The Eurasia Proceedings of
Educational & Social Sciences (EPESS)
ISSN: 2587-1730

The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2016

Volume 5, Pages 434-454

ICRES 2016: International Conference on Research in Education and Science

EVALUATION OF EDUCATIONAL SOFTWARE AND PAPER-BASED RESOURCES FOR TEACHING LOGICAL-THINKING SKILLS TO GRADE SIX AND SEVEN STUDENTS

Peter Fenrich
British Columbia Institute of Technology

Abstract: To determine whether there were any differences between the effectiveness of educational software and paper-based materials in teaching the logical-thinking skills of classification, analogical reasoning, sequencing, patterning, and deductive reasoning, a quantitative assessment was conducted using a pre-test, post-test, experimental design. One-way ANOVAs were used to compare an experimental group learning from educational software (32 students), an experimental group learning from paper-based materials (32 students), and a control group (32 students). Tukey HSD Post Hoc Tests were performed since a significant difference was found between the groups. For each test, the subjects taught through educational software and those taught through paper-based materials scored significantly higher in logical-thinking ability than the control group, except for the subskill of deductive reasoning for both experimental groups. There were no significant differences between subjects taught through educational software and those taught through paper-based materials on any test. Results from paired samples t-test results showed that the subjects learning from educational software and those learning from paper-based materials had significant percentage gains on all of their pre-test to post-test scores, except the subjects learning through paper-based materials showed no significant gains on the subskill of deductive-reasoning.

Keywords: Logical thinking, instructional design, educational software

Introduction

Based on the literature, researchers, administrators, teachers, employers, and parents from around the world seem to agree that students do not leave the school system with the higher-order thinking skills needed in both the workplace and life (Burkhart, 2006; Lee, 2008).

There are difficulties in teaching higher-order thinking skills. One reason is that many traditional teaching methods do not inherently address those skills (Lee, 2010). As well, some teachers do not have the skills needed to teach specific higher-order thinking skills (Jeremiah, 2012), while other teachers only have minimal skills in in that domain (Almatrodi, 2007; Burkhart, 2006). Another concern is that some school districts require students to complete standardized tests that focus on lower-order thinking skills (Sondel, 2009). Consequently, standardized tests can lead teachers towards teaching lower-order thinking skills (Burkhart, 2006). Given these factors, it is not surprising that teachers and the education system itself have been criticised and are under pressure to improve the teaching and learning of higher-order thinking skills.

Historically, there has been little research on teaching higher-order thinking skills (Sondel, 2009). The literature provided few details about the instructional strategies used in interventions (Wruck, 2010). This has led to uncertainty in how to teach higher-order thinking skills and the more specific logical-thinking skills.

Objectives and Research Questions of the Study

This study was based on two objectives. The first was to assess whether there were any differences between the ability of comparable paper-based materials and educational software in teaching logical-thinking skills. The

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Peter Fenrich E-Mail: Peter_Fenrich@bcit.ca

second was to determine whether the paper-based materials and educational software could effectively teach logical-thinking skills.

To achieve the objectives, the following research questions guided the study:
1. Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those taught using paper-based materials?
2. Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those not being exposed to any intervention?
3. Were there significant differences in the logical-thinking ability between subjects taught using paper-based materials compared to those not being exposed to any intervention?

Literature Review

Introduction

Since researchers have provided little information on how to design and develop materials that teach higher-order thinking skills, there is no clear method for effectively teaching higher-order thinking skills or the more specific logical-thinking skills. As well, there is even less research if the factors of being taught in a stand-alone mode and the subjects being grade six and seven students are considered. Related research has tended to focus on traditional classroom practices (Semper Scott, 2005), test scores, (Mintz, 2000), the cost-effectiveness of the intervention (Mintz, 2000), and/or post-secondary school subjects (Burkhart, 2006).

Based on the research questions, a number of topics are discussed within this literature review.

Logical-thinking Skills

Logic is a “way of reasoning correctly, or without making mistakes, to solve problems” (Rivière, 1990, p. 13). Mains (1997), in a similar way, defined logical thinking as “reasoning in a clear and consistent manner based on earlier or otherwise known statements, events, or conditions (Mayer, 1983)” (p. 8). Lamb (2001) stated that logical thinking and reasoning includes numerous skills such as comparison, classification, sequencing, cause/effect, patterning, webbing, analogies, deductive and inductive reasoning, forecasting, planning, hypothesizing, and critiquing. For practical reasons, the construct of logical thinking for this study was more narrowly defined. The specific logical-thinking skills that were taught and measured were the skills of classification, analogical reasoning, sequencing, patterning, and deductive reasoning. Since these specific skills can be defined in different ways, each skill was specifically defined for this research study. The classification skill entailed determining which word, from a list of words, had the same thing in common as three given words. Analogical reasoning was the ability to discover a specific similarity between a given pair of words and using that similarity to match another given word to a word in a list. The sequencing skill required subjects to determine a repeating pattern within a sequence of numbers to predict the next two numbers in that sequence. Patterning required learners to determine which examples in a third series matched the common pattern of the first series. Deductive reasoning required subjects to draw conclusions based on given information, then draw new conclusions based on the current information, and repeat that process until the problem is solved.

Assessment of Thinking Skills

The instruments commonly used to measure higher-order thinking skills would not inherently measure all of the specific logical-thinking skills addressed in this research and not in the same context for the same age group. Consequently, no commercial instrument fit this study. This matches the opinion of Lee (2008), who stated that current testing instruments tend to be too general to be useful.

Educational Software

The research on teaching higher-order thinking skills through an educational-software intervention has shown positive findings, mixed findings, and no significant differences.

Significant Findings Through Educational Software

The studies discussed below show that learning through educational software can increase a learner's higher-order thinking skills, although some of the studies had mixed findings.

With five-year and six-year old students, Bradberry-Guest (2011), assessed whether the supplemental Webber Interactive WH Questions software program could significantly increase a subject's ability to answer "why" questions. The experimental group scored significantly higher than the control group.

Katzlberger (2006) compared two educational-software approaches to teach problem-solving skills to grade-six subjects. The educational-software approaches were identical except one approach also included the task of the subjects teaching a computer-based agent the skills. Both approaches led to a significant increase in problem-solving skills. There were no significant differences between the two approaches.

The Higher-Order Thinking Skills (HOTS) software program for elementary-school children does not directly teach higher-order thinking skills but rather relies on initiating student to student and student to teacher interactions for students to solve problems and interpret events (Pogrow, 2005). Pogrow stated that the HOTS program resulted in significant gains in mathematics, reading comprehension, and metacognition. In a different study of the HOTS program utilized by students in grade four and five, as described by (1999) in a literature review, grade-five students scored significantly higher in sequential-synthesis skills. However, no significant differences were found in abstract-relation skills.

Campbell (2000) evaluated the Computer Curriculum Corporation's Successmaker educational software with grade four and five students. The software contained challenging multimedia-based problems and specific questions based on critical thinking. Campbell found that an experimental group in one school showed an increase in critical-thinking ability while an experimental group in another school had a decrease in critical-thinking ability as compared to the control group.

Lewis (1998) conducted a study with grade-four students to determine whether learning from the educational-software program entitled The Yukon Trail would lead to an increase in critical-thinking skills. Lewis found significant positive gains in critical-thinking skills.

With grades one through six subjects, Shiah (1994) assessed whether subjects taught through educational software would have an increase in their ability to solve math word problems. Within the educational software, one experimental group was given an explicit cognitive strategy to follow and had animated images to support the strategy while the other experimental group received the same cognitive strategy but with static images to support the strategy. The control group received a different computer-delivered problem-solving approach and supporting static images. Shiah found significant gains in problem-solving skills in all three groups with no significant differences between the groups.

Allison (1993) compared an experimental group that was taught critical thinking via a problem-solving focussed software program, called the Super Math and Reading Thinking Skills, to a control group receiving traditional education. The subjects were in grade three, four, and five. The experimental subjects performed significantly better than the control group in math skills.

Leiker's (1993) research, with grade three and four students, compared mathematical and higher-order thinking skills in students who were taught with supplemental educational software to those taught only through traditional means. Leiker found that the experimental subjects achieved higher scores in both mathematics and higher-order thinking skills.

With grade one and two students, Orabuchi (1992) compared an experimental group that had supplemental use of numerous educational-software packages to a traditionally-taught control group. The software packages included Animal Rescue, Trading Post, Muppet Math, Mickey's Magic Reader, What's in a Frame?, Memory Building Blocks, Odd One Out, Hands On Math, Patterns, Reading Magic Library: Flodd The Bad Guy, and Choices Choices: Taking Responsibility. Orabuchi found that the experimental group scored significantly higher on inferences, generalisations, and math problem-solving.

Cotton (1991) reviewed five research studies that assessed the effectiveness of educational software and found that all of the educational-software programs were able to significantly increase a student's thinking skills. However, the overall findings were mixed in four of the five studies. The educational-software programs focussed on analogical reasoning, logical reasoning, and inductive and deductive thinking.

Swan (1990) assessed whether grade four to eight students would perform differently on logical-thinking skills when provided with different types of problems to solve using the Logo programming language. One group received graphic problems, another group was given lists problems, while a third group received both graphics and lists problems. Swan found that there were no significant differences between the groups in the skills of sub-goal formation, analogical reasoning, forward chaining, backward chaining, systematic trial and error, and alternative representation. However, there were significant gains in the five skills of sub-goal formation, analogical reasoning, forward chaining, systematic trial and error, and alternative representation. Due to the lack of a control group, these findings must be carefully interpreted.

Duffield (1989) assessed whether grade three and four students could learn problem-solving skills from the educational-software packages called the King's Rule, which taught specific skills used in math and social studies, and Safari Search, which taught problem solving through puzzles. Duffield found that the subjects gained problem-solving skills that the software specifically addressed and devised their own unique strategies to solve the presented problems. Duffield did not have a control group.

With grade seven and eight students, Galinski (1988) evaluated the effectiveness of a software program called The Factory. Both the experimental and control groups received traditional mathematics instruction. Galinski found that both groups gained in mathematical problem-solving, analysis, and synthesis ability. Galinski also found no differences between the two groups with respect to mathematical problem-solving, analysis, and synthesis ability but did find that the control group had an increase in spatial ability. The gains could have been due to the traditional instruction rather than the software.

Judy (1987) assessed whether a direct instructional strategy and an inquiry-based instructional strategy would lead to an increase in analogical reasoning skills in grade-six subjects. Judy found that both instructional strategies led to a significant increase in analogical reasoning skills. She also found that the direct instructional strategy led to significantly better results than the inquiry-based instructional strategy.

Non-significant Findings Through Educational Software

The following studies illustrate that learning through an educational-software intervention can result in no significant differences with respect to a learner's higher-order thinking skills.

Baumer (2009) compared experimental subjects taught in a traditional way but also received supplementary computer-delivered instruction to control-group subjects taught in a traditional way but who also received additional traditional materials. The software analyzed, with respect to metaphor usage, the subject's answers to typical questions and then provided questions leading students to create their own metaphors. Baumer found no significant differences with respect to metaphorical reasoning or creativity.

Mintz (2000) compared an experimental group of grade four and five students learning mathematics from the Successmaker educational software to a control group learning mathematics through traditional teaching methods. Mintz found no significant differences between the groups.

Rothman (2000) evaluated whether an educational-software game, entitled The Voyage of the Mimi, could lead to increases in critical thinking compared to traditional methods of teaching grade-five students. Although the experimental group had a higher positive trend in critical thinking, the results were insignificant.

Ellingwood (1999) evaluated the effects of Logo programming language instruction on higher-order thinking skills with grade-one students. Experimental students were initially taught keyboarding and other basic computer skills. These students were then taught mathematics through Logo programming. The control group was taught mathematics by traditional methods. Ellingwood found that the experimental group had higher mean gains in higher-order thinking skills but these gains were insignificant.

With grade-four subjects, Lafferty (1996) determined whether experimental-group subjects who worked through the Structure of Intellect educational-software program in addition to their normal coursework would perform better than the control group that only received their normal coursework. The software is a training program that aims to develop over twenty cognitive skills. Lafferty found no differences in performance scores.

Schmidt (1991) assessed whether grade-seven science students who used the Weather Prediction expert system would gain in weather-prediction skills. The subjects also learned about weather through textbook readings, discussions, taking actual weather measurements, and entering measurements into the expert system to see the system's weather prediction. They were asked, based on data, to make their own predictions and compare those to the computer's predictions. Schmidt found no significant differences in pre-test to post-test scores.

Repman (1989) assessed whether different commercially-available software packages (Bank Street Writer, U.S. History Databases, Where in the U.S.A. is Carmen San Diego?, Ten Clues, Crossword Magic, and SuperPrint!) could effectively teach critical-thinking skills to grade-seven students. Repman found no significant differences between the experimental and control groups.

Bass and Perkins (1984) used educational software to attempt to teach grade-seven students verbal analogies, logical reasoning, inductive/deductive reasoning, and word-problem analysis. As compared to traditional teaching methods, Bass and Perkins did not find any statistically significant differences.

With kindergarten students, vonStein (1982) compared an educational-software tutorial to traditional teaching for how to sequence shapes. No significant differences were found between groups.

Paper-based Materials

Only two studies were found that used a stand-alone paper-based intervention that aimed to teach and assess gains in higher-order thinking skills.

Using the Primary Education Thinking Skills™ (PETS) program as an intervention with elementary-school subjects, Thomson (2009) assessed whether experimental subjects exposed to the supplemental use of the program would gain in reasoning skills as compared to control subjects who did not use the PETS program. The pre-test was given as the subjects entered grade one and the post-test was given as the subjects completed grade three. Thomson found no significant differences in reasoning skills.

With adults taking an introductory educational-technology course, Lee (2008) wrote online-discussion transcripts for each group to read. The control group read expository text whereas one experimental group read high-level discussion transcripts while the other experimental group read low-level discussion transcripts. The participants of each group wrote their responses to the transcripts in an essay. Lee found that the experimental group that read high-level discussion transcripts scored significantly higher in critical-thinking skills than the other two groups.

Findings for Comparable Interventions

This research compared gains in logical-thinking skills through an educational-software intervention to a highly similar paper-based intervention.

In a landmark paper, Clark (1983) stated,
Consistent evidence is found for the generalization that there are no learning benefits to be gained from employing any specific medium to deliver instruction. Research showing performance or time-saving gains from one or another medium are shown to be vulnerable to compelling rival hypotheses concerning the uncontrolled effects of instructional method and novelty (p. 445).

This matches Semper Scott (2005) who stated that minor differences in the learning experience do not result in significant performance differences. Larson, Britt, and Kurby (2009), Shiah (1994), and Titterington (2007) each had findings that supported Clark's (1983) conclusion given they found no significant differences with comparable interventions. Heo (2012) and Cott (1991) also supported Clark's conclusion. Heo (2012) assessed whether there were any comprehension differences between college students taught with video-based instruction, audio-based instruction, or text-based instruction. Heo found no differences in comprehension between the three groups. Cott (1991) designed a study with college students, where the experimental group was taught logic skills via a programmed instruction booklet and the control group was taught the same skills via traditional text. The programmed instruction and text versions were designed to closely match each other. Cott found no significant differences in achievement between the two groups.

Educational Software Compared to Paper-based Interventions

There were relatively few research studies on teaching higher-order thinking skills that compared an educational-software intervention to a paper-based intervention.

Larson et al. (2009) assessed whether claim-reason argumentation could be taught to both secondary and post-secondary students through stand-alone paper-based and educational-software tutorials. As compared to the control group that did not receive the tutorial, the subjects in both experimental groups had a significant increase in the ability to detect flawed arguments. The increases between the two groups were comparable. Similarly, Larson et al. found similar gains with a web-based tutorial that closely paralleled the software version.

With post-secondary allied-health students, Titterington (2007) quantitatively found significant but comparable gains in higher-order thinking skills through both a traditionally-delivered paper-based intervention and a comparable online-delivery intervention.

With first-year post-secondary students, Bessick (2008) compared the critical-thinking skills from learning from an educational-software package called the Rationale Argument Mapping Program to learning from a paper-based package called The Thinker's Guides. The Rationale Argument Mapping Program was based on an interactive approach that guided, built, and evaluated an individual's ability to make arguments. The Thinker's Guides were based on Richard Paul's model of critical thinking and were designed to help students "identify general concepts related to critical thinking and specific critical-thinking skills necessary to think effectively for different disciplines and tasks" (Bessick, 2008, p. 57). The two interventions are not directly comparable in that the instructional strategies of these two resources are distinctly different. Bessick found that neither experimental group showed a significant increase in critical-thinking skills.

Of the above studies, none closely matched this research study in terms of aiming to teach the same logical-thinking skills via both an educational-software and a paper-based intervention, following a related instructional strategy, and having subjects in grade six and seven. None of the researchers provided details about the instructional strategies that they used in their interventions.

Principles of Instructional Design

This section first illustrates how an instructional development cycle model, a systematic process of instructional design, and Gagné's Nine Events of Instruction (Gagné, Wagner & Briggs, 1988) were amalgamated to create the Combined Instructional Design and Development Model. This section then describes the learning theories that need to be applied to help ensure that instructional strategies will lead to effective learning.

Combined Instructional Design and Development Model

Instructional development cycle models guide instructional development processes for creating materials that enhance learning. A nonlinear instructional development cycle model makes it possible to evaluate and revise the materials as they are being designed and developed (Fenrich, 2014). The instructional development cycle followed in designing the interventions of this study began with the analysis phase and continued with the planning, design, development, and implementation phases. The evaluation and revision phase was ongoing throughout the cycle. After each phase, the outputs were evaluated and revisions were made.

The nonlinear instructional development cycle model needs to contain a systematic instructional design process (Fenrich, 2014). Steps of the systematic instructional design process, which were based on Dick and Carey's (1990) model, included identifying the instructional goal, conducting a goal analysis, conducting a subordinate skills analysis, identifying entry skills and characteristics, writing learning outcomes, developing criterion-referenced test questions, developing an instructional strategy, developing and selecting instructional materials, and conducting formative evaluations, with getting feedback and revising at each step.

The systematic instructional design process does not in itself provide all of the attributes that effective instructional materials need. To help ensure effective learning, educational materials should be created based on principles of instructional design and follow a model such as Gagné's Nine Events of Instruction. Gagné's Nine Events of Instruction are gaining attention, informing the learner of the learning outcome, stimulating recall of prerequisites, presenting the material, providing learning guidance, eliciting the performance, providing feedback, assessing performance, and enhancing retention and transfer (Gagné et al., 1988; Singh, 2010).

The non-linear instructional development cycle model was merged with the systematic instructional design process and Gagné's Nine Events of Instruction to create a Combined Instructional Design and Development Model, as depicted in Figure 1.

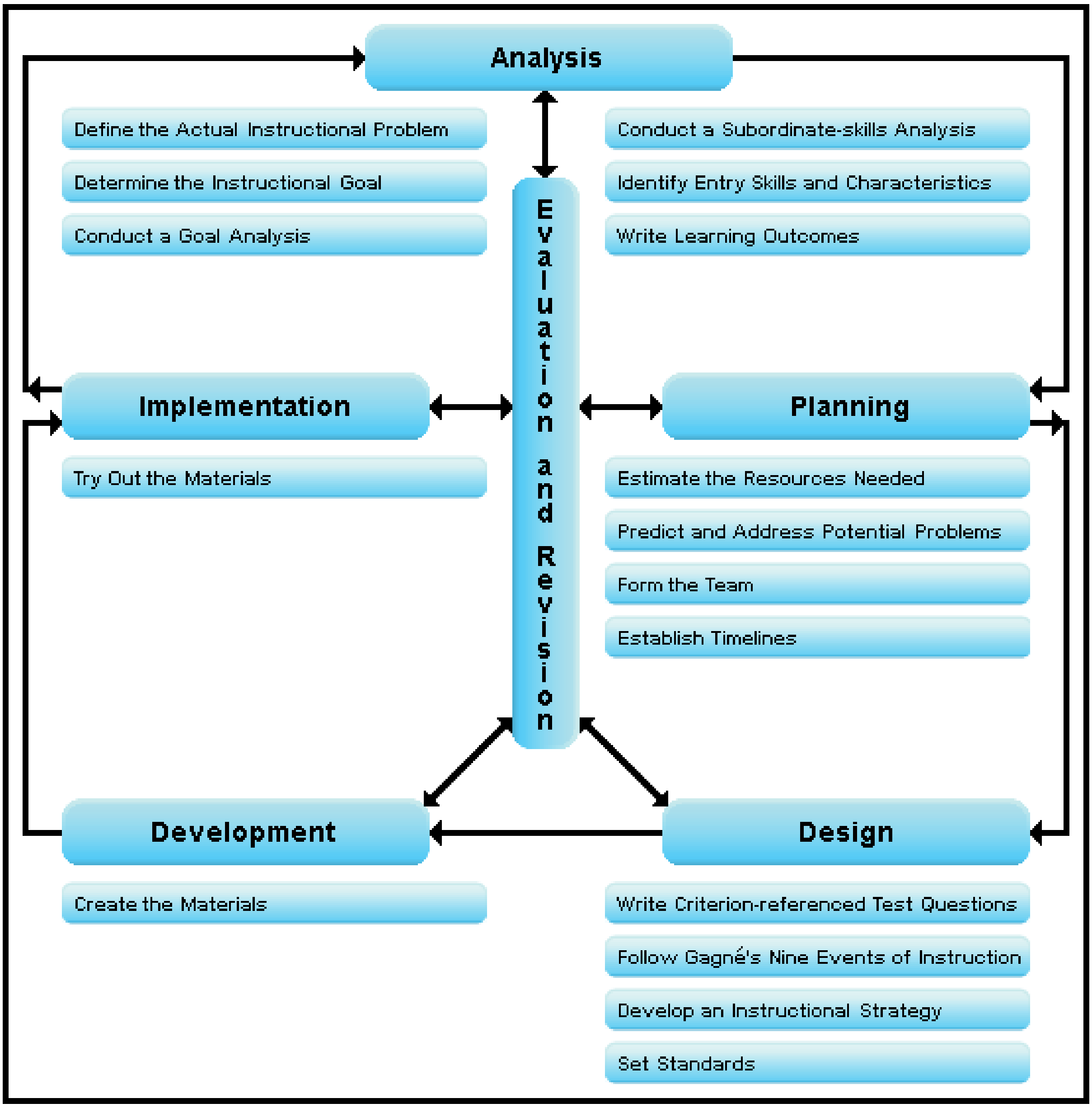

Figure 1. Combined instructional design and development model

Creating Effective Instructional Materials

The Combined Instructional Design and Development Model provided the framework for designing effective instructional materials.

The first step in the analysis phase is to identify the instructional goal(s). Instructional goals are general learning outcomes that can be broken down into precise measureable skills. Instructional goals should be based on a definition of the actual instructional problem. The next activity is to conduct an instructional analysis, which includes a goal analysis and subordinate skills analysis. In the goal analysis, create a general but precise visual statement of the consecutive steps that a learner will do to achieve the goal. Complete a subordinate skills analysis for any step of the goal analysis that is too large to be taught in one step or to determine if the learners need more information prior to learning a step. This may result in some or all of the steps being broken into smaller elements. Identify entry behaviours and learner characteristics next. For effective learning, there must be a match between the instructional materials and the capabilities of the widest practical range of learners that consider the learners' abilities, language level, motivation, and interests. This leads to a list of skills that are not taught because, in general, the target audience has already mastered them. The next step is writing learning outcomes, which are specific measurable skills that describe what learners should be able to do in a more specific way than instructional goals. Learning outcomes form the basis of all of the subsequent steps. The analysis phase ends with conducting formative evaluations and making revisions (Fenrich, 2014; Reddy, 2008).

The planning phase helps the instructional development process proceed smoothly. In the planning phase, it is important to provide initial estimates of the resources needed, even though the estimates may be inaccurate until there is more information. The estimates can help to determine if there may be resource constraints. Another important part of the planning phase entails identifying and addressing potential problems. Other tasks include assembling the team and setting timelines. The last step in the planning phase is to conduct formative evaluations and make revisions (Fenrich, 2014; Reddy, 2008).

For each learning outcome, an initial step in the design phase is to develop criterion-referenced test questions. A part of this step is to determine whether the test actually measures what it is intended to measure. Being able to test a learning outcome helps to confirm that the outcome is written correctly. A challenging part of the design phase is to develop instructional strategies. Instructional strategies are needed for each learning outcome and are created by following principles of instructional design. This task is the heart of the instructional solution as flaws in instructional design compromise learning. Within the instructional strategies, determine what media are needed to support learning. As is appropriate, set standards on the writing style and tone, menus, orientation information, headings, image locations, text locations, prompt locations, prompt wording, error messages, navigation text and/or buttons, branching details, fonts, font sizes, highlighting methods, colours, scoring for tests, and criteria for passing. For an educational-software intervention, programming the above standards into templates and sub-routines can begin. Another step is to confirm or refine information regarding the needed resources, including personnel, equipment, and software, based on what is now known about the design. The last step in the design phase is to conduct formative evaluations and make revisions (Fenrich, 2014; Reddy, 2008).

The development phase entails creating and selecting the instructional materials and evaluating the materials. The first step is to develop and select the instructional materials, including the media, based on the instructional strategies and the design specifications. If the instructional materials will be paper-based, the core part of the work is already done. If the instructional materials are delivered in a different way, such as with educational software, the paper-based content is used as the foundation for developing those materials. A critical step of the development phase is to conduct formative evaluations and make revisions. This should first be done with a prototype, which may be based on the materials for one or more learning outcomes. Changes made in the prototype can be applied, as is appropriate, to all future content (Fenrich, 2014; Reddy, 2008).

Implementation entails the step of trying the materials in a real setting to determine what works and what needs to change. Ensure that the personnel, facilities, equipment (as needed), and instructional materials are available when scheduled. For educational software, the software should be run on the equipment that will be used for the full implementation. A team member should work through every screen and question answer choice. If needed, problems should be addressed. Once everything is set up, based on a representative sample of target audience learners, conduct an evaluation and make revisions. After the instructional materials are thoroughly evaluated and revised, full implementation can begin (Fenrich, 2014; Reddy, 2008).

Evaluation is the systematic collection and analysis of information to support decision making and planning. Formative evaluation should be ongoing throughout each phase of the instructional design and development process. Typically, educational software should be tested on different computer systems with varying speeds, memory, screen resolutions, and monitors. This should be done early in the process with initial prototypes. Revisions should be based on the feedback and data obtained in each iteration of evaluation. Summative evaluation should occur after the instructional solution is fully implemented (Fenrich, 2014; Reddy, 2008).

Designing Learning for Higher-order Thinking

The effectiveness of instructional materials is limited by the instructional design. To ensure instructional effectiveness, educational software, paper-based content, and other instructional materials need to be designed based on principles of instructional design that are grounded on learning theory. The instructional design process should ensure many things including that the content is targeted to the intended audience, learning outcomes are stated, measurable, and at the highest appropriate level, learners are motivated to learn, the instructional strategies are designed to solve the specific instructional problem, and the materials are highly interactive (Fenrich, 2014; Gagné et al., 1988; Wu, 2009). Given that higher-order thinking skills are complex and challenging to teach, focussing on the instructional design of the materials is paramount. If this is not done, learning gains may be insignificant.

General Methods for Designing Learning for Higher-order Thinking Skills

Higher-order thinking skills need to be specifically taught in that learners do not tend to acquire these skills passively (Burkhart, 2006; Lee, 2008; Wruck, 2010). Cotton (1991), in her review of research studies where thinking skills were explicitly taught, found that each intervention resulted in improved performance. However, explicitly teaching the skills is not a simple matter. According to Sondel (2009), “There is a dearth of information which evaluates the best ways to improve critical thinking” (p. 15). Beyond the limited information on how to teach higher-order thinking skills (Thomson, 2009), there is a challenge to teach the skills because they tend to be abstract and there is no agreement on how to define them (Jeremiah, 2012). In other words, how do you teach and assess something if you are not sure what it is? A further problem is that there does not seem to be one “right” method for teaching higher-order thinking skills (Bessick, 2008).

The instructional strategies should be based on the specific learning outcomes that need to be taught as well as the thinking processes that the learner needs to invoke to achieve the learning outcomes (Rukavina, 2003). One common theme from the literature is that the instructional strategies needed for teaching higher-order thinking skills must focus on higher-order thinking skills. One reason for this is that students often only superficially interact with typical instructional materials, which suggests a lower thinking level is being used (Enniss 2006).

Teaching higher-order thinking skills requires planning, such as gradually building the skills in incremental steps to help ensure success (Bessick, 2008) and intense practice (Burkhart, 2006) as these skills require significant effort to learn. In contrast, the stereotypic didactic instructional approach and traditional drill-and-practice have been relatively ineffective in teaching higher-order thinking skills (Almatrodi, 2007).

Instructional Strategies for Teaching Analogies

For subjects in grade six and seven and the specific logical-thinking skills addressed within the interventions of this study, the literature only provided basic ideas for teaching analogical-reasoning skills. Masteron and Perrey’s (1999) instructional strategy entailed learners determining a relationship between the first pair of words, pairing the third word with a word in the list that has a relationship like the relationship between the first pair of words, and creating a sentence that incorporated those relationships. Masteron and Perrey presented sample analogies, had the subjects practice analogies in groups, and then had the subjects practice individually.

Relevant Learning Theories

Following Piaget’s theory of cognitive development, the highest thinking level within the content needs to be at an achievable level. Consequently, the skills learned must gradually build in small incremental steps from the student’s current ability to the highest skill level desired for each learning outcome (Leiker, 1993; Reddy, 2008).

Within the content, important concepts must be emphasized so that the learner focuses on the material. Techniques for motivating learners can help with this. When a learner concentrates, the learner can encode new concepts stored in short-term memory with existing concepts from long-term memory and then store that linked information in long-term memory for later retrieval. When the new concepts match an existing mental model, the concepts are assimilated into long-term memory. When the new concepts extend an existing mental model, the concepts are accommodated into long-term memory. While factoring in the cognitive load theory, which suggests that only a manageable amount of content should be presented at a time so as to not overload the short-term memory, instructional materials need to facilitate both the assimilation and accommodation of new concepts into long-term memory. Assimilation and accommodation allow learners to construct knowledge, which is a tenet of the constructivism theory (Fenrich, 2014; Heo, 2012; Katzlberger, 2006).

Also, based on the constructivism theory, the instructional strategies should provide practice opportunities with immediate and elaborative feedback to help learners construct accurate mental models of the concepts in long-term memory, be highly interactive, and allow learners to proceed at their own pace to enable the learners to spend the time needed to master the materials. Allowing learners to proceed at their own pace supports the important concept of metacognition, where learners monitor their progress and decide what is needed to support their learning goal (Bessick 2008; Burkhart, 2006; Cott, 1991; Fenrich, 2014; Katzlberger, 2006).

Learning theory also suggests that the materials have variation so that each learner has some preferred activities, support the corresponding learning outcome, direct the students to learn the content deeply, guide the students in how to solve each of the skills, include questions that are at the highest appropriate thinking level, have elaborative feedback that is accurate and complete, and include summaries (Fenrich, 2014; Katzlberger, 2006).

Gagné's Nine Events of Instruction

With respect to this study, the researcher designed the materials based on Gagné's Nine Events of Instruction, as it is a solid foundation from which to design the instructional strategies (Gagné et al., 1988; 2008; Singh, 2010).

Gaining Attention

The learners' attention should be gained to get the learners involved and motivated. As well, a learner's attention should be maintained throughout the learning materials (Fenrich, 2014; Gagné et al., 1988; Singh, 2010). In this study, the techniques used to gain and maintain attention were based on the reviewers' recommendations to "ask the students to obtain high scores", "stress the importance of thinking carefully", "make the materials highly interactive", and "pose challenging statements and questions".

Informing the Learner of the Learning Outcome

Informing learners of the learning outcome enables them to focus their efforts (Fenrich, 2014; Singh, 2010). The learning outcomes were written to be clear, measureable, and at the highest appropriate level of difficulty. The learner was informed of each learning outcome.

Stimulating Recall of Prerequisites

In this study, recalling prerequisites was not needed as a part of the instructional strategy. The skills were taught from a foundational level that was suitable for grade six and seven students.

Presenting the Material

There are many principles for effectively presenting learning materials. The total amount of material presented in a lesson should be based on the learners' age, the learner's expected attention span, the material's complexity, the activities needed to teach the skills effectively, and the time needed for all of the instructional events. In general, material should be presented in increasing difficulty and in small incremental steps. This helps ensure learner success and increases the learner's confidence (Fenrich, 2014; Gagné, et al., 1988). Providing a variety of instructional strategies and activities can generate interest. The activities provided must support the learning outcomes being taught (Fenrich, 2014; Semper Scott, 2005). Learners should actively participate in his or her learning (Baumer, 2009; Fenrich, 2014; Gagné et al., 1988). Learning by doing becomes particularly important for instructional materials that require extensive reading since learners, due to the limitations of memory, cannot remember all that they read (Wu, 2009). Consequently, the instructional strategy must direct the learner's focus to the deeper learning concepts that support higher-order thinking skills (Wu, 2009). The learner can be directed to focus on deeper learning by making the material interactive through activities that require higher-order thinking (Enniss, 2006; Fenrich, 2014). With respect to presenting the material, there were a number of instructional features in each of the lessons created for this study. Some of these features were based on the reviewers' recommendations for teaching logical-thinking skills, which included that it was important to "ask numerous questions", "stimulate a high level of thinking", and "gradually increase the difficulty of the content". To support academically weak through strong target-audience students, the materials were designed for the expected level of cognitive development and attention spans of students in grade six and seven. The concepts gradually increased in difficulty and each sample of each skill was explained in a series of manageable steps. A variety of instructional activities and strategies were provided. Each logical-thinking skill was taught through presenting samples, practice questions, a summary, a self-test, and challenge questions. In particular, the samples, practice questions, and summary were inherently different and thus there were varied activities. As well, each of the five logical-thinking skills required a different instructional strategy. There was a high degree of active learning in the samples, practice questions, self-test, and challenge questions. Learners were highly engaged in each of these activities, especially since answering the questions required high-order thinking. The samples provided for each of the logical-thinking skills required the learners to think about each step, rather than the learners simply being told how to solve the problems. In all of the practice questions for each skill, the learners were engaged through trying to determine each answer. After a learner initially answered a question incorrectly, rather than immediately providing the answer, which would limit thinking, hints were given that further stimulated thinking. Each hint provided information to think about and directed the learner to apply that information to reach a solution. One or more successive hints were provided that gave yet more clues towards answering the question. The learner's focus was directed to the deeper learning concepts that supported higher-order thinking skills. As described above, this was accomplished through presenting the learner with numerous questions and ideas to consider that directly related to the higher-order thinking skill that was being taught. Assimilation was supported by having concepts learned in the sample problems directly relate to concepts addressed in the practice questions. Accommodation was supported by presenting new relationships in problems within the practice questions. Principles of the constructivism theory were followed. Some examples include supporting assimilation and accommodation, the high degree of active learning, the practice and feedback provided, and learning through self-paced materials. Metacognition and self-reflection were encouraged through a number of statements that supported those concepts.

Providing Learning Guidance

Providing learning guidance is used to help students learn the material, such that the content is stored into long-term memory in a meaningful way that allows the student to retrieve the content from long-term memory (Fenrich, 2014; Gagné et al., 1988; Singh, 2010). In the materials of this study, guidance for how to solve each of the skills was provided through presenting the initial samples, stating what needed to be done, and providing increasingly informative hints. The guidance was designed to help the student store the concepts in his or her long-term memory. Guidance was also provided through directly stating, in the introduction to the samples, introduction to the practice questions, and summary, what needed to be done to answer the questions.

Eliciting the Performance and Providing Feedback

The purpose of eliciting the performance and providing feedback is that learners must find out how well they are doing and how they can improve their performance. This can be done by asking questions or providing opportunities to practice the skill and then giving elaborative feedback. The questions asked should be at the highest appropriate thinking level to promote that level of skill development and should be asked throughout the learning rather than massed together, such as at the end of the materials. Eliciting the performance and providing elaborative feedback are typically integrated together so that students can immediately see and understand consequences of their actions (Fenrich, 2014; Gagné et al., 1988). Answering questions is often an important part of an instructional strategy as it is more effective than being told information (Baumer, 2009). When teaching higher-order thinking skills, it is particularly important to provide practice and feedback as most students do not fully learn the skills without practice and feedback (Mains, 1997). In the materials created for this research, subjects were given numerous questions throughout the materials to practice the concepts at the highest needed thinking level. For each question, feedback in the form of hints of increasing detail was given. As well, for both correct and incorrect answers of the practice, self-test, and challenge questions, elaborative feedback was provided that detailed what the right answer was, why the answer was right, and why the other answers were incorrect. Elaborative feedback was provided for correct answers in case the learner guessed the right answer or got it right for the wrong reasons. Students only received a recommendation to proceed in the lesson if he or she scored a perfect mark or only made one mistake on the self-test. The rationale was that higher expectations tend to lead to increased learning.

Assessing Performance

Learners are formally tested in the assessing performance event. This step is more formal than eliciting the performance. All learning outcomes and only the learning outcomes should be tested. Test performance should be based on achieving the specified learning outcomes (i.e., criterion-referenced) as opposed to comparing students to each other (i.e., normative-referenced) (Fenrich, 2014; Gagné et al., 1988; Singh, 2010). For this study, a formal assessment was done on each logical-thinking skill.

Enhancing Retention and Transfer

In the enhancing retention and transfer event, students should be supported in retaining the information and transferring the skills beyond the specific ideas presented in the learning materials. Retention increases as exposure increases. Retention can also be supported through providing summaries. Retention activities should occur at spaced intervals and occur before more complex skills are learned. Transfer should be deliberately addressed to make transfer more likely to occur. Transfer can be facilitated by providing real-life, novel, and/or varied problems and solutions. As well, transfer is more likely as the amount of practice and feedback increases and if the skills measured are of the near transfer type (Fenrich, 2014; Gagné et al., 1988). For each of the logical-thinking skills of the instructional materials for the interventions of this study, retention was reinforced through numerous interactions and with a summary that reiterated the strategy of how to solve the logical-thinking skill.

Methods

Introduction

The topics discussed in the methodology are research design, the population and sample, research instruments, data collection procedures, and data analysis procedures.

Research Design

A quantitative assessment was done using a pre-test, post-test, experimental design to assess the effectiveness of the closely-matching paper-based and educational-software materials in teaching logical-thinking skills.

Table 1. Differences between the paper-based and educational-software interventions

Paper-based Intervention	Educational-software Intervention
The menu page listed page numbers with each menu item so that a learner could easily find each section.	The menu page did not list page numbers. Clicking on any menu item immediately branched a learner to that section.
If subjects wanted to view a specific section or page, the subject had to spend time to manually find it.	If subjects wanted to branch to a specific section or screen, the subject could click buttons to quickly branch there.
Black and white due to the extra cost of printing thousands of pages in colour. Correct answers were highlighted in bold.	Yellow orientation information, white text, light blue prompts, correct answer feedback in yellow, incorrect answer feedback in magenta, and blue background. Correct answers were highlighted with a box.
Subjects manually flipped to the next page to see more content.	Subjects clicked on a button to move to the next screen to see more content.
Subjects read further down the page or the next page for successive hints. Subjects were told to not read the hints until the hints were needed. The hints were either separated by white space or on a subsequent page so that it was a deliberate choice to read them.	A subject was automatically provided with a hint if he or she answered a question incorrectly. A subsequent hint was provided if another mistake was made.
Subjects manually flipped to the next page to see a correct answer. They compared their answer to the provided answer.	Subjects received instant feedback on the same screen when they selected an answer.
For the main learning activity and self-test, subjects had to count their own scores and determine what they should do based on the statement provided.	For the main learning activity and self-test, the software recorded and displayed scores and provided specific advice based on a score.
Learners were challenged to see how many questions they could answer correctly on the first try but had to keep track of their own score.	Learners were challenged to see how many questions they could answer correctly on the first try. The software prompted with words like, "Correct on first try: 12 of 15".
For the deductive-reasoning skill, the learners had to manually compare their work to a series of matrices that gradually showed the solution. If a learner made a mistake, they would have to erase some of their work and try again.	For the deductive-reasoning skill, the learners received instant feedback about whether an answer was possible to know based on the information in the matrix. The learner clicked "Try again" to remove his or her last input.

The paper-based and educational-software interventions closely paralleled each other as the educational software was created from the paper-based content. There were differences between the interventions based on their own limitations or costs, as listed in Table 1. It was not believed that these differences would create a significant difference between the two experimental groups.

Pre-test and Post-test Design

An experimental pre-test, post-test, control group design was used as this design inherently has relatively high internal validity. In a foundational book, Campbell and Stanley (1966) called the pre-test, post-test, control group design a "true experimental design" and recommended the design strongly. This design has strong internal validity with respect to history, maturation, testing, instrumentation, regression, selection, mortality, and interaction of selection and maturation (Campbell and Stanley, 1966). The control group is critical. Without a control group, one cannot be sure that changes were a result of the intervention or other factors (Bessick, 2008).

As defined by Campbell and Stanley (1966) a true experimental pre-test, post-test, control group design is as follows where R represents randomization, O represents the test (pre-test and post-test), and X represents the intervention:

R	O	X	O
R	O		O

The design used for this study is depicted below, where XP represents the paper-based intervention and XE represents the educational-software intervention.

O	R	XP	O
O	R	XE	O
O	R		O

Each participant was given five pre-tests; one on each of the five logical-thinking skills that would be taught through the intervention. As recommended (Abramis, 2008), stratified random sampling, based on the cumulative results of all of the pre-test scores, was used to ensure that proportionate numbers of male and female and grade six and seven students of equivalent overall logical-thinking ability were assigned to each experimental group and the control group.

Throughout this study, all of the students were in their regular classrooms except for when the intervention took place. During the time of the intervention, the control students were in a classroom where they completed the post-test and then participated in a traditional classroom activity that was unrelated to the logical-thinking skills being taught to the experimental groups. At the same time, one experimental group worked through the educational-software intervention in the school's computer lab while the other experimental group worked on the paper-based intervention in another school classroom. The experimental students worked independently of each other and were given the time they needed to learn from the resources. Immediately after the intervention, the students in the experimental groups were given the post-test. If the post-test was not immediately completed, the students of the same or different groups could have talked with each other about the interventions when they returned to their regular classrooms. This could potentially have introduced a factor of cooperative learning that may have impacted the results (Asamani, 1998).

The independent variable was the treatment or no treatment. The dependent variable was the score in logical-thinking ability.

The Treatment

The experimental groups worked through either the educational software or the equivalent paper-based materials that contained instructional content designed to enhance logical-thinking ability. For the students of the experimental groups, the interventions taught the logical-thinking skills of classification, analogical reasoning, sequencing, patterning, and deductive reasoning.

Population and Sample

The student population was public-school grade six and seven students from one urban elementary school in Burnaby, British Columbia, Canada. The sample was drawn from the entire grade six and seven population of this school, which consisted of 102 students. 96 participants completed both the pre-tests and post-tests. Two participants, who wrote the pre-tests, did not participate in the rest of the study. One was absent due to medical reasons while the other voluntarily withdrew from the study. Three students were not invited to participate based on a decision by their teachers. One student was deficient in English language skills and two students instead attended their regular special instructional session due to their need for learning assistance. One student was not given permission to participate in the study.

Table 2. Gender of sample

Gender	Educational Software Intervention Group	Paper-based Intervention Group	Control Group	Total
Male	17	17	16	**50**
Female	15	15	16	**46**
Total	**32**	**32**	**32**	**96**

Table 3. Grade and age of sample

Grade	Age	Educational Software Intervention Group	Paper-based Intervention Group	Control Group	Total
Six	Eleven	16	16	17	**49**
Seven	Twelve	16	16	15	**47**
Total		**32**	**32**	**32**	**96**

Table 2 describes the gender of the sample. 50 participants were male while 46 were female. Table 3 describes the grade level and ages of the sample. 49 were in grade six and were eleven years old, while 47 were in grade seven and were twelve years old.

Research Instruments

As discussed in the literature review, commercial assessment tools were not suitable for the quantitative analysis. Consequently, the researcher created the assessment instruments.

Data Collection Procedures

For the pre-tests and post-tests, the input from the students of the control and experimental groups was marked and recorded manually. The data recorded was whether each question was completely right or wrong. The pre-tests and post-tests were conducted on paper. The questions were designed to be solved using the logical-thinking strategies that were presented to the experimental groups. The pre-test and post-test questions were different but comparable to the material presented in the interventions, and contained ten questions for each of the specific logical-thinking skills.

Data Analysis Procedures

The statistical tests assumed a significance of 0.05. One-way ANOVAs were conducted using the overall scores based on the sum of the test scores from all of the questions on the five logical-thinking ability tests and test scores on each logical-thinking ability test. If significance was found between the groups, a Tukey HSD Post Hoc Test was done. Paired samples t-tests were calculated to determine the percent gained for each group. The t-tests compared the pre-test and post-test scores on each specific logical-thinking skill.

Results and Findings

Introduction

The findings for the following research questions are presented below:

1. Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those taught using paper-based materials?
2. Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those not being exposed to any intervention?
3. Were there significant differences in the logical-thinking ability between subjects taught using paper-based materials compared to those not being exposed to any intervention?

where, ESG is the educational-software group, PBG is the paper-based group, and CG is the control group.

Learning Scores

Table 4 Pre-test and post-test (df=2,93) learning scores

						y
	(SD)	(SD)	(SD)			
est Total	(8.71)	(8.78)	(8.04)			
test Total	(6.06)	(6.34)	(8.21)			ESG > CG PBG > CG
Classification Skill Pre-test	6.44 (2.65)	6.03 (2.74)	5.81 (2.78)	.434	.649	

Classification Skill Post-test	7.81 (1.31)	7.88 (1.60)	6.19 (2.71)	7.58	.001	ESG > CG PBG > CG
Analogical-reasoning Skill Pre-test	7.06 (2.20)	7.19 (2.01)	6.97 (2.22)	.084	.920	
Analogical-reasoning Skill Post-test	8.68 (1.23)	9.06 (0.91)	7.45 (2.17)	9.65	.000	ESG > CG PBG > CG
Sequencing Skill Pre-test	7.72 (1.80)	7.75 (2.06)	7.28 (2.20)	.534	.588	
Sequencing Skill Post-test	8.47 (1.65)	8.47 (1.70)	7.16 (2.44)	4.74	.011	ESG > CG PBG > CG
Patterning Skill Pre-test	4.72 (2.26)	4.78 (2.47)	4.75 (2.26)	.006	.994	
Patterning Skill Post-test	6.15 (1.90)	6.28 (2.64)	4.52 (2.43)	5.62	.005	ESG > CG PBG > CG
Deductive-reasoning Skill Pre-test	7.16 (3.69)	8.63 (2.52)	9.06 (1.24)	4.44	.014	CG > ESG
Deductive-reasoning Skill Post-test	8.43 (2.02)	8.75 (1.54)	8.97 (1.43)	.84	.437	

Table 4 presents the pre-test and post-test One-Way ANOVA results from comparing the three groups (ESG, PBG, and CG) on the total scores as well as the scores on each specific logical-thinking skill test. The following findings are based on post-hoc Tukey's HSD analysis. On the post-test total scores, the ESG scored significantly higher than the CG ($p = .009$) and the PBG scored significantly higher than the CG ($p = .002$). For the classification skill, the ESG scored significantly higher than the CG ($p = .004$) and the PBG scored significantly higher than the CG ($p = .003$). For the analogical-reasoning skill, the ESG scored significantly higher than the CG ($p = .005$) and the PBG scored significantly higher than the CG ($p = .001$). For the sequencing skill, the ESG scored significantly higher than the CG ($p = .024$) and the PBG scored significantly higher than the CG ($p = .025$). For the patterning skill, the ESG scored significantly higher than the CG ($p = .017$) and the PBG scored significantly higher than the CG ($p = .010$). There were no significant differences between the ESG and the PBG on any of the above tests.

Table 5. Deductive-reasoning skill scores t-test results

Test	N	Mean difference (Post–Pre)	t (df)	p
ESG Post-test versus Pre-test	32	1.27	2.12 (31)	.042
PBG Post-test versus Pre-test	32	.125	.275 (31)	.785
CG Post-test versus Pre-test	32	–.094	–.407 (31)	.687

For the deductive-reasoning skill, post-hoc Tukey's HSD tests showed that the CG scored significantly higher than the ESG ($p = .014$) on the pre-test while there were no significant pre-test score differences between the PBG and the CG or between the ESG and the PBG. An ANCOVA, using the CG pre-test scores as a covariate, showed no significant differences between the groups. Given the CG scored significantly higher than the ESG, a Pearson's correlation coefficient was calculated. The correlation between the pre-test and post-test score was significant; $r(95) = .394$, $p = .000$. Given a significant Pearson's correlation, a paired samples t-test was calculated to assess whether there were significant differences between the pre-test and post-test scores for each group. Based on the t-test results for the deductive-reasoning skill, as summarized in Table 5, the ESG score was significantly higher on the post-test than the pre-test ($p = .042$). Neither the PBG nor CG had a significant difference between their pre-test and post-test scores.

ESG Amount of Learning

Table 6. ESG amount of learning

Test	Post-test n=32 df=1,31 Mean (SD)	Pre-test n=32 df=1,31 Mean (SD)	t	p	Percent Gain
Total Score	39.54 (6.06)	33.09 (8.71)	5.36	.000	19.5%
Classification Skill	7.81 (1.31)	6.44 (2.65)	2.78	.009	27.2%
Analogical-reasoning Skill	8.68 (1.23)	7.06 (2.20)	5.00	.000	22.9%
Sequencing Skill	8.47 (1.65)	7.72 (1.80)	3.41	.002	9.7%
Patterning Skill	6.15 (1.90)	4.72 (2.26)	3.16	.004	30.4%

Deductive-reasoning Skill	8.43 (2.03)	7.16 (3.69)	2.12	.042	17.8%

Table 6 presents the paired samples t-test results from comparing the ESG pre-test and post-test total scores and the scores on each specific logical-thinking skill. For the ESG students, there was a significant percentage gain in every pre-test to post-test score.

PBG Amount of Learning

Table 7. PBG amount of learning

Test	**Post-test n=32 df=1,31 Mean (SD)**	**Pre-test n=32 df=1,31 Mean (SD)**	**t**	**p**	**Percent Gain**
Total Score	40.43 (6.34)	34.38 (8.78)	5.90	.000	17.6%
Classification Skill	7.88 (1.60)	6.03 (2.74)	4.43	.000	30.6%
Analogical-reasoning Skill	9.06 (0.91)	7.19 (2.01)	5.92	.000	26.1%
Sequencing Skill	8.47 (1.70)	7.75 (2.06)	2.24	.033	9.3%
Patterning Skill	6.28 (2.64)	4.78 (2.47)	3.36	.002	31.3%
Deductive-reasoning Skill	8.75 (1.54)	8.63 (2.52)	.28	.785	1.4%

Table 7 presents the paired samples t-test results from comparing the PBG pre-test and post-test total scores and the scores on each specific logical-thinking skill. For the PBG students' total scores and the scores on the classification, analogical-reasoning, sequencing, and patterning skills, there was a significant percentage gain in pre-test to post-test scores. There was no significant pre-test to post-test change in the deductive-reasoning skill.

CG Amount of Learning

Table 8. CG amount of learning

Test	**Post-test n=32 df=1,31 Mean (SD)**	**Pre-test n=32 df=1,31 Mean (SD)**	**t**	**p**	**Percent Gain**
Total Score	34.29 (8.21)	33.88 (8.04)	.81	.422	1.2%
Classification Skill	6.19 (2.71)	5.81 (2.78)	1.79	.083	6.5%
Analogical-reasoning Skill	7.45 (2.17)	6.97 (2.22)	1.90	.067	6.9%
Sequencing Skill	7.16 (2.44)	7.28 (2.20)	–.49	.627	–1.6%
Patterning Skill	4.52 (2.43)	4.75 (2.26)	–.67	.510	–4.9%
Deductive-reasoning Skill	8.97 (1.43)	9.06 (1.24)	–.41	.687	–1.0%

Table 8 presents the paired samples t-test results from comparing the CG pre-test and post-test total scores and the scores on each specific logical-thinking skill. For the CG students, there was no significant pre-test to post-test change in any score.

Discussion

Introduction

This research led to the development of instructional materials that aimed to teach specific logical-thinking skills and the assessment of whether the instructional materials could effectively teach those skills. This section presents a discussion on the findings for each research question.

Discussion Regarding Research Question 1

Research Question 1: Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those taught using paper-based materials?

Through ANOVA and Tukey HSD Post Hoc Test calculations, for every test, there were no significant differences between the ESG and PBG. These findings are harmonious with Clark (1983) who stated,

Consistent evidence is found for the generalization that there are no learning benefits to be gained from employing any specific medium to deliver instruction. Research showing performance or time-saving gains from one or another medium are shown to be vulnerable to compelling rival hypotheses concerning the uncontrolled effects of instructional method and novelty (p. 445).
This matches the literature where Cott (1991), Heo (2012), Larson et al. (2009), Shiah (1994), and Titterington (2007) have found no significant differences between comparable interventions. No studies were found where comparable interventions led to significant differences between the experimental groups.

Discussion Regarding Research Question 2

Research Question 2: Were there significant differences in the logical-thinking ability between subjects taught using educational software compared to those not being exposed to any intervention?

Through ANOVA and Tukey HSD Post Hoc Test calculations, for the total post-test scores and the classification, analogical-reasoning, sequencing, and patterning skills scores, the ESG scored significantly higher than the CG. For the deductive-reasoning skill scores, the CG scored significantly higher than the ESG on the pre-test. However, for the deductive reasoning test, a paired samples t-test showed that the ESG score was significantly higher on the post-test than the pre-test (a 17.8% gain) while the CG pre-test to post-test score was insignificantly different (a 1.0% decrease).

With respect to logical-thinking skills, as were taught in the interventions of this study, Larson et al. (2009) and Swan (1990) have also found at least some significant gains with stand-alone educational software. However, in comparison to this study, neither of these researchers assessed the same logical-thinking skills and neither of the studies only included grade six and seven students.

Discussion Regarding Research Question 3

Research Question 3: Were there significant differences in the logical-thinking ability between subjects taught using paper-based materials compared to those not being exposed to any intervention?

Based on ANOVA and Tukey HSD Post Hoc Test calculations, for the total scores and the classification, analogical-reasoning, sequencing, and patterning skills scores, the PBG scored significantly higher than the CG. For the deductive-reasoning skill scores, the PBG and CG did not have a significant difference between their pre-test and post-test scores. This may have been due to the PBG being overwhelmed with the amount of material. The students seemed to lose their concentration about halfway through the intervention. The intervention was 159 pages of double-sided printing, which was far thicker than typical assignments (D. Moore, personal communication, May 12, 2009). The PBG may have lost their motivation to work through the materials (Fenrich, 2014). A lack of motivation is consistent with the results where the PBG and CG pre-test to post-test score changes were insignificant (respectively a 1.4% gain and a 1.0% decrease), whereas the ESG had significant gains (respectively a 17.8% and 19.1% gain) with comparable materials, based on paired samples t-test results. In stand-alone educational software interventions, the amount of content is not as readily apparent as with paper-based materials where learners can visually see the thickness of the handout. With respect to higher-order thinking skills taught through paper-based materials, Larson et al. (2009) and Titterington (2007) also found significant gains. However, in comparison to the interventions of this study, neither of these researchers assessed the same thinking skills and neither of the studies only included grade six and seven students.

Discussion Regarding Gains in Amount Learned

Paired samples t-tests were conducted that compared the pre-test and post-test scores on each test. For the ESG students' total scores and the scores on each test, there was a significant percentage gain in every pre-test to post-test score. For the PBG students' total scores and the scores on each test, there was a significant percentage gain in every pre-test to post-test score, except for the deductive-reasoning skill, as discussed above. For the CG students, there was no significant pre-test to post-test change in any score.

In general, both the educational-software and paper-based interventions led to significant gains in logical-thinking ability. With respect to the significant pre-test post-test gains found from the total scores, these findings are consistent with other researchers who reported that learners could be taught a significant amount of higher-order thinking skills. These researchers include Allison (1993), Bradberry-Guest (2011), Campbell (2000), Cotton (1991), Duffield (1989), Galinski (1988), Judy (1987), Katzlberger (2006), Larson et al. (2009), Leiker

(1993), Lewis (1998), Orabuchi (1992), Pogrow (2005), Shiah (1994), Swan (1990), 1999), and Titterington (2007).

In this study, there were significant differences in the scores. However, the findings were not significant in a few of the tests for specific logical-thinking skills. Lee (2008) had similar findings where there was a significant difference based on the whole score but no significant differences in critical thinking based on some individual items. Mixed findings are consistent with other researchers. These researchers include Campbell (2000), Swan (1990), and (1999). Explanations for the lack of a significant difference on specific logical-thinking tests are discussed above.

It was expected that the experimental groups would perform significantly higher on the post-tests as compared to the pre-tests because of the Combined Instructional Design and Development Model followed and the instructional materials contained many features that support effective teaching. The techniques used to gain and maintain attention included asking the students to obtain high scores, stressing the importance of thinking carefully, posing challenging statements and questions, and making the materials highly interactive. These techniques also helped to motivate the students. The learner was informed of each learning outcome as the learning outcome was directly presented within the instructional materials. The content was matched to the anticipated cognitive development of the students. The concepts gradually increased in difficulty in that the material was presented in small incremental steps. A variety of instructional activities and strategies were created. The activities supported each learning outcome. There was a high degree of active learning in the samples, practice questions, self-test, and challenge questions. Active learning is a cornerstone of the constructivist theory of learning. The learner’s focus was directed to the deeper learning concepts that supported higher-order thinking skills. The assimilation and accommodation of knowledge into long-term memory was supported. Metacognition and self-reflection were encouraged. There was guidance for how to solve each of the skills through the presentation of the initial samples, and direct statements of what needed to be done to answer the questions. Feedback contained hints with increasing detail and the final elaborative feedback explained why the answer was correct and why other answer choices were incorrect, as was appropriate (Fenrich, 2014; Gagné et al., 1988; Wu, 2009).

It was expected that the control group would not have any significant gains because students tend not to improve higher-order thinking skills when there is no explicit intervention in place (Jeremiah, 2012; Wruck, 2010).

Implications

Based on the findings of this study, there are a number of implications:

The instructional interventions were created through following the Combined Instructional Design and Development Model, which was based on a nonlinear instructional development cycle, a systematic process of instructional design, and Gagné’s Nine Events of Instruction. Others can use this model to create interventions that teach logical thinking or other higher-order thinking skills.

Logical-thinking skills can be taught through standalone educational-software and paper-based materials. This suggests that other logical-thinking and higher-order thinking skills can be taught through stand-alone educational-software and paper-based materials. Presumably, other stand-alone delivery methods, such as web-based interventions, could also be used to teach logical-thinking and higher-order thinking skills. Given effective instructional design on stand-alone materials, students can learn logical-thinking skills and likely other higher-order thinking skills, whether or not their teachers address the skills or have the ability to teach the skills.

Improving logical-thinking skills needs an explicit intervention in that the control group did not have any significant increase in pre-test to post-test scores, whereas both experimental groups gained a significant amount of logical-thinking skills.

Limitations

With respect to this study, there were a number of factors that limit its usefulness:

The instructional interventions were developed through following the created Combined Instructional Design and Development Model. Conducting research using a different instructional design and development model could impact the results.

If this study was to be repeated with different reviewers and a researcher with a different instructional design background, the resulting interventions could look substantially different.

In terms of generalizability, the findings are limited to logical-thinking skills rather than higher-order thinking skills in general. As well, the findings are limited to the five logical-thinking skills of classification, analogical reasoning, sequencing, patterning, and deductive reasoning, as specifically defined, rather than all logical-thinking skills. The findings are also limited to grade six and seven students.

Suggestions for Further Research

With respect to further research, a number of studies could contribute to the limited knowledge base regarding teaching higher-order thinking skills in a stand-alone mode:

For grade six and seven learners, develop and assess materials that teach other logical-thinking or higher-order thinking skills. Use the Combined Instructional Design and Development Model or something comparable as the foundation for creating the materials.

Develop and assess materials that teach logical-thinking or higher-order thinking skills to younger and/or older learners. The age-appropriateness of the materials can be assessed beforehand. When doing this with younger learners, determine whether the learners have the maturity to thoroughly complete the tasks.

Measure attitudes towards the materials before and after the intervention to determine if there is a difference between an educational-software intervention and paper-based materials.

Assess whether gains in logical-thinking and higher-order thinking skills are maintained over time.

Compare learners working through the materials cooperatively in dyads and triads to those learning individually to determine if cooperative learning should be recommended as a part of future interventions.

Determine if there are differences in using stand-alone materials that teach logical thinking or other higher-order thinking skills in schools where rote memorization is the norm through to schools where thinking skills are regularly emphasized. Determine what, if anything, needs to be done to create effective materials for individuals who have different histories with respect to higher-order thinking skills.

Determine what needs to be done to change educational systems, ranging from government standards or requirements through to the classroom teacher, so that higher-order thinking development becomes a high enough priority so that typical graduates leave the school system with the higher-order thinking skills needed in both the workplace and life itself.

Conclusion

This study's objectives were to quantitatively assess the efficacy of the materials. Given the limitations of this study, the quantitative findings showed that both the educational software and paper-based interventions led to significant gains in logical-thinking skills, although there were no significant gains in a specific skill.

This research helps to fill the gap in research regarding teaching logical-thinking skills, particularly with standalone educational software and paper-based materials that aim to teach grade six and seven students logical-thinking skills. However, more research and interventions are needed to fully solve the problem of students leaving the school system without the level of logical-thinking skills needed to reach one's full potential in both the workplace and life in general.

References

Abrami, P., et al. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research 78*(4), 1102-1134.

Allison, A. K. (1993). *Critical thinking/problem-solving skills for the at-risk student* (Doctoral dissertation, Northern Arizona University, 1993).

Almatrodi, D. (2007). *Effective critical thinking teaching strategies as perceived by affiliated program evaluation faculty* (Doctoral dissertation, Western Michigan University, 2007).

Asamani, J. B. (1998). *An experimental study of the relationship between cooperative learning and critical thinking* (Doctoral dissertation, Tennessee State University, 1995).

Bass, G. M., & Perkins, H. W. (1984). Teaching critical thinking skills with CAI. *Electronic Learning 14*(2), 32, 34, 96.

Baumer, E. (2009). *Computational metaphor identification to foster critical thinking and creativity.* (Doctoral dissertation, University of California, 2009).

Bessick, S. (2008). *Improved critical thinking skills as a result of direct instruction and their relationship to academic achievement* (Doctoral dissertation, Indiana University of Pennsylvania, 2008).

Bradberry-Guest, K. (2011). *Effects of computer-based intervention on higher order thinking skills and implications for response to intervention (RTI)* (Doctoral dissertation, Walden University, 2011).

Burkhart, L. M. (2006). *Thinking critically about critical thinking: Developing thinking skills among high school students* (Doctoral dissertation, Claremont Graduate University, 2006).

Campbell, D., & Stanley, J. (1966). *Experimental and quasi-experimental designs for research*. Chicago, Il: Rand McNally Publishing Company.

Campbell, J. P. (2000). *A comparison of computerized and traditional instruction in the area of elementary reading* (Doctoral dissertation, University of Alabama, 2000).

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445-459.

Cott, A. M. (1991). *On the efficacy of programmed instruction on training creative problem solving strategies* (Doctoral dissertation, Hofstra University, 1991).

Cotton, K. (1991). *Close-Up #11: Teaching Thinking Skills*. Retrieved April 2, 2016, from http://educationnorthwest.org/sites/default/files/TeachingThinkingSkills.pdf

Dick, W., & Carey, L. (1990). *The systematic design of instruction* (3rd ed.). Glenville, IL: Harper Collins Publishers.

Duffield, J. (1989). *Problem-solving software: What does it teach?* (Doctoral dissertation, Florida State University, 1989).

Ellingwood, M. (1997). *The effects of Logo instruction on the higher-order thinking skills and mathematical achievement on male and female first grade students* (Doctoral dissertation, University of South Alabama, 1997).

Enniss, I. (2006). *A qualitative analysis of the kinds of higher-order thinking processes present in the reflective journals of graduate students of reading theory* (Doctoral dissertation, Wayne State University, 2006).

Fenrich, P. (2014). *Practical principles of instructional design, media selection, and interface design with a focus on computer-based training / educational software*. Santa Rosa, CA: Informing Science Press.

Gagné, R., Briggs, L., & Wager, W. (1988). *Principles of instructional design* (3rd ed.). New York, NY: Holt, Rinehart and Winston.

Galinski, M. (1988). *An investigation of the relationship of microcomputer problem-solving software and critical thinking abilities of junior-high students* (Doctoral dissertation, Northeastern University, 1988).

Heo, J. (2012). *The effect of using three types of instructional media on comprehension and motivation of Korean college students in an online course* (Doctoral dissertation, University of Nevada, 2012).

Jeremiah, K. (2012). *Understanding approaches to teaching critical thinking in high school classrooms* (Doctoral dissertation, Walden University, 2012).

Judy, J. E. (1987). *Effects of two instructional approaches and peer tutoring on gifted and nongifted sixth graders' analogy performance* (Doctoral dissertation, Texas A&M University, 1987).

Katzlberger, T. (2006). *Learning by teaching agents* (Doctoral dissertation, Vanderbilt University, 2005).

Lafferty, I. (1996). *Analysis of the computerized structure of intellect teaching thinking intervention* (Doctoral dissertation, Havard University, 1996).

Lamb, A. (2001). *Critical and Creative Thinking – Bloom's Taxanomy*. Retrieved April 2, 2016, from http://eduscapes.com/tap/topic69.htm.

Larson, M. E. (2010). *Transfer of learning: The effects of different instruction methods on software application learning* (Doctoral dissertation, Capella University, 2010).

Larson, A., Britt, A., & Kurby, C. (2009). Improving students' evaluation of informal arguments. *The Journal of Experimental Education 77*(4), 339-366.

Lee, J. S. (2010). *Outcomes and implications of designing and implementing a project-based learning geometry unit* (Doctoral dissertation, Indiana University, 2010).

Lee, Y. (2008). *The effects of different levels of vicarious interactions on critical thinking* (Doctoral dissertation, Purdue University, 2008).

Leiker, V. (1993). *The relationship between an integrated learning system, reading and mathematics achievement, higher-order thinking skills and certain demographic variables: A study conducted in two school districts* (Doctoral dissertation, Baylor University, 1993).

Lewis, R. B. (1998). *Developing critical thinking through an interdisciplinary approach with social studies simulations and technology in fourth grade classrooms* (Doctoral dissertation, Tennessee State University, 1998).

Liu, C. (2006). *Connecting the dots between critical thinking skills and dispositions for middle school students* (Doctoral dissertation, Wayne State University, 2006.

Mains, M. G. (1997). *The effects of learning a programming language on logical thinking skills* (Master's degree thesis, University of Nevada, 1997).

Masteron, J., & Perrey, C. (1999). Training analogical thinking skills in children with language disorders. *American Journal of Speech-Language Pathology 8*(1), 53-61.

Mintz, K. (2000). *A comparison of computerized and traditional instruction in elementary mathematics* (Doctoral dissertation, University of Alabama, 2000).

Orabuchi, I. (1992). *Effects of using interactive CAI on primary students' higher-order thinking skills: Inferences, generalizations, and math problem-solving* (Doctoral dissertation, Texas Woman's University, 1992).

Pogrow, S. (2005). HOTS revisited: A thinking development approach to reducing the learning gap after grade 3. *Phi Delta Kappan 87*(1), 64-75.

Reddy, S. (2008). *Developing, implementing and evaluating computer-based instruction* (Doctoral dissertation, Indiana State University, 2008).

Repman, J. L. (1989). *Cognitive and affective outcomes of varying levels of structured collaboration in a computer-based learning environment* (Doctoral dissertation, Louisiana State University, 1989).

Rivière, F. (1990). *Developing problem solving abilities with artificial intelligence* (Master degree thesis, University of Windsor, 1993).

Rothman, A. H. (2000). *The impact of computer-based versus "traditional" textbook science instruction on selected learning outcomes* (Doctoral dissertation, Temple University, 2000).

Rukavina, P. (2003). *The effect of scaffolding movement challenges on students' task-related thoughts and performance* (Doctoral dissertation, Louisiana State University, 2003).

Schmidt, D. (1991). *An investigation of the effect of gender, cognitive level, and attitude of seventh-grade science students on an identification of key variables task using an expert system* (Doctoral dissertation, Florida Atlantic University, 1991).

Semper Scott, L. D. (2005). *The cognitive emphasis of instruction as demonstrated by assessment practices* (Doctoral dissertation, American University, 2005).

Shiah, R. (1994). *The effects of computer-assisted instruction on the mathematical problem solving of students with learning disabilities* (Doctoral dissertation, Purdue University, 1994).

Singh, O. B. (2010). *Development and validation of a web-based module to teach metacognitive learning strategies to students in higher education* (Doctoral dissertation, University of South Florida, 2009).

Sondel, H. B. (2009). *The effects of curricular programs on aspects of critical thinking as applied to writing* (Doctoral dissertation, University of Virginia, 2009).

Swan, K. (1990). *Logo programming and the teaching and learning of problem solving* (Doctoral dissertation, Columbia University, 1989).

A. (1999). *The effects of computer-based instruction on the academic achievement of at-risk students*. (ERIC Document Reproduction Service No. ED457855).

Thomson, S. (2009). *Effects of Primary Education Thinking Skills™ on achievement: Mixed-method evaluative study* (Doctoral dissertation, Aurora University, 2009).

Titterington, L. C. (2007). *Case studies in pathophysiology: The development and evaluation of an interactive online learning environment to develop higher order thinking and argumentation* (Doctoral dissertation, Ohio State University, 2007).

vonStein, J. H. (1982). *An evaluation of the microcomputer as a facilitator of indirect learning for the kindergarten child* (Doctoral dissertation, Florida Atlantic University, 1982).

Wruck, L. M. (2010). *Computer-mediated communication: Instructional design strategies that support the attainment of Bloom's higher order cognitive skills in asynchronous discussion questions* (Doctoral dissertation, Capella University, 2010).

Wu, X. (2009). *Improving children's engagement and learning through free-flowing discussions: Impact of collaborative reasoning discussion* (Doctoral dissertation, University of Illinois, 2009).