The Impact of the Affinity Learning Authoring Tool on Student Learning

Leen-Kiat Soh*, David Fowler**, and Art I. Zygielbaum***

* National Center for Information Technology in Education, Department of Computer Science and Engineering, University of Nebraska, 256 Avery Hall, Lincoln, NE 68588-0115, USA. E-mail: lksoh@cse.unl.edu

** Department of Teaching, Learning and Teacher Education, University of Nebraska, 16 Henzlik Hall, Lincoln, NE 68588-0355, USA. E-mail: dfowler1@unl.edu

*** University of Nebraska, Lincoln, USA.

Abstract

Affinity Learning is a system that allows the user to build a lesson on a subject matter by breaking it down into concepts, misconceptions, assessments, and remediation steps. Examples and questions can also used in these components. Affinity Learning has been found to be effective and can offer critical insights to student learning strategies. Authoring Affinity Learning lesson plans or hierarchies, however, is non-trivial. We have developed two authoring tools: the first tool provides an overall view of the hierarchy with a *graphical display* of the nodes and links; but the second tool does not. This paper reports on a study conducted to test whether the graphical authoring tool can help produce better-quality hierarchies and also help the users learn about the subject matter better than the non-graphical authoring tool. Results show that while the graphical authoring tool can result in better-quality hierarchies, it does not result in better learning.

1.   Introduction

The Affinity Learning system is similar to, but distinguishable from, other electronic tutor development efforts. Many educational simulation, electronic tutor or virtual classroom programs, such as the IMMEX system (www.immex.ucla.edu) developed at UCLA, provide opportunities for students to participate in active learning. Affinity can accommodate these types of presentations and assessments. In addition, Affinity Learning system adds customized support to help students overcome misconceptions, missing knowledge, or missing skills. Like IMMEX, Affinity provides data allowing instructors to understand how students used the system and how well they learned. Affinity-based lessons in mathematics, geosciences, and medical education have shown positive learning achievement (Zygielbaum and Feese, 2005). The system can also distinguish student learning strategies – such as those used by male and female students (Zygielbaum and Grandgenett, 2001).

Affinity Learning was designed, from its earliest phases, to harness the knowledge of master teachers and to provide information that better enables them to meet the needs of learners. A number of educational design principles have guided Affinity Learning's development:

1. Leverage the knowledge and experience of teachers.

2. Target instruction to match student understanding.

3. Use technology to guide learning using real life data and applying real life problem solving techniques.

4. Create an ongoing understanding of student mastery through systematic embedded assessments.

5. Cultivate active learning through deep conceptual questions that encourage dynamic, critical thinking.

6. Enhance the relationship among teachers and students.

7. Provide a non-threatening environment that helps to clarify thinking processes, identify approaches to problems, and facilitates tests of potential solutions.

Affinity Learning is based on predicting the likely misconceptions of students. Misconceptions are detected through embedded assessments and resolved through remediation or augmentation learning nodes. Lessons are prepared by breaking down the topic to be taught into small learning experiences or "learning nodes." Lessons appear as a hierarchical network of interconnected nodes. Each node consists of a web-based presentation of material and an assessment of student understanding of the material. As a student completes one learning node, he or she is directed to a subsequent node based on the outcome of the assessment. The student is either directed to continue in the learning progression or diverted to an augmentation node. If the outcome is unanticipated or not covered by a succeeding node, the system notifies the faculty member and the student. The faculty member then enhances the lesson by adding new node(s) and interconnections to better serve the student (Bruning, Zygielbaum and Grandgenett, 2001).

Students' progress through the network of learning nodes is monitored by the web-serving computer. Each node outcome is assigned a weight by the author that indicates the educational value of that response. A large weight denotes an element critical to understanding the overall topic. A small weight indicates ancillary information. If the instructor wishes to track learning efficiency (progress versus time), he or she may assign negative weights to particular outcomes. Data collected include the order of node visitation, the related sequence of outcome weights, time spent on each node, and the results of assessments. These data are presented graphically to the instructor.

The complexity of authoring continues to be a challenge as Affinity Learning is applied in additional teaching and learning environments. Establishing the structure of Affinity Learning lessons takes two major steps: Decomposition and Interconnection. The instructor must first decompose a lesson into the set of learning nodes. At the same time, he or she must anticipate potential misconceptions and add nodes to respond to the misconceptions. The initial decomposition has proven a significant challenge for new authors. It is very difficult for even experienced STEM educators to decompose their lessons into the small "learning nodes" that drive Affinity Learning. Teachers have automated their teaching skills through experience. Significant effort must be expended to capture the steps they take, how they recognize misconceptions and how they apply prescriptions.

The current authoring approach involves creating the course in the classical order of presentation. Following the research of Minstrell, it has been suggested that the process be inverted; the author first enumerating misconceptions and then constructing learning nodes and structure to avoid them (Minstrell, 1989; Kraus and Minstrell, 2003). Misconceptions could be identified through instructor experience, focus groups, or experimentation with students. The second step, establishing an interconnected network for the nodes, also presents challenges for all but simple subjects. As the breadth and difficulty of the topic covered increases, the network

interconnecting the learning nodes becomes increasingly complex. Keeping track of all of the nodes, and the conditions that lead to their selection, has been shown to be a formidable task as instructors have difficulty maintaining a mental image of the network of nodes, the purposes of interconnections, and the network structure in general.

In this paper, we report on an evaluation of two versions of the Affinity Learning authoring tools: graphical and non-graphical. In the graphical version, the users are presented with a graphical display of the nodes and links created so far. On the other hand, the non-graphical version is basically text-based with no visual display of the Affinity Learning hierarchy. We conducted a study in a course of 30 students in mathematics education—learning how to teach secondary mathematics. We used a pre- and post-test setup to measure the learning gains of the treatment and control groups. We also collected data on other criteria to measure the quality and mechanics of the Affinity Learning hierarchies produced. Results show that while the graphical authoring tool can result in better-quality hierarchies, it does not result in better learning.

In the following, we first discuss related work in terms of authoring tools and our focus in misconceptions in algebra in Section 2. Section 3 describes the two Affinity Learning authoring tools: graphical and non-graphical. Then, we describe our study setup in Section 4. We present and discuss the results and our hypotheses in Section 5 before concluding our paper.

2. Related Work

2.1. Authoring Tools

Because of the advances and popularity of intelligent tutoring systems (ITSs), there has been research in authoring tools for these systems (e.g., Virvou and Moundridou 2000, Dabbagh 2002, Brusilovsky 2003).

Murray (1998, 1999, 2003) reported authoring tools on design issues and tradeoffs and the knowledge acquisition and representation tissues. Murray also provided a seven-part categorization of a handful of authoring systems for ITS: for curriculum sequencing and planning, tutoring strategies, device simulation and equipment, and so on. The Affinity Learning authoring tool in this paper is for the authoring of tutoring strategies in the form of affinity learning hierarchies. Moundridou and Virvou (2002) also pointed out that the construction of an authoring tool is associated with many problems, such as the generality of the techniques incorporated, domain-independence, effectiveness for the prospective authors (instructors), and effectiveness for the students who will use the resulting ITSs. In their studies, Moundridou and Virvou (2002) found that the most important feature of a generically applicable authoring tool is an instructor modeling component that supports adaptivity. In EASE, Aroyo et al. (2004) also touched upon an authoring task ontology that at a meta-level defines and controls the configuration and tuning of an authoring tool for a specific authoring process.

Van Rosmalen et al. (2006) further introduced a standards-based model for adaptive e-learning and investigated the conditions and tools required by authors to implement this model. They defined adaptation in the context of e-learning as creating a learner experience that purposely adjusts to various conditions over a period of time with the intention of increasing pre-defined success criteria, and that each of the conditions for adaptation has to be represented in a rigorous way. Thus, it might be critical for authoring tools to provide such support to authors.

Major et al. (1997) proposed that for authoring tools to be truly effective, the environments must be based on psychologically plausible account of teaching. To illustrate this, the REDEEM software suite was used to exploit the symbiotic relationship between psychology and authoring environments. REDEEM (Ainsworth et al. 2000, 2003) is an ITS authoring environment that creates simple ITSs by overlaying teaching expertise on existing domain material. Studies have shown that the REDEEM environment can cater for different instructional goals (Ainsworth et al. 2000).

In our Affinity Learning tool, our authoring task ontology is basically the terms of concepts, misconceptions, assessments, and remediation steps. The authoring tool is also domain-independent and supports adaptivity. However, presently, the authoring tool does not yet cater for different instructional goals as in REDEEM. Right now, the authoring tool is specific to the affinity learning approach – teaching through the use of misconceptions. In addition, in the literature, though the evolution of authoring tools has been discussed, the process of designing an authoring tool has been investigated, and the impact of the tools on knowledge acquisition and representation has been studied, the impact of such tools on the users who use them to design ITS is not clear.

The Affinity Learning authoring tool has the potential to be possibly an effective "learning tool" because it deals with concepts and misconceptions in a logical manner that solicits and then *maps* critical insights from the authors. Indeed, such concept maps have been viewed as a learning tool, and not as a teaching mechanism (Novak 1998). The notion of concepts and misconceptions has also been previously incorporated into ITSs (e.g., the Algebrator of (Jurkovic 2001), and the Buggy system of (Brown and Burton 1978)).

2.2. Misconceptions in Algebra

In this discussion, we will take *algebra* to mean the subject as usually represented in secondary school mathematics classes. Within that domain, we will make an informal distinction between procedural misconceptions and conceptual misconceptions. In doing so we set aside many implicit misconceptions about algebra that may be held even by teachers, including teachers at the college level. As an example of an implicit misconception, consider the tradition of using single letters to represent a variable, so that any combination of two letters, such as "tz" would be considered the product of "t" and "z".

An example of a conceptual misconception, consider a student who is asked to graph the velocity of a cyclist climbing a hill and then descending, with time on the horizontal axis. The student erroneously believes that the graph must represent the image of a hill (Clement 1989). As an example of a procedural misconception, we offer the notorious *Zorro cancellation* error (Figure 1), particularly troublesome to students because it can lead to a correct result when $x = 1$.
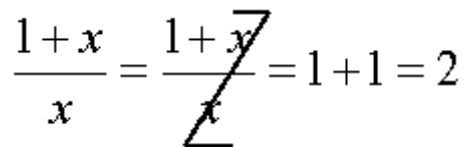
$$\frac{1+x}{x} = \frac{1+\cancel{x}}{\cancel{x}} = 1+1 = 2$$

Figure 1. "Zorro cancellation."

Misconceptions in algebra arise in part from the conventions of two-dimensional mathematical notation. This notation is itself ambiguous and dependent upon context and historical convention. As an example, consider this simple definition of a function of two variables (Figure 2). The experienced algebra teacher knows that (1) f represents a function, since f always stands for function. (2) The letters x and y are used for variables, by the same convention. (3) the number 2 is a factor that will be distributed across terms to the right. (4) The superscript 2 is an exponent indicating the squared product of the expression in parenthesis. Convention dictates that this product must be expanded before the factor 2 can be distributed. Finally, the parenthesis in (5) is used to group the arguments of the function, while the parenthesis in (6) is used to group the elements of the algebraic expression. Depending upon context, these parentheses follow completely different rules.

$$\underset{5\ \ \ \ \ \ \ \ \ \ \ \ \ \ 6}{\overset{1\ \ \ 2\ \ \ \ \ \ \ 3\ \ \ \ \ \ \ \ 4}{f(x, y) = 2(x + y)^2}}$$

Figure 2. Misconceptions from the conventions of two-dimensional mathematical notation.

Misconceptions can also arise from the traditional communication of mathematics through hand-written media. The distinction between the sign used to represent a negative number and the symbol for subtraction may be made clear in a typeset book, but to students in school, they are the same symbol, again depending upon context. The customary abuse of language in using these symbols leads to students saying things like "minus 2 minus minus 2," which can quickly lead to the ambiguous statement "a minus and a minus is a plus." (The distinction between the typeset minus, the hyphen-minus, and the en-dash may confuse even non-TEX using mathematicians.) Finally, mathematics must always be written. The number of mathematicians who can carry on an extended symbolic discussion in purely oral mode is probably smaller than the number of mathematicians who can play blindfold chess.

Given these and other systemic sources of confusion in mathematics (e.g., see (Lakoff and Nunez, 2000) for extensive analysis), mathematics is still a workable language, and many students do learn to abide by its conventions. Unfortunately, many students don't. These students will stumble through high school algebra, somehow pass a school graduation requirement, and then go on to college algebra, where failure rates—"F or D or withdraw" (FDW) rates—are estimated to be in the 40% to 60% range (Small 2007).

Clearly then, mathematics teachers must learn to identify and remediate misconceptions in their students. A number of studies have been made to throw some light on the problem of student misconceptions in algebra. The study of misconceptions about the use of variables has received particular attention, both from classroom teachers and university researchers (e.g., Rosnick 1981; Fisher 1998; Dede 2004; Selitto 1997). Although general categories of misconceptions, including the study of reasoning "bugs" (VanLehn 1990), teachers must develop

their own sense of when a student holds a conceptual or procedural misconception. Some misconceptions are general and well known, while some may be idiosyncratic, such as the anecdote about the student who thought a leading zero on a decimal representation required a second decimal point, because he had seen the phrase "...we also write 0.3."

Preservice teachers prior to their field experiences may not be aware of the problems of misconceptions, unless they have experience as a remedial tutor. Students planning to become secondary mathematics teachers were usually good math students when they were in high school, and they have a well-rehearsed set of algebraic skills from hours of calculus courses. When these pre-service teachers have their first practicum experiences, they are often surprised by the difficulties they encounter in students. Lacking systematic knowledge of misconceptions, the pre-service teachers must improvise remediation for misconceptions as these errors are encountered. Preparation for working with student misconceptions is obviously highly desirable.

We studied the understanding of misconceptions and remediation in pre-service teachers who had not yet gone through formal field experiences. Our approach used the Affinity Learning system, and required the students to imagine how they would test for misconceptions and plan remediations within the Affinity Learning system.

3. Two Designs: Graphical and Non-Graphical

Affinity Learning helps the instructor in setting up a structured curriculum that is adjusted for individual students depending on the feedback obtained from assessments on learning material. This is accomplished through nodes, and the assessment based paths guiding through a network of nodes to ensure that students achieve specified level of competency in every topic.

Activity Builder is developed to support the creation of Affinity Learning activity modules offline. Each activity is represented as a network of nodes and stores the activity in a widely adopted standard format (XML) so that the content can be reused and transferred to the existing Web-based system.

Table 1 lists the terms and their definitions used in Affinity Learning.

Table 1. Terms and definitions used in Affinity Learning.

| Terms	Definitions
Activity	An activity is a network of *nodes* to allow personalized learning, with each node containing some instructional content about the topic.
Node	A node is the basic component of the activity that contains some instructional content like text, images, etc., covering a brief topic and typically a short *assessment* to test the student's competency in the provided material.
Link	A link is a direct path from one node to another node in the activity. Usually, a link leads to a node that contains the next part of the topic. The node at which link originates is called the *source* node and the node at which link ends is called the *destination* node. A link also typically includes a *label* that is the text that is displayed on the screen to represent the link for the student.
Instructional Elements	The educational materials which make up a lesson on a node. Each node in the activity may or may not contain some instructional material. Material such as text, image, audio or video can be embedded in a node.

Assessment	If an instructor decides to test the competency of a student based on the material provided as part of the node, assessment is included. Assessment is made up of questions and answers. Types of questions include *multiple choice*, *true-false*, *short answer*, *fill in the blank*, and *multiple select*.
Outcome	An outcome provides a connection between nodes based on student responses to an assessment. They are intended to represent the student's conception of the material in the assessment and may be as simple as "right" and "wrong". An outcome prescribes a node which the student will be taken to next. Outcomes also have a progress level adjustment value that can be used to indicate the progress made by the student by reaching a given outcome.
Responses	Define the circumstances under which an outcome is reached. They represent specific answers to assessments. Several responses will often be used for an outcome.

3.1.   Non-Graphical

In the non-graphical authoring tool, the user traverses from one component to the next on his or her own, *without a visual display of the overall view of the hierarchy built so far*. Figure 3 shows the front graphical user interface (GUI) of the tool, with options for the user to set up different components of the hierarchy. Figure 4 shows how the user adds and insert a node to the hierarchy. This version of the Activity Builder is essentially "text-based", where the user needs to, for example, draw their own hierarchy and keep track of the nodes, links, outcomes, etc. on their own.
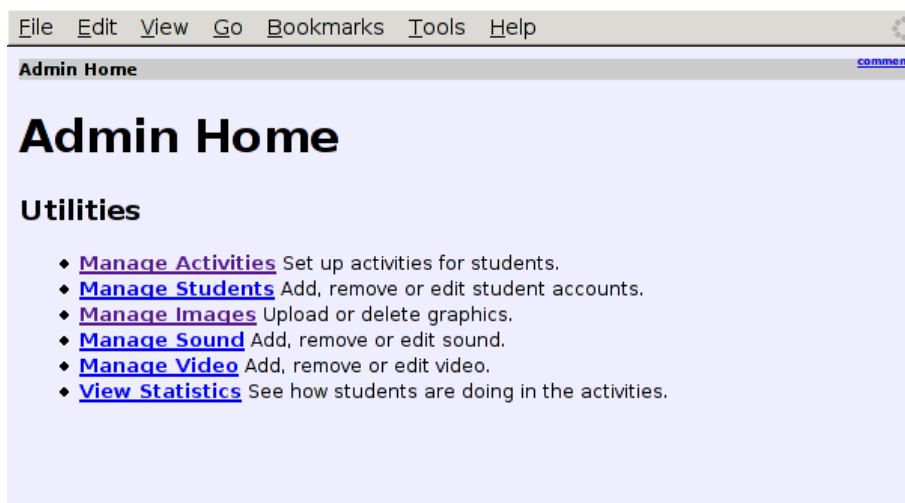
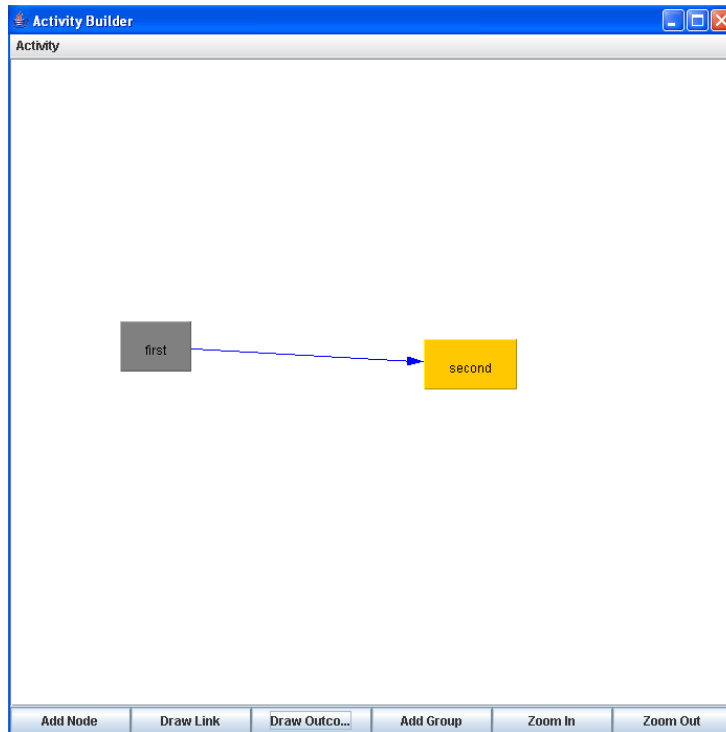**Figure 3.** The front graphical user interface (GUI) of the non-graphical Activity Builder tool.
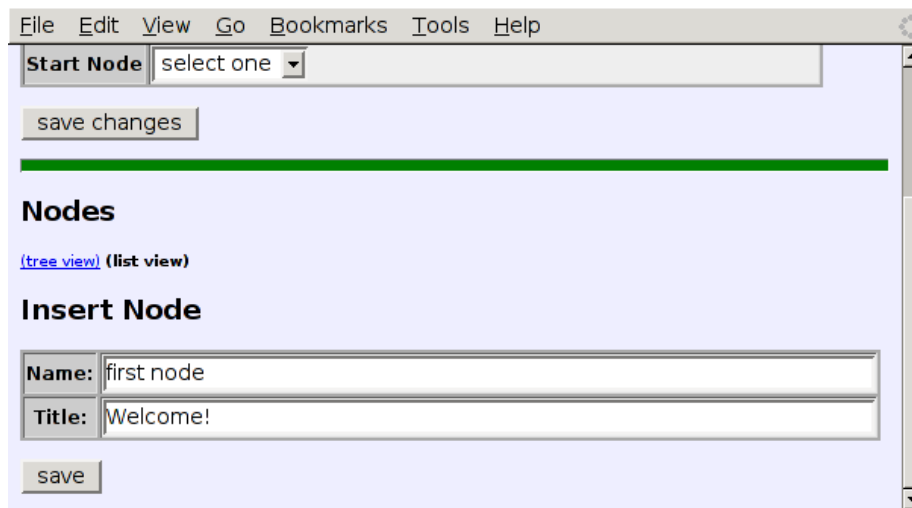
Figure 4. An example: The GUI for the user to add and insert a node to the hierarchy using the non-graphical Activity Builder tool.
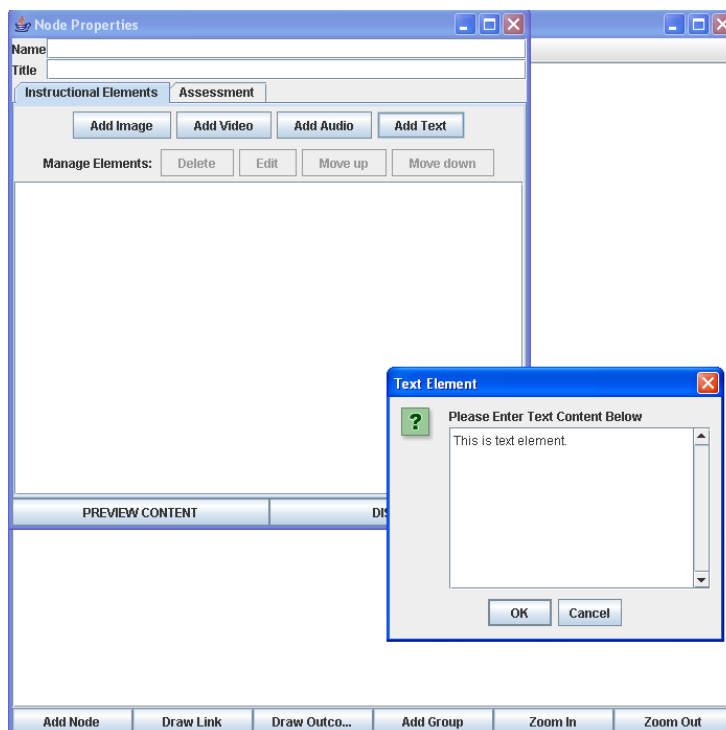
3.2. Graphical

In the graphical authoring tool, the user builds an Affinity Learning hierarchy by clicking and dragging nodes on a digital whiteboard. The user may connect two nodes with a link, click on the link to add details, and so on. Thus, there is a *visual display of the overall view of the hierarchy built so far*. Figure 5 shows the graphical user interface (GUI) of the tool, with "button" options for the user to create and annotate nodes, links, and outcomes. When the user clicks on each component, the system pops up a new window that solicits the user's entry that specifies the component. For example, Figure 6 shows the pop-up window when the user clicked on the "first" node. This window allows the user to enter instructional elements and assessments. An instructional element could be as simple as a few lines of text explaining the concept represented by the node. As a result, the graphical version helps the user organize the nodes and relationships among nodes.

Figure 5. The graphical user interface (GUI) of the graphical Affinity Builder system, with two nodes and a link created.

Figure 6. The pop-up window when a node is clicked in the graphical Affinity Builder system.

4.   Study Setup

4.1. Students and Groups

Participants were undergraduate students in the TEAC 451P course at the Department of Teaching, Learning, and Teacher Education (TLTE) at the University of Nebraska. These students were being trained to be teachers in secondary school mathematics.

The students were randomly assigned to one of the two groups. In the *treatment* group, students used the graphical Affinity Learning authoring tool. In the *control* group, students used the non-graphical authoring tool. There were 30 students altogether and 15 students in each group.

4.2. Steps

Step 1: The instructor of the TEAC 451P course announced and presented the study. Please see Appendix A for the "Technology Project Description" given to the students in class. This step prepared the students for the study.

Step 2: Students were given a 30-minute pre-test (see Appendix D for the pre-test, which was also used as the post-test in Step 4 later). The test asked students to come up with an inventory on student algebra misconceptions. In the pre-test, students were asked to come up with common misconceptions, the correct solutions, and the remediation steps as well.

Step 3: A week after the pre-test, the 2-hour laboratory exercise took place where students used the Affinity Learning authoring tools to come up with an inventory on student polynomial misconceptions. The exercise was preceded with a "pep talk" by the course instructor. See Appendix B for the handout on the project. Then, the students moved to two separate computer rooms according to their assigned groups. Each room had a technical helper familiar with the use of the Affinity Learning tools. Each student was also given a user manual of the Affinity Learning authoring tool to be used. The technical helper also gave a 15-minute step-by-step demonstration in the beginning of the laboratory exercise to acclimate the students to the tool and the terms used in Affinity Learning. At the end of the exercise, the students were asked to fill out a questionnaire (see Appendix C for the user feedback questionnaire).

Step 4: A week after the laboratory exercise, the students were given a 30-minute post-test. Once again, the test asked students to come up with an inventory on student algebra misconceptions. It was the same test as that used in Step 2 above.

The pre-tests, post-tests, and the saved Affinity Learning hierarchies (outcomes of the laboratory exercise) were graded by two graduate students in mathematics.

Note that the students were tested on algebra while the laboratory exercise was on polynomials. This was designed to eliminate the content impact of the laboratory exercise so that we could more accurately measure the direct effects on coming up with misconceptions and remediation steps.

4.3. Hypotheses and Outcome Measures

Hypothesis 1: The graphical Affinity Learning authoring tool helps the users develop better Affinity Learning hierarchies than the non-graphical Affinity Learning authoring tool.

The outcome measures used for testing Hypothesis 1 included the following criteria deemed to represent the quality of an Affinity Learning hierarchy, as shown in Table 2.

Table 2. Criteria used in computing the quality of an Affinity Learning hierarchy.

Criteria	Description
Verbosity	Whether the content on each node was verbose
Correctness	Whether the content was technically correct (e.g., correct definition of polynomials and correct solutions)
Insights	Whether the content included insights to pitfalls and common misconceptions, and tips to understand a concept
Feedback	Whether the feedback on a student response serves to reinforce student understanding or correct student misconceptions
Breadth	Whether the content covered sufficient subareas in polynomials
Depth	Whether the content covered sufficient depth on polynomials
Flow	Whether the nodes in the hierarchy has a logical flow (e.g., from easy to more difficult subareas, from definitions to examples, etc.)
Modularity	Whether different issues ore subareas were discussed in different nodes
Chart flow	A technical criterion, whether the nodes were linked correctly

We also used the criteria shown in Table 3 to compute the quality of the mechanics of the hierarchies, more as supporting measures to see whether there is noticeable difference between the two sets of hierarchies.

Table 3. Criteria used in measuring the mechanics of an Affinity Learning hierarchy.

Criteria	Description
#nodes	The number of nodes
#elements	The number of elements altogether
#links	The number of links used to connect the nodes
#outcomes	The number of outcome nodes
#questions	The number of questions posed
longest path	The length of the longest path in number of nodes
#orphaned nodes	The number of nodes that do not have an incoming link
#incomplete links	The number of links that do not lead to another node
#incomplete outcomes	The number of outcomes that do not a student response to a question

Hypothesis 2: The graphical Affinity Learning authoring tool improves the users' understanding of a subject matter more significantly than the non-graphical Affinity Learning authoring tool.

The outcome measures were the pre-test and post-test scores. For each test score, we also break down the total score into four sets of subscores to provide a more detailed look: examples, solutions, misconceptions, and remediation steps.

5. Results

Here we first present the data and results that we collected from our study in Section 5.1. Then we discuss the results with respective to the two hypotheses described in Section 4.3 above.

5.1. Presentation and Observations of Results

Pre-Test and Post-Test Scores. Figure 7 shows the pretest and post-test scores, broken down to the four subsets of items. General observations are:

- In both groups, student improved from before to after the laboratory exercise (52.39% to 60.45% in the treatment group; and 45.70% to 58.80% in the control group). The students in the control group seemed to improve more significantly on average. In particular, the students in the control group seemed to improve more significantly than students in the treatment group in the quality of remediation (44.72% to 50.61% in treatment, and 33.56% to 47.07% in control). This could be because with the graphical support, the students might have to concentrate on and visualize the hierarchy on their own, leading to better learning.

- Both groups improved in terms of the quality of the examples, solutions, misconceptions, and remediation steps that they came up with before and after the use of online tools, in terms of the solutions. This indicates that the tool might have helped their understanding of the subject matter.

- The biggest improvements are in the quality of examples (43.56% to 56.06% in the treatment group, and 33.52% to 50.27% in the control group) and the quality of misconceptions (54.72% to 66.06% in the treatment group, and 51.12% to 64.49% in the control group). We see that this is a direct result of the laboratory exercise where students were sort of "required" to come up with these items. That motivated them to cognitively process the examples and misconceptions better.
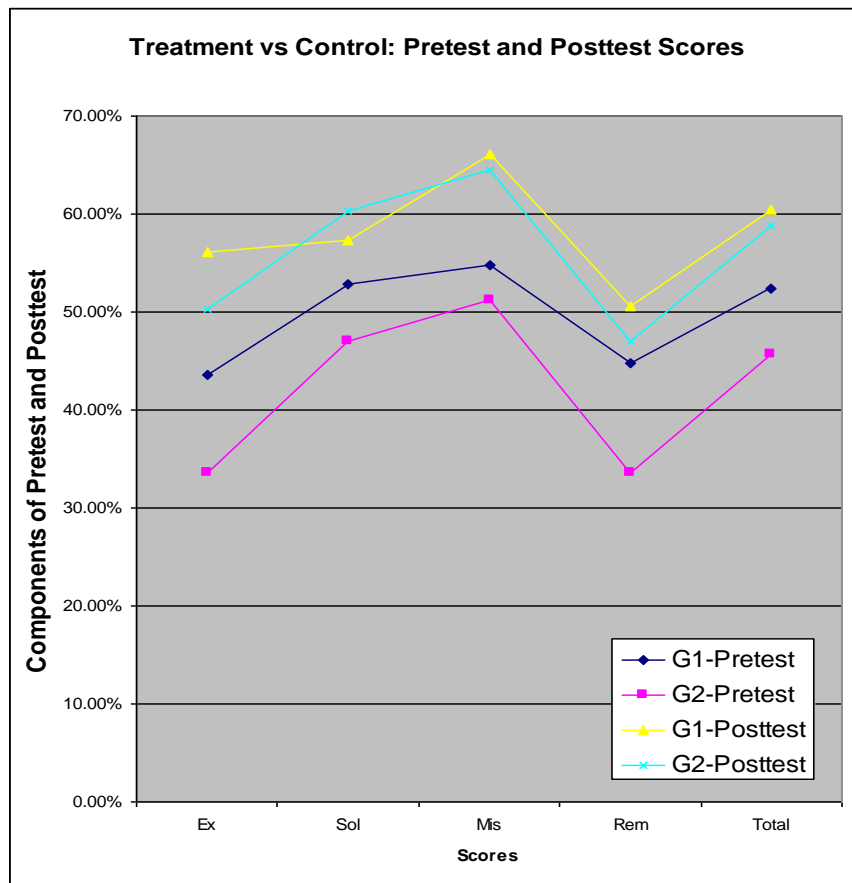
Figure 7. Ex: examples, Sol: solution, Mis: misconceptions, Rem: remediation, and Total: the total scores. G1: Treatment group, using graphical affinity. G2: Control group, using non-graphical affinity.


Individual Algebraic Subtopical Analysis. The pre-test and post-test were the same. There were 16 subtopics altogether as shown in Table 4. Each topic had a total score of 10 points, with 2.5 points for each of its four components: example, solution, misconception, and remediation. Comprehensive statistical analyses yielded neither significant patterns nor trends.

Table 4. Subtopics of the test on algebra.

ID	Subtopics	Info
1	numerical fractions – addition	Example, solution, misconception, remediation all provided
2	symbolic fractions – addition ($\frac{1}{a} + \frac{1}{b}$)	Example provided
3	symbolic fractions – addition ($\frac{b}{a} + \frac{a}{b}$)	Example provided
4	numerical fractions – division	
5	translating a sentence into symbolic form	Example provided

|		
6	Evaluating simple expressions of numbers raised to powers	
7	Evaluating simple expressions of square roots of numbers	
8	Algebraic manipulations. Simplify an expression with both + and − signs	Example provided
9	Expressions with zero	
10	The Distributive Law	
11	Simplifying expressions using inequalities	
12	Identifying polynomials	
13	[Student choice]	
14	[Student choice]	
15	[Student choice]	
16	[Student choice]	

Quality of Affinity Learning Hierarchies. As outlined in Table 3 earlier, we graded the quality of the Affinity Learning hierarchies along nine criteria. We used a scale of 0-5, with 5 being the maximum score. Figure 8 shows the average values for the two groups. Once again, we see no significant difference between the two groups. The treatment group seemed to perform better in terms of correctness, insights, feedback, flow, modularity and chart flow; while the control group seemed to perform better in terms of the breadth and depth of the hierarchies they produced.
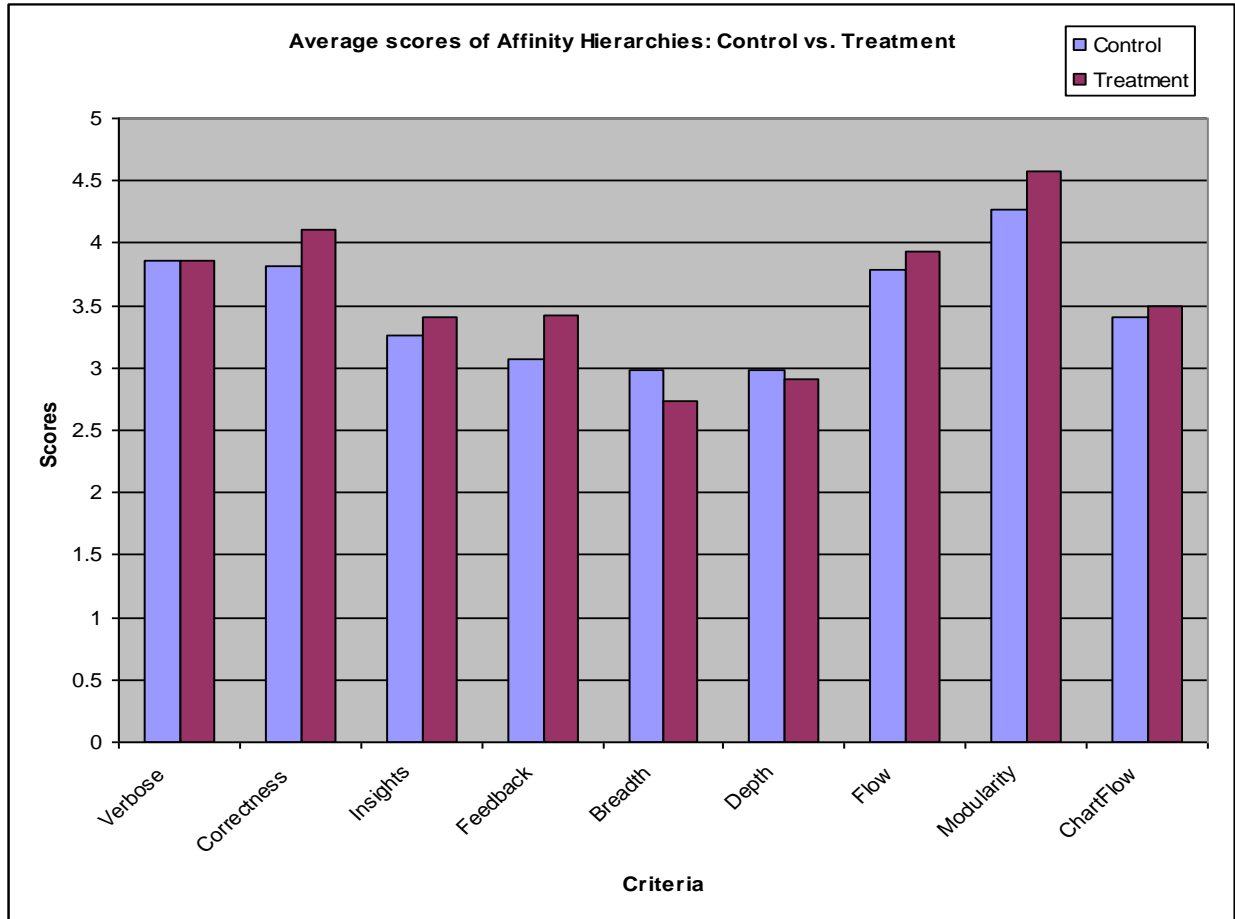
Figure 8. Average scores of the quality of Affinity Learning hierarchies produced: control vs. treatment groups

Mechanics of Affinity Learning Hierarchies. We also looked at the mechanics of the affinity learning hierarchies using the criterion described in Table 3. Figure 9 shows the average measurements for the two groups. Though no statistically significant differences were observed, we see that:

- In general, students in the treatment group included more nodes, elements, and questions than students using the non-graphical tools.

- In general, students in the control group included more links and outcomes, but also incurred more errors, in terms of the number of orphaned nodes, incomplete links and incomplete outcomes. We see this as a direct result of using the non-graphical Activity Builder system as opposed to the graphical version.
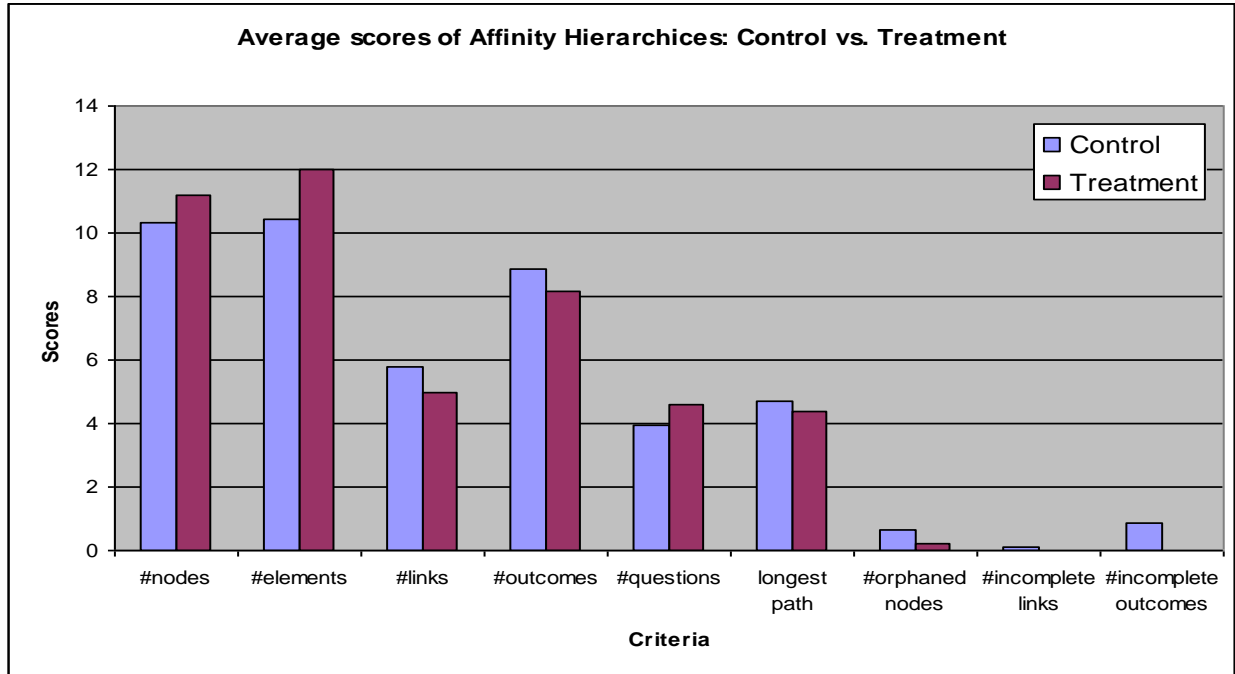
Figure 9. Average scores of the mechanics of Affinity Learning hierarchies produced: control vs. treatment groups.

System Evaluation. At the end of the laboratory exercises, students were asked to fill out a questionnaire (Appendix C). Figure 10 shows the average scores for the control and treatment groups. Accessibility, friendliness, logical, content, layout, and appeal are part of the feedback on the design of the interface. Ease of use, usefulness, effectiveness, speed, correctness, completeness, suitability, and technical correctness contribute to the overall functionality average. Finally, four criteria indicate the overall quality of the user manual.

- Based on a pair t-test, we found that there were two statistically significant differences between the two designs (alpha = 0.05): the layout (part of Interface) attribute and the effectiveness (part of Functionality) attribute. The graphical version was rated significantly higher than the non-graphical version (4.21 vs. 3.07 for layout; 4.14 vs. 3.50 for effectiveness).

- In general, the students thought that the graphical design had a better overall interface (3.95 vs. 3.48), better overall functionality (4.03 vs. 3.47), and better overall user manual (4.07 vs. 3.40). The difference in the user manual was primarily due to the complexity in describing the authoring process in the non-graphical Activity Builder.

- And surprisingly, even though the non-graphical design was not as convenient, students did not think too much worse of it (since a score of 3 would be the "neutral" score), compared to those students using the graphical design. One possible reason could be that the non-graphical design was a step forward in terms of (the lack of) computer software built for designing concept nodes to which the students had been exposed. According to the instructor, the students were also motivated in participating and learning from this exercise. These students chose to take this course and thus there was background eagerness. The tools

16

were also novel to them. In addition, the instructor integrated the study into his course well and was able to convince the students to want to do well in this exercise.
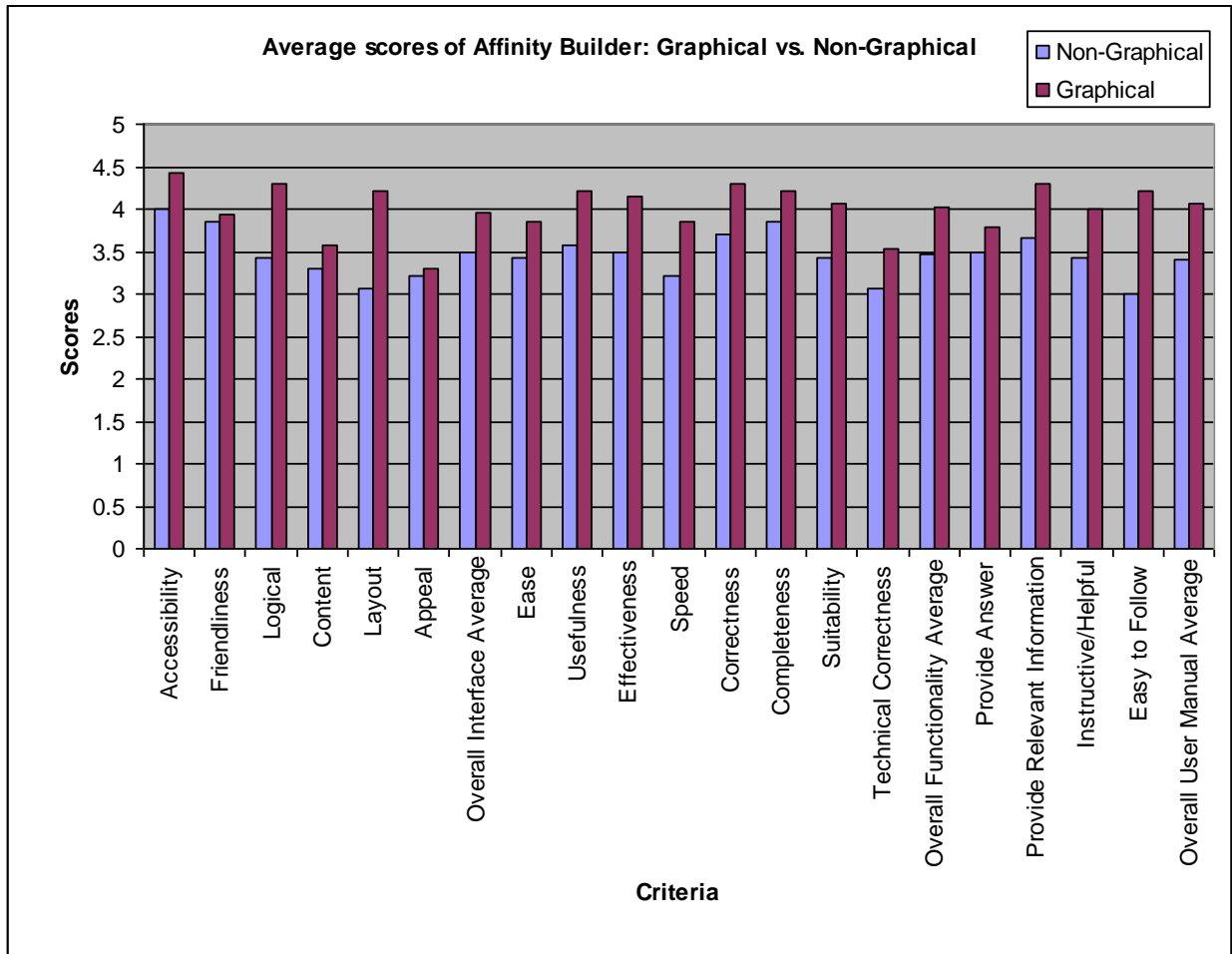
Figure 10. Average scores of user feedback on Affinity Builder: non-graphical vs. graphical versions.

5.2.   Hypotheses

Hypothesis 1: The graphical Affinity Learning authoring tool helps the users develop better Affinity Learning hierarchies than the non-graphical Affinity Learning authoring tool.

We found no statistically significant evidence to validate Hypothesis 1. However, there are indications that the graphical Affinity Learning authoring tool helps the users develop better-quality hierarchies (as shown in Figure 8) and hierarchies with more complete mechanics (as shown in Figure 9).

Hypothesis 2: The graphical Affinity Learning authoring tool improves the users' understanding of a subject matter more significantly than the non-graphical Affinity Learning authoring tool.

We found no statistically significant evidence to validate Hypothesis 2. Both authoring tools were found to be improve the users' understanding of a subject matter (i.e., how to teach algebra to secondary school students). However, the non-graphical Affinity Learning authoring tool was found to outperform the graphical Affinity Learning authoring tool, though the difference was statistically insignificant. This concludes that the convenience provided by the

graphical version, though facilitating better-quality hierarchies as well as hierarchies with better mechanics, does not lead to better student learning. We speculate that the students who used the non-graphical version relied on their cognitive organization of the misconceptions and remediation steps, and having to visualize the hierarchies on their own, were able to transfer what they learned from the laboratory exercise better to their post-tests than their counterparts in the treatment group.

This above finding has significant implications.

- First, a laboratory exercise during which students use a software authoring tool to build Affinity Learning hierarchies for a particular topic is able to help improve the students' understanding on how to develop better Affinity Learning hierarchies for another topic. Thus, continuing work in this direction—using the Affinity Learning authoring tool as a knowledge engineering and content development tool—is viable and feasible.

- Second, whether the tool is more convenient or less convenient, it does not make the students learn significantly different. On the contrary, there are indications that the students using a less convenient tool might learn more even though they might produce hierarchies of less quality. Thus, the graphical version of Activity Builder is the preferred version if the goal is to generate good quality hierarchies; and the non-graphical version is probably the preferred version if the goal is to use the authoring tool to train students to learn about misconceptions and remediation steps about a subject matter.

6.   Conclusions

We have described an evaluation of two authoring tools for Affinity Learning. We found that both tools help improve the users' understanding of the subject matter. However, though the graphical version can yield better-quality hierarchies, it does not help improve student learning more than the non-graphical version. In fact, the non-graphical version seemed to lead to higher student performance gains. This could be due to the convenience provided by the graphical version, which led to less cognitive processing of the Affinity Learning hierarchies in the students, and thus less effective transfer of the Affinity Learning concepts.

As for future work, we plan to conduct further studies to validate our findings and apply Affinity Learning to other disciplines such as humanities and engineering.

7.   Acknowledgements

8.   References

Ainsworth, S., J. Underwood, and S. Grimshaw (2000). Using an ITS Authoring Tool to Explore Educators' Use of Instructional Strategies, in Proceedings of the 5th International Conference on Intelligent Tutoring Systems (ITS'2000),

Ainsworth, S., N. Major, S. Grimshaw, M. Hayes, J. Underwood, B. Williams, and D. Wood (2003). REDEEM: Simple Intelligent Tutoring Systems from Usable Tools, Chapter 8, in Murray, T., S. Ainsworth, and S. Blessing (eds.): *Authoring Tools for Advanced Technology Learning Environment*. Dordrecht: Kluwer Academic Publishers, 205-232.

Aroyo, L., A. Inaba, L. Soldatova, and R. Mizoguchi (2004). EASE: Evolutional Authoring Support Environment, in *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (*ITS'2004*).

Brown, J. and Burton, R. (1978). Diagnostic Models for Procedural Bugs in Basic Mathematics, *Cognitive Science*, **2**(2):155-192.

Bruning, R., Zygielbaum, A., and Grandgenett, N. (2001). Using Online Learning Resources to Promote Deeper Learning, in *Proceedings of the Symposium on Assessing the Quality of Online Instruction*, Monterey, CA.

Brusilovsky, P. (2003) Developing Adaptive Educational Hypermedia Systems: From Design Models to Authoring Tools. In: T. Murray, S. Blessing and S. Ainsworth (eds.): *Authoring Tools for Advanced Technology Learning Environment*. Dordrecht: Kluwer Academic Publishers, 377-409.

Clement, J. (1989). The Concept of Variation and Misconceptions in Cartesian Graphing, *Focus on Learning Problems in Mathematics*, **11**(1-2):77-87.

Dabbagh, N. (2002). The Evolution of Authoring Tools and Hypermedia Learning Systems: Current and Future Implications, *Educational Technology*, **42**(4):24-31.

Dede, Y. (2004). The Concept of Variable and Identification its Learning Difficulties, *Educational Sciences: Theory & Practice*, **4**(1):50-56.

Fisher, K. M. (1988). The Students-and-Professors Problem Revisited, *Journal for Research in Mathematics Education*, **19**(3):260-262.

George, L. and R. E. Nunez (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*, Basic Books.

Jurkovic, N. (2001). Diagnosing and Correcting Student's Misconceptions in an Educational Computer Algebra System, in *Proceedings of the International Symposium on Symbolic and Algebraic Computation (ISSAC'2001)*, July 22-25, Ontario, Canada, pp. 195-200.

Kraus, P., and Minstrell, J. (2003). Designing Diagnostic Assessments. Paper presented at the

2002 Physics Education Research Conference, Boise, ID.

Major, N., Ainsworth, S. and Wood, D. (1997). REDEEM: Exploiting Symbiosis between Psychology and Authoring Environments, **8**:317-340.

Minstrell, J. (2001). The Role of the Teacher in Making Sense of Classroom Experiences and

Effecting Better Learning. In D. Klahr, Carver, S. (Ed.) *Cognition and Instruction: 25 Years of*

Progress. Mahwah: Lawrence Erlbaum Associates.

Moundridou, M. and M. Virvou (2002). Analysis and design of a Web-based authoring tool generating intelligent tutoring systems, *Computers and Education*, **40**:157-181.

Murray, T. (1998). Authoring Knowledge-Based Tutors: Tools for Content, Instructional Strategy, Student Model, and Interface Design, *Journal of the Learning Sciences*, **7**(1):5-64.

Murray, T. (1999). Authoring Intelligent Tutoring Systems: An Analysis of the State of the Art, *International Journal of Artificial Intelligence in Education*, **10**:98-129.

Murray, T. (2003). An Overview of Intelligent Tutoring System Authoring Tools: Updated analysis of the state of the art, Chapter 17 in Murray, T., Blessing, S. and Ainsworth, S. (Eds.) Authoring Tools for Advanced Technology Learning Environments, Dordrecht: Kluwer Academic Publishers.

Novak, J. D. (1998). *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*, Mahwah, NJ: Lawrence Erlbaum Associates.

Rosnick, P. (1981). Some Misconceptions Concerning the Concept of Variable: Are you Careful about Defining Your Variables?, *Mathematics Teacher*, **74**(6):418-420.

Selitto, G . L. (1997). *An Identification of Problematic Algebraic Concepts and the Understanding Possessed by Students in Elementary Algebra*, Ph.D. Dissertation, Columbia University.

Small, D. (2007). An Urgent Call to Improve Traditional College Algebra Programs, U.S. Military Academy, West Point, NY 10996. MAA OnLine. www.maa.org.

Van Rosmalen, P., Vogten, H., Van Es, R., Passier, H., Poelmans, P., & Koper, R. (2006). Authoring a Full Life Cycle Model in Standards-Based, Adaptive E-Learning, *Educational Technology & Society*, **9**(1):72-83.

VanLehn, K. (1990). *Mind Bugs: The Origins of Procedural Misconceptions*, The MIT Press.

Virvou, M. and M. Moundridou (2000). A Web-Based Authoring Tool for Algebra-Related Intelligent Tutoring Systems, *Educational Technology and Society*, **3**(2):

Zygielbaum, A.I., and Grandgenett, N. (2001). Affinity Learning in Mathematics, Invited Paper, *NSF Workshop on Pre-Calculus Education Reform*.

Zygielbaum, A.I., and Feese, R. (2005), Affinity Learning – Fostering Critical Thinking Skills

Online, in PytlikZillig, M. Bodvarsson, and R. Bruning (Eds.) *Web-based Education: Bringing Practitioners and Researchers Together*, Information Age Publishing.