Flavell, J. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.
- Jain, A., & Dubes, R. (1988). *Algorithms for clustering data*. Upper Saddle River, NJ: Prentice Hall.
- Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development*, 53(3), 83–93.
- Hacker, D., Dunlosky, J., & Graesser, A. (Eds.). (2009). *Handbook of metacognition in education*. New York, NY: Taylor & Francis.
- Kinnebrew, J. S., Loretz, K., & Biswas, G. (2013). A contextualized, differential sequence mining method

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

- to derive students' learning behavior patterns. *Journal of Educational Data Mining*, 5(1), 190–219.
- Kinnebrew, J. S., Mack, D. L. C., Biswas, G., & Chang, C. K. (2014). A differential approach for identifying important student learning behavior patterns with evolving usage over time. *Proceedings of the Mobile Sensing, Mining and Visualization for Human Behavior Inference Workshop at the 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 281–292). Tainan, Taiwan. Retrieved from http://link.springer.com/chapter/10.1007%2F978-3-319-13186-3_27#page-1
- Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2014). Analyzing the temporal evolution of students' behaviors in open-ended learning environments. *Metacognition and Learning*, 9(2), 187–215.
- Land, S. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, 48(3), 61–78.
- Land, S., Hannafin, M., & Oliver, K. (2012). Student-centered learning environments: Foundations, assumptions and design. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (pp. 3–25). New York, NY: Routledge.
- Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education*, 18(3), 181–208.
- Luckin, R., & du Boulay, B. (1999). Ecolab: The development and evaluation of a Vygotskian design framework. *International Journal of Artificial Intelligence in Education*, 10(2), 198–220.
- Luckin, R., & Hammeron, L. (2002). Getting to know me: Helping learners understand their own learning needs through metacognitive scaffolding. In S. Cerri, G. Gouardres, & F. Paraguau (Eds.), *Lecture Notes in Computer Science: Intelligent Tutoring Systems*, 2363 (pp. 759–771). Berlin: Springer-Verlag.
- Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, 59(1), 14–19.
- McNamara, D. (2004). SERT: Self-explanation reading training. *Discourse Processes*, 38(1), 1–30.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125.
- Murtagh, F. (1983). A survey of recent advances in hierarchical clustering algorithms. *The Computer Journal*, 26(4), 354–359.
- Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), 1226–1238.
- Quintana, C., Reiser, B., Davis, E., Krajcik, J., Fretz, E., Duncan, R., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, 13(3), 337–386.
- Roscoe, R., Segedy, J., Sulcer, B., Jeong, H., & Biswas, G. (2013). Shallow strategy development in a teachable agent environment designed to support self-regulated learning. *Computers & Education*, 62, 286–297.

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

- Sabourin, J., Mott, B., & Lester, J. (2013). Discovering behavior patterns of self-regulated learners in an inquiry-based learning environment. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Lecture Notes in Computer Science: Artificial Intelligence in Education, 7926* (pp. 209–218). Berlin: Springer-Verlag.
- Sabourin, J., Shores, L., Mott, B., & Lester, J. (2013). Understanding and predicting student self-regulated learning strategies in game-based environments. *International Journal of Artificial Intelligence in Education*, 23, 94–114.
- Schraw, G., Crippen, K., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36(1), 111–139.
- Segedy, J. R. (2014). *Adaptive scaffolds in open-ended computer-based learning environments* (Doctoral dissertation). Department of EECS, Vanderbilt University, Nashville, TN.
- Segedy, J. R., Biswas, G., & Sulcer, B. (2014). A model-based behavior analysis approach for open-ended environments. *The Journal of Educational Technology & Society*, 17(1), 272–282.
- Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2011). Modeling learners' cognitive and metacognitive strategies in an open-ended learning environment. *Advances in Cognitive Systems: Papers from the AAAI Fall Symposium* (pp. 297–304). Menlo Park, CA: AAAI Press.
- Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2013). The effect of contextualized conversational feedback in a complex open-ended learning environment. *Educational Technology Research and Development*, 61(1), 71–89.
- Sengupta, P., Kinnebrew, J., Basu, S., Biswas, G., & Clark, D. (2013). Integrating computational thinking with K–12 science education using agent-based computation: A theoretical framework. *Educational and Information Technologies*, 18(2), 351–380.
- Snow, E. L., Jackson, G. T., & McNamara, D. S. (2014). Emergent behaviors in computer-based learning environments: Computational signals of catching up. *Computers in Human Behavior*, 41, 62–70.
- Spires, H., Rowe, J., Mott, B., & Lester, J. (2011). Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, 44(4), 453–472.
- Suh, J., & Moyer, P. (2007). Developing students' representational fluency using virtual and physical algebra balances. *The Journal of Computers in Mathematics and Science Teaching*, 26(2), 155–173.
- van Joolingen, W., de Jong, T., Lazonder, A., Savelsbergh, E., & Manlove, S. (2005). Co-lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior*, 21(4), 671–688.
- Veenman, M. (2011). Learning to self-monitor and self-regulate. In R. Mayer & P. Alexander (Eds.), *Handbook of Research on Learning and Instruction* (pp. 197–218). New York, NY: Routledge.
- Veenman, M. (2013). Assessing metacognitive skills in computerized learning environments. In R. Azevedo & V. Alevin (Eds.), *International handbook of metacognition and learning technologies* (pp. 157–168), Springer International Handbooks of Education, Vol. 26. New York, NY: Springer

(2015). Using Coherence Analysis to Characterize Self-Regulated Learning Behaviours in Open-Ended Learning Environments. *Journal of Learning Analytics*, 2(1), 13–48.

Science+Business Media.

- Whitebread, D., & Cárdenas, V. (2012). Self-regulated learning and conceptual development in young children: The development of biological understanding. In A. Zohar & Y. J. Dori (Eds.), *Metacognition in science education: Trends in current research* (pp. 101–132), Contemporary Trends and Issues in Science Education, Vol. 40. Netherlands: Springer Science+Business Media.
- Winters, F., Greene, J., & Costich, C. (2008). Self-regulation of learning within computer-based learning environments: A critical synthesis. *Educational Psychology Review*, 20(4), 429–444.
- Young, A., & Fry, J. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning*, 8(2), 1–10.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3–17.
- Zimmerman, B., & Schunk, D. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York, NY: Routledge.