Moving beyond learning progressions to dynamic learning maps: A validation study

Russell E. Swinburne Romine

&

Jonathan Schuster

University of Kansas

Center for Educational Testing and Evaluation

Abstract

This study examined the proposed structure of a small section of a learning map consisting of nodes that represent cognitive skills and connections between nodes that represent the directional relationship between skills. The section of a learning map used for analysis represented skills for comprehension of argumentation in informational text. Two assessment forms were developed based on the hypothesized structure of the learning map aligned to portions of the Common Core State Standards in reading informational texts and delivered to middle school students via an online, formative testing tool. Structural equation modeling was used to examine the relationship between items written to assess nodes in the learning map. The factors that emerged from the analysis did not support the hypothesized relationships. Several explanations for the unexpected pattern of results are discussed.

Purpose

As the Common Core State Standards (CCSS) are implemented, instructors throughout the nation face the challenge of identifying the skills that their students possess and can demonstrate as well as the skill gaps or deficiencies that require further instruction. In short, teachers need instructionally relevant classroom assessments that align with the CCSS (Conley, 2011; Hess, 2012). Many traditional classroom assessments are based on escalated or linear learning progression frameworks that treat one skill as building on another single skill (Salinas, 2009). Although linear progressions are conceptually accessible to educators, the frameworks do not always represent the most effective or accurate descriptions of learning pathways, nor do they account for variability among students' individual learning progressions (Kingston & Erickson, 2010).

In light of the limitations of linear progressions and with the availability of statistical procedures capable of computing the contributions of multiple skills on a subsequent skill, research in assessment and instruction is moving from linear learning progression frameworks (Popham, 2008; 2011) to learning map frameworks that can better represent what students know and can do (e.g., Bechard, Hess, Camacho, Russell, & Thomas, 2012; Lee, Tidwell-Scheuring, & Barton, 2006). Learning maps, which have been referred to as landscape learning progressions, learning progressive networks, or cognitive learning models (Bechard et al., 2012; Salinas, 2009; Tatsuoka, 2009), can illustrate the many-to-many relationships between skills as well as multiple pathways for a student to build skills. Assessment and instruction based on learning maps, therefore, can allow for multiple pathways to target skills and provide finer grain representations of knowledge and skill acquisition using a node and connection web-like structure. Kingston and Erickson (2010) stated that, "Compared to other methods of organizing content, such as grade-level standards, the complex arrangement of learning targets in learning maps is conducive to the diagnosis of concept and skill deficits and prescriptive learning" (p. 8). In addition, learning maps may represent such transitional steps on a learning pathway as partially correct knowledge or incomplete mastery (Lee, et al., 2006). For the purposes of classroom assessment, a learning map of a complex set of cognitive skills and content area knowledge has the potential to yield useful diagnostic information for teachers about where to target instruction for individual students.

Therefore, the purpose of this study is to examine a small section of a Dynamic Learning Maps (DLM) English Language Arts (ELA) map that is aligned with the CCSS. Indeed, a learning map that is well aligned to CCSS standards may have the potential to yield more instructionally relevant information than a traditional progression. A learning map is especially applicable for English Language Arts teachers who need relevant assessments to support instructional decisions in reading and writing skills like argumentation and evaluation. DLM is the product of a U.S. Department of Education grant to develop an alternative assessment system based on alternate achievement standards (CFDA 84.373X, 2010). DLM constructed learning maps in both math and ELA based on a comprehensive, systematic process of research synthesis that represented developmental and educational research findings, largely from the general education population, on cognitive and conceptual progressions in student knowledge and skills. The DLM includes more than 3,500 nodes representing cognitive abilities and more than 7,500 connections between nodes representing progressions between those abilities. While developed as a framework for the design and implementation of an alternate assessment, DLM is a hypothesized framework of cognitive skills and relationships for all students. "The DLM uses the learning map to drive assessment item development and ultimately to make inferences about what students know and can do from assessment data" (Bechard et al., 2012, p. 8). There are currently no data in existence using the ELA portion of DLM that support testing of the hypothesized relationships represented in the map. By examining the relationships between skills represented in a small section of a learning map we aimed to develop and argument for validity for this part of the DLM. Our parallel intention was to support classroom use of instructionally-relevant assessments for this and other future applications of learning maps.

Theoretical Framework

Two theoretical frameworks are relevant to the present study. The first framework is that of learning maps, and the second is an instructional method for teaching the evaluation of arguments described by Bulgren and Ellis (2012), who based their ideas on Toulmin's (1958) theory of the components of argument. A brief overview of each follows.

Learning Maps

Learning maps are networks of sequenced learning targets worded in terms of the knowledge required for achievement. Unlike linear learning progressions, learning maps may contain alternate paths to a learning objective, when they exist, thus representing more than one possible learning progression for a student to follow to achieve mastery (Tatsuoka, 2009). The map arranges learning targets in what are referred to as learning-order relationships.

Learning maps are typically developed on the basis of learning-order relationships evident in research of linear relationships. The DLM attempts to integrate these learning-order relationships in various topical domains within English Language Arts to develop a more interconnected map of the domain. These extra-linear relationships provide the foundation for multiple pathways to the acquisition of academic content. However, due to the nature of the DLM construction, the network of relationships are only hypothetical, based on topical research available to the map developers. Therefore, the network of relationships, as a whole, needs validation if it is to be the basis for instructionally relevant assessment. The present study examines student performance on assessments constructed to assess the cognitive content that is represented by a small section of the learning map.

Figure 1 shows a section of the DLM learning map that contains nodes describing the skills aligned to elements of the standards from the CCSS in Reading Information strand in grades 6, 7 and 8. However, a limitation of viewing isolated sections of a learning map is that other relationships are invisible apart from the network of nodes and connections constituting the larger DLM learning map. A significant challenge in the development of the learning map in areas that represent complex cognitive content has been to represent the core cognitive components of a skill in a way that is applicable across a variety of contexts. One primary aim of this study was to produce tasks for students that were developed based on the cognitive skills represented for the nodes in the learning map using different texts. If items from both forms, which were written to different content from the texts, but written for the same cognitive skills specified by the nodes functioned similarly it would potentially provide evidence to support a validity argument for the section of the learning map that contained the nodes.

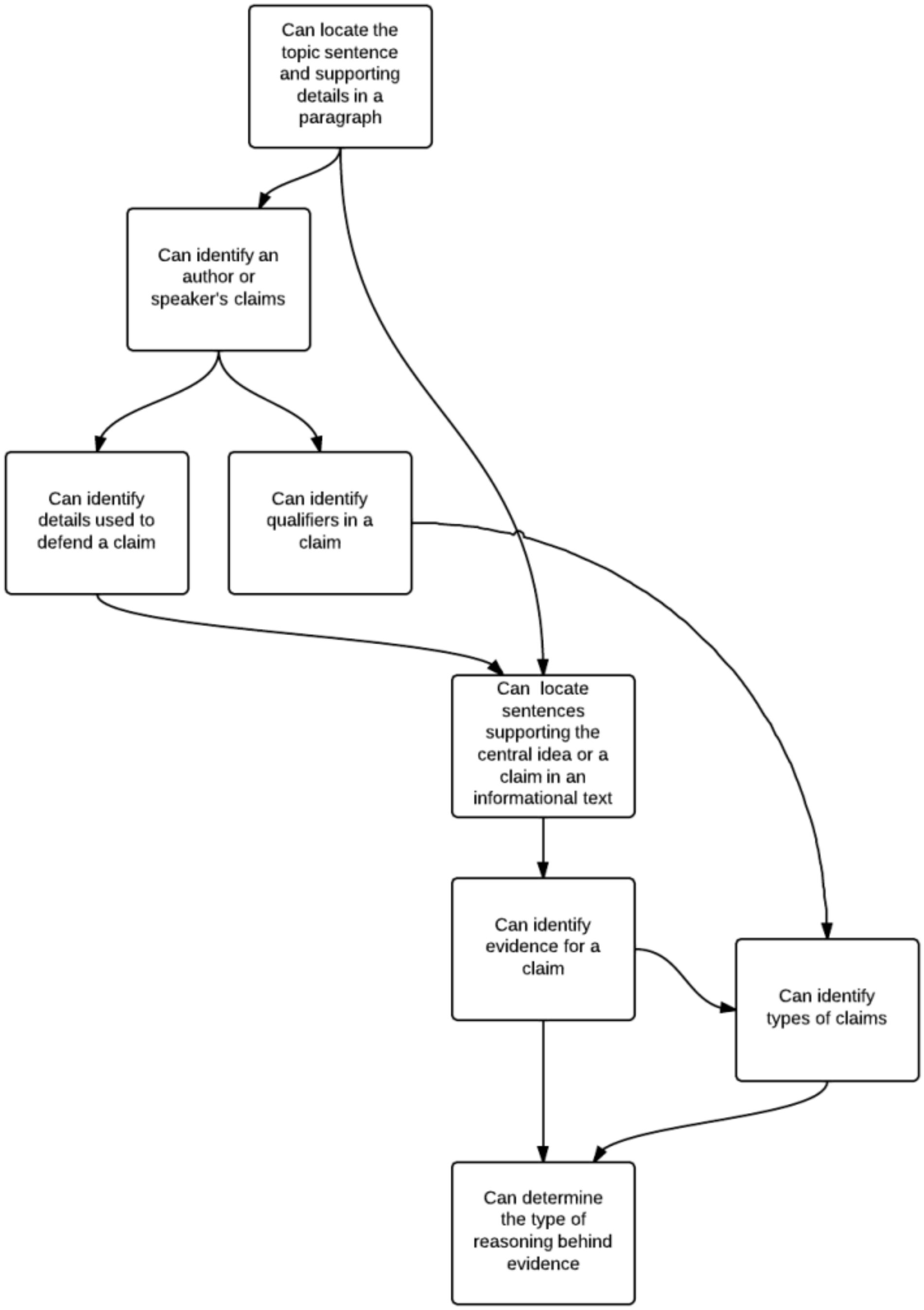

Figure 1. A section of the DLM English Language Arts map for evaluation and argumentation.

Evaluation and Argumentation

The CCSS place particular emphasis on integrated literacy development. Students in English Language Arts achieve an interconnected set of objectives designed to be mutually reinforcing across a range of texts and disciplines. This cross-disciplinary focus on literacy development is especially apparent in the CCSS emphasis on students' ability to read, comprehend, and analyze informational texts. Another CCSS focus, on identifying the cognitive processes that underlie comprehension rather than more traditional knowledge-based targets, presents a challenge for classroom teachers (Beach, 2011; Kober & Rentner, 2011). One example of these sets of cross-disciplinary skills is represented in Figure 1, which shows eight nodes that model relationships between the student critical thinking skills involved in evaluating the claims and arguments in texts. The Framework for K-12 Science and Engineering (National Academy Press, 2012) and the CCSS (CCSS, 2010) suggest that argumentation provides the basis for developing the high-order thought processes critical for academic success (Conley, 2008; Heller & Greenleaf, 2007; National Research Council, 2012).

An argument is a reason or group of reasons used to persuade someone to accept that some conclusion is self-evident. It typically includes a number of components that contribute to its meaning. At the base level, data are the details and evidence used to support the claims made in the argument. Warrants then combine the details and evidence using specific reasons and link them together into coherent units. The claims are then derived from these combined units of evidence. Thus, claims are the conclusions made in the argument based on the specific examples and reasons used in the argument. Because the reasons impact the relationship between the evidence and ultimately the meaning of the claim itself, the rationale for using the reasons to support the claim, the backing of an argument, is important (Toulmin, 1958; Toulmin et al., 1984). Additional argument components specify the validity of the claims. Qualifiers provide information as to the conditions under which the claims will be true. On the other hand, rebuttals are counter-arguments that indicate the conditions under which the claims will be false (Toulmin, 1958). In summary, arguments are supported by claims arising from the interaction of specific reasons, evidence, and backing, while the qualifiers and rebuttals determine the situations when the arguments are true and false, respectively.

Arguments can vary in their difficulty according to whether they include two components which influence an argument's quality: the presence of reasons and a rebuttal. Reasons, which include the evidence, warrants, and backing, substantiate a claim by providing a physical and rational basis for it. Without reasons, claims are just opinions and do not have the same justification as do claims with reasons. Similarly, rebuttals provide the basis for undermining other claims and are important for persuading others to accept the argument. Thus, arguments containing a rebuttal are of a higher quality than are arguments without one (Erduran, Simon, & Osborne, 2004; Pontecorvo & Erduran, 2003). Arguments containing reasons and a rebuttal are more difficult to comprehend than are arguments without reasons and a rebuttal.

Arguments can be ordered according to whether they contain reasons and a rebuttal. The easiest arguments contain only a simple claim and counterclaim. A slightly more difficult argument contains a claim with evidence, warrants, or backings and a counterclaim but no rebuttal. Adding to the previous argument, a complex argument includes a claim with evidence, warrants, or backings, a counterclaim, and a rudimentary rebuttal. From here, arguments increase in difficulty based on the specificity of the rebuttals and the number of claims and rebuttals. An argument with multiple claims and a specific rebuttal would be difficult to grasp, while an argument with multiple claims and rebuttals would be the most complex argument to comprehend (Erduran, Simon, & Osborne, 2004; Pontecorvo & Erduran, 2003). Thus, students will have more difficulty understanding arguments when contain reasons and multiple claims and rebuttals.

Determining the quality of arguments is a difficult process and requires higher-order thought processes. To evaluate an argument, students must realize the importance and quality of the evidence and reasoning (Kuhn, Amsel, & O'Loughlin, 1988; Latour & Woolgar, 1986). A high quality argument contains evidence and reasoning that is reliable (Schauble, 1996), valid (Koslowski et al., 1989; Kuhn et al., 1995; Schauble, 1996), and objective (Klahr et al., 1993; Kuhn et al., 1995; Penner & Klahr, 1996; Schauble, 1996). Thus, determining how and to what extent the evidence and reasoning actually support the claims and argument is an important skill in argumentation (Blair & Johnson, 1987; Latour & Woolgar, 1986). Additionally, the handling of alternative claims, counter-arguments, and rebuttals can influence an argument's quality (Kuhn, 1993; Marttunen, 1994). Students must determine whether the author explained away the challenges provided by alternative claims, counter-arguments, and rebuttals. In summary, ascertaining the quality of an argument is a critical analytical skill for adjudging whether an argument should be accepted or rejected.

The section of learning map (Figure 1) synthesizes research conducted by Bulgren and Ellis (2012) on the development of students' understanding of questions, claims, evidence and argument in the context of scientific reasoning. Their study used an instructional strategy, the Argumentation Evaluation Intervention (AEI) which emphasized Toulmin's (1958) elements of argumentation. In the intervention, the students are first asked to identify the claims and their qualifiers. The students are then asked to identify the evidence supporting the specific claims and its type (data, fact, opinion, or theory) and to evaluate the quality of the evidence. Thirdly, students are asked to identify the reasoning connecting the evidence to the claims and the type of reasoning (theory, authority, or logic) and to evaluate the quality of the reasoning. The students then identify possible counter-arguments and rebuttals. Lastly, students make a conclusion as to the quality of the argument and provide reasons for this conclusion based on the information they collected in the previous steps.

Bulgren and Ellis (2012) showed that the AEI helped students understand and evaluate arguments presented in scientific texts. We propose that this model of the evaluation of claims and arguments is generalizable to the larger context of reading and analyzing the content of different informational texts. A model of skill development based on a learning map synthesizing and representing empirical research on reading and analyzing claims and arguments has the potential to provide the foundation for building a useful set of instructionally relevant assessments across a variety of disciplinary contexts. A learning map can also provide educators with valuable information regarding individual students' skill development in the area of evaluation and argumentation.

Methods

In order to examine the relationships between skills represented in the learning map, we administered two forms through the Kansas Assessment Program formative assessment tool. Teachers are able to search and use test forms made publicly available through the formative test tool. Using this tool, we delivered two forms of multiple choice, reading comprehension assessments for informational middle school education texts with items assigned to specific nodes in the learning map. Data collection and procedures were approved by the University of Kansas institutional review board.

Materials

Two assessment forms, using different informational texts were administered to students in grade 6, 7 and 8 as part of the Kansas Assessment Program voluntary formative assessment tool. Teachers were invited to use the two formative assessments based on the identified section of the learning map as a part of their regular instruction. Since participation was voluntary based on teachers' interest and student consent we are limited to a convenience sample. We used convenience sampling, because we intended to examine the predicted pattern of relationships between skills and not the extent of the relationship for all students. Our primary interest is in examining the structure of the identified structure of the learning map while providing a formative assessment that may be useful to teachers and students.

Two forms were delivered as stand-alone formative assessments of students' ability to understand arguments and claims in text. The nodes from the learning map represented content from three grade level CCSS in reading informational texts. Table 1 shows the standards addressed in grades 6, 7 and 8.

Table 1. Standards Addressed in Forms A and B

Standard Name	Standard Description
CCSS.ELA-LITERACY.RI.6.8	Trace and evaluate the argument and specific claims in a text, distinguishing claims that are supported by reasons and evidence from claims that are not.
CCSS.ELA-LITERACY.RI.7.8	Trace and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient to support the claims.
CCSS.ELA-LITERACY.RI.8.8	Delineate and evaluate the argument and specific claims in a text, assessing whether the reasoning is sound and the evidence is relevant and sufficient; recognize when irrelevant evidence is introduced.

Informational texts for both forms were selected from the CCSS text exemplar list for grade 7. Form A used "The Evolution of the Grocery Bag," by Henry Petroski, an informative text using technical and theoretical language. Form A included 21 items, each written to align to one of the eight nodes identified for examination from the learning map. Form B used an excerpt from “The Story of the Montgomery Bus Boycott,” by Russell Freedman, a historical text. Form B included 17 items, each written to assess the cognitive skills described by one of seven nodes from the learning map. Form B did not include items which assessed students’ ability to identifiy qualifiers in claims. Items were reviewed for alignment to the content of the nodes for both forms. Since different texts were used for the two forms, items differed in their content and focus across the forms, but were constructed to measure the same skills, with the exception of the node that addressed the ability to identify qualifiers in claims. Kansas teachers of English language arts or social studies were invited to use either or both forms as a part of their regular instruction.

Participants

There were 235 students who took Form B. There were 24 Grade 6 students, 201 Grade seven students, and 10 Grade 8 students. There were 384 students who took Form A. There were 74 Grade 6 students, 220 Grade 7 students, and 90 Grade 8 students.

Results

For Form A, n = 384 M= 10.83 correct, *SD* = 2.96. For Form B, n = 235, M = 9.57 correct, *SD* = 4.28. To evaluate the measurement properties of both forms, a Structural Equation Modeling (Muthén & Muthén, 1998; 2011) approach was used. The structure tested was informed by the existing learning map nodes and connections, but when convergence was not achieved, a simplified model was tested, in which adjoining nodes were viewed as components of a single factor rather than distinct factors. Twelve test items were used in two factor model representing Form B, while 13 items produced a 3 factor model for the Form A. As additional items are added, a more nuanced model may be possible. The grouping of items by factors is presented in Table 2.

Table 2

Form A		Form B	
Factor 1			
Item	Loading	Item	Loading
ITEM3	1.000	ITEM9	1.000
ITEM1	0.680	ITEM1	0.666
ITEM4	0.837	ITEM2	0.523
ITEM19	0.905	ITEM4	0.512
ITEM21	0.668	ITEM6	0.821
		ITEM8	0.953
		ITEM3	0.721
		ITEM13	0.524
Factor 2			
ITEM9	1.000	ITEM12	1.000
ITEM6	0.809	ITEM5	0.549
ITEM7	0.706	ITEM7	0.845
ITEM8	0.565	ITEM14	0.912
ITEM20	0.454		
Factor 3			
ITEM15	1.000		
ITEM12	0.757		
ITEM14	0.542		

Form A, the longer form, contained sufficient numbers of items to constitute a third factor, but the factors identified by the analysis did not match the cognitive content represented in the learning map. Each of the three factors that emerged from the analysis contained multiple nodes and no apparent pattern of similarity in question content. On the other hand, Form B included factors aligned to constructs of claim identification, and identification of information that supports an author’s claim. Each factor was initially identified theoretically, and then evaluated quantitatively. Theoretically identified factors that did not prove to be distinct quantitatively were combined.

Factor 1 in Form B included items written to align to three nodes from the identified section of the learning map. Items 1, 2 and 3 were written to assess “can identify an author or speaker’s claims.” Items 4 and 6 were written to assess “can identify details that may be used to defend a claim.” Items 8, 9 and 13 were written to assess “can identify evidence for a claim.” Factor 2 in Form B included items written to align to 4 different nodes from the proposed learning map structure. Item 5 was written to assess “can identify details that may defend a claim.” Item 7 was written to assess “can identify evidence for a claim.” Item 12 was written to assess “Locate the topic sentence and supporting details in a paragraph.” Item 14 was written to assess “Can locate sentences supporting the central idea or a claim in an informational text.”

Form A contained items for a node that was not included in the items for Form B. Those items assessed students’ ability to identify qualifiers in a claim. In Form A, Items 10 and 11 were answered correctly less often than would be expected by chance. Item 12, which assessed the same node, saw students answer correctly at a rate only slightly greater than chance.

Discussion

The preliminary analysis presented here does not provide evidence in support of the learning map structure used to develop the two forms used in this study. Items that loaded in factors in the two forms did not match specifications for the node contents that were used to develop the items. The factors found for Form A were not distinct based on the specific items aligned to each factor. On the other hand, the factors for Form B are distinct and do represent the skills associated with argumentation. The first factor covered the identification of the claims, while the second factor contained identifying the details and evidence supporting the claims.

However, there was overlap between the factors in the items covering these skills. In summary, the results suggest that the skills represented by the nodes in the learning map as measured by the items are not distinct enough to warrant separate and clear factors.

A number of elements could have caused the factors corresponding to the skills associated with argumentation to be less defined. The unexpected patterns that emerged from the analyses could be the result of a smaller number of actual nodes than were present in the hypothesized structure. It is also highly possible that we do not have sufficient numbers of well-aligned items to provide the nuanced differentation between skills represented in the nodes. More items would have provided more data with which to find a clearer distinction between the factors and the skills involved in argumentation. Additionally, the items could also range in difficulty, so a better understanding of the extent to which the students can identify the claims and evidence would be possible.

Another potential reason for the observed lack of variance is the limitation imposed by using only two texts as the source material for evaluation of student ability to comprehend claims and arguments. Additional texts could be used with varying structures. Texts with a clear argument structure as indicated by linguistic signals could be used along with more difficult texts with a less explicit argument structure. Students might be better able to identify the claims and evidence for the explicitly structured texts than they would for the text with the more inferred structure. These texts might have allowed for more distinct factors associated with the skills involved in argumentation.

A final reason for the less transparent factors for the two forms could revolve around the argumentation development. Children use arguments in their spoken language before being able to understand them in written text. At seven years, students can produce a basic argument

(Golder & Coirier, 1994; Miller, 1986, 1987; Stein & Miller, 1993) a claim and a supporting reason (Coquin-Viennot & Coirier, 1992; Golder & Coirier, 1994). At twelve years, they include more reasons to their arguments and indicate potential counter-arguments (Golder & Coirier, 1994). Despite their progression in argument production, argumentation continues to develop throughout their education and extends into university (Golder & Coirier, 1994; Knudson, 1992; 1994).

Argumentation requires higher-order processing and reasoning abilities, and it typically develops later than most other reading-related processes. Students begin to understand the global text structure of arguments and develop schemata regarding them in the fifth grade (Chambliss & Murphy, 2010). Although middle and high school students can recognize a clear argument (Chambliss, 1995; Chambliss & Murphy, 2002; McCann, 1989), they have not acquired an adequate understanding of the components associated with arguments (Chambliss, 1995; Chambliss & Murphy, 2002). Only 12% of high school seniors are proficient in comprehending arguments and their elements (National Assessment of Educational Progress, 1996). Additionally, college students have difficulty identifying the elements of arguments with less explicit structures. As a result of this difficulty, they also have problems identifying the reasons and evidence supporting the claims, and they sometimes consider counterclaims as one of the claims made in the argument (Larson, Britt, & Larson, 2004). Because sixth, seventh, and eighth grade students are just starting to acquire argumentation, they might not have enough knowledge of arguments to identify the claims and supporting evidence, causing the ambiguous factors.

Since both forms were constructed for use as a part of a formative assessment tool, it is presumable that many students took the tests for a formative purpose. The scores on both forms indicate the possibility that students had not received instruction on the concepts being tested.

With such low mean scores, it is difficult to justify using the results of this study as evidence for or against the structures represented in the learning map that were used to produce both test forms. At this time both forms remain available to teachers. At the end of the school year, some teachers may use the form that their students have not taken as an opportunity to assess student learning since they took the first form in January or February.

In summary, items were created to assess the area of the learning map representing the skills associated with argumentation. Little evidence was found that the skills are related in the same way as they are in the learning map for middle school students. The derived factors from the analysis were ambiguous but suggestive of two factors on one of the forms: the identification of claims and identification of the evidence supporting them. The number of items and texts and the nature of argument development might have influenced these findings. Future research could extend this study by using more items and text and by testing it on both middle and high school students.

References

Beach, R. W (2011). Issues in Analyzing Alignment of Language Arts Common Core Standards With State Standards. Educational Researcher, 40, 179-182.

Bechard, S., Hess, K., Camacho, C., Russell, M., & Thomas, K. (2012). *Why should cognitive learning models be used as the foundation for designing next generation assessment systems?* Symposium 2011 Topic 1 White Paper. Menlo Park, CA, and Lawrence, KS: SRI International and Center for Educational Testing and Evaluation (CETE).

Blair, J. A., & Johnson, R. H. (1987). Argumentation as dialetical. *Argumentation, 1,* 41–56.

Bulgren, J. A. & Ellis, J. D. (2012) Argumentation and evaluation intervention in science classes: Teaching and learning with Toulmin. In M. S. Khine (Ed.), *Perspectives on Scientific Argumentation: Theory, Practice and Research,* (pp. 135-154). New York: Springer Science+Business Media.

Chambliss, M. J. (1995). Text cues and strategies successful readers use to construct the gist of lengthy written arguments. *Reading Research Quarterly, 30*(4), 778–807.

Chambliss, M. J., & Murphy, P. K. (2002). Fourth and fifth graders representing the argument structure in written texts. *Discourse Processes, 34*(1), 91–115.

Chambliss, M. J., & Murphy, P. K. (2010). Fourth and fifth graders representing the argument structure in written texts. *Discourse Processes, 34*, 91-115.

Common Core State Standards Initiative (CSSS) (2010). Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects. Retrieved from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Conley, D.T. (2008). *College knowledge: what it really takes for students to succeed and what we can do to get them ready*. New York, NY: Wiley.

Conley, D. T. (2011). Building on the Common Core. *Educational Leadership*, *68*(6), 16-20.

Coquin-Viennot, D., & Coirier, P. (1992). Structures of argumentative discourse: Effects of type of referential space. *European Journal of Psychology of Education*, *7*(3), 219–229.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, *88*(6), 915-933.

Freedman, R. Freedom Walkers: The Story of the Montgomery Bus Boycott. New York: Holiday House, 2006.

Golder, C., & Coirier, P. (1994). Argumentative text writing: Developmental trends. *Discourse Processes, 18*, 187–210.

Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.

Hess, K. K. (2012). *Learning progressions in K-8 classrooms: How progress maps can influence classroom practice and perceptions and help teachers make more informed instructional decisions in support of struggling learners*. Retrieved from http://www.cehd.umn.edu/NCEO/OnlinePubs/Synthesis87/SynthesisReport87.pdf

Kingston, N. & Erickson, K. (2010). *Response to the Technical Assistance on State Data Collection IDEA General Supervision Enhancement Grant Alternate Academic Achievement Standards, CFDA 84.373X*. Available at http://dynamiclearningmaps.org/pdf/DLMNarrative.pdf

Klahr, D., Fay, A.L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*, 111–146.

Kober, N., & Rentner, D. S. (2011). States' progress and challenges in implementing common core standards. Center on Educational Policy. Retrieved from http://www.cep-dc.org/displayDocument.cfm?DocumentID=391

Koslowski, B., Okagaki, L., Lorenz, C., & Umbach, D. (1989). When covariation is not enough: The role of causal mechanism, sampling method, and sample size in causal reasoning. *Child Development, 60*, 1316–1327.

Knudson, R. E. (1992). The development of written argumentation: An analysis and comparison of argumentative writing at four grade levels. *Child Study Journal*, *22*, 167–184.

Knudson, R. E. (1994). An analysis of persuasive discourse: Learning how to take a stand. *Discourse Processes*, *18*, 211–30.

Kuhn, D. (1993). Science argument: Implications for teaching and learning scientific thinking. *Science Education*, *77(3)*, 319–337.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. New York, NY: Academic Press.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monograph of Social Research and Child Development, 60*, 1–127.

Larson, M., Britt, M. A., & Larson, A. A. (2004). Disfluencies in comprehending argumentative texts. *Reading Psychology*, *25*, 205-224.

Latour, B., & Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. (2nd ed.). Princetown, NJ: Princetown University Press.

Lee, J., Tidwell-Scheuring, S., & Barton, K. (2006). Online assessment and instruction using learning maps: A glimpse into the future. In D. D. Williams, S. L. Howell, & M. Hricko (Eds.), *Online assessment, measurement and evaluation: Emerging practices*, (pp. 227-237). Hershey, PA: Information Science Publishing.

Marttunen, M. (1994). Assessing argumentation skills among Finnish university students. *Learning and Instruction*, *4*, 175–191.

McCann, T. M. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English*, *23*, 62-76.

Miller, M. (1986). Argumentation and cognition. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 225–249). San Diego, CA: Academic Press.

Miller, M. (1987). Culture and collective argumentation. *Argumentation, 1*, 127–154.

Muthén, L. K., & Muthén, B. O. (1998; 2011). *Mplus user's guide*, Sixth Edition. Los Angeles, CA: Muthén & Muthén.

National Academy Press (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Washington, DC.

National Assessment of Educational Progress. (1996). *NAEP 1994 U.S. history report card: Findings from the National Assessment of Educational Progress*. Princeton, NJ: Educational Testing Service.

National Research Council (2012) *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

Penner, D., & Klahr, D. (1996). The interaction of domain-specific knowledge and domain-general discovery strategies: A study with sinking objects. *Child Development, 67*, 2709–2727.

Petroski, H. “The Evolution of the Grocery Bag.” American Scholar 72.4 (Autumn 2003).

Pontecorvo, C. (1987). Discussing for reasoning: The role of argument in knowledge construction. In E. De Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and instruction: A publication of the European Association for Research on Learning and Instruction* (vol. 1, pp. 71–82). Oxford: EARLI.

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.

Popham, W. J. (2011). *Transformative assessment in action: An inside look at applying the process*. Alexandria, VA: ASCD.

Salinas, I. (2009, June). *Learning progressions in science education: two approaches for development*. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*, 102–119.

Stein, N. L., & Miller, C. A. (1993). The development of memory and reasoning skill in argumentative contexts: Evaluating, explaining, and generating evidence. In R. Glaser (Ed.), *Advances in instructional psychology 4* (pp. 285–335). Hillsdale, NJ: Erlbaum.

Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the Rule Space Method*. New York: Routledge.

Toulmin, S. E. (1958). *The uses of argument.* London: Cambridge University Press.

Toulmin, S., Rieke, R., & Janik, A. (1984). *An introduction to reasoning*. New York, NY: MacMillan.