Journal of Educational Psychology
1999, Vol. 91, No. 2, 334-341

0022-0663/99/$3.00

A Comparison of Cognitive Load Associated With Discovery Learning and Worked Examples

Juhani E. Tuovinen
Charles Sturt University

John Sweller
University of New South Wales

This article reports experimental work comparing exploration and worked-examples practice in learning to use a database program. Exploration practice is based on discovery learning principles, whereas worked-examples practice arose from the development of cognitive load theory. Exploration practice was expected to place a considerable load on working memory, whereas a heavy use of worked examples was hypothesized to lead to more effective processing by reducing extraneous mental load. Students with no previous domain familiarity with databases were found to substantially benefit from worked examples in comparison to exploration. However, if students had previous familiarity with the database domain, the type of practice made no significant difference to their learning because the exploration students were able to draw on existing, well-developed domain schemas to guide their exploration.

Discovery learning requires learners to discover concepts and procedures that might otherwise be communicated by direct instruction. Pure discovery involves almost no structure or guidance and is rarely favored with most support reserved for variations of guided discovery that include some degree of structure to guide learners in the discovery process. The procedure has been extensively promoted as a desirable approach to learning since its development during the 1950s and 1960s (Anthony, 1977; Bell, 1991; Blanchette & Brouard, 1995; Bruner, 1959, 1961, 1962; Maor & Taylor, 1995; Nuthall & Snook, 1973; Sfondilias & Siegel, 1990; Shulman & Tamir, 1973; Whitehead, 1978). Since its early days, it has mutated into exploration learning (Charney, Reder, & Kusbit, 1990; Crawford, 1995; diSessa, Hoyles, Noss, & Edwards, 1995; Hoyles, 1995; Hsu, Chapelle, & Thompson, 1993; Kamouri, Kamouri, & Smith, 1986; Njoo & de Jong, 1993; Vitale, 1995), inductive learning (Holzman, Pellagrino, & Glaser, 1983; Rieber & Parmley, 1995), and other variants (Greeno, Collins, & Resnick, 1996).

However, it has also been attacked since its early evolutionary period for being poorly defined (Wittrock, 1966), inefficient (Carlson, Lundy, & Schneider, 1992; Cronbach, 1966; Wittrock, 1966), confusing means with ends (Wittrock, 1966), requiring many trials for concept learning (Gagné, 1966), being slower than reception learning for learning principles and rules (Gagné, 1966), developing heuristics of discovery that are not generalizable (Ausubel, 1961), requiring substantial guidance (Gagné & Brown, 1961), and therefore being of only limited value in the educational process (Carlson et al., 1992; Cronbach, 1966; Maor & Taylor, 1995; Mayes, 1992; Wittrock, 1966). Indeed, Bruner (1966, p. 101), referring to students learning material that has become part of a cultural heritage, commented that discovery learning is "... the most inefficient technique possible for regaining what has been gathered over a long period of time..." However, it is still supported by many teachers and educational writers (Maor & Taylor, 1995). In this article, we report experimental work in the context of current cognitive theory to assess some aspects of discovery or exploration learning.

Cognitive Load Theory

Cognitive load theory (see Sweller, van Merriënboer, & Paas, 1998) derives instructional design principles from aspects of our cognitive architecture. The theory assumes a very limited working memory (Miller, 1956), an effectively unlimited long-term memory (Simon & Gilmartin, 1973) holding large numbers of schemas (Chi, Glaser, & Rees, 1982) that can vary in their degree of automaticity (Kotovsky, Hayes, & Simon, 1985). This architecture interacts with instructional material in various ways.

First, different learners will process the material in different ways. If the elements of material that require processing are incorporated in an automated schema, working-memory load (or cognitive load) will be low. Schemas allow many elements to be treated as a single element in working memory, and automatic processing limits working-memory demands compared with controlled, conscious processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). As a consequence, if a learner has acquired appropriate automated schemas, cognitive load will be low, and substantial working-memory resources are likely to be free. In contrast, if the elements of material that require processing must each be considered as a discrete element in working memory because no schema is available, cognitive load will be high. Working memory may be entirely occupied in processing large numbers of individual elements.

Juhani E. Tuovinen, School of Education, Charles Sturt University; John Sweller, School of Education Studies, University of New South Wales.

We thank the students of Charles Sturt University for participating in this study.

Correspondence concerning this article should be addressed to Juhani E. Tuovinen, Centre for Learning and Teaching Support, Monash University, Churchill, Victoria, Australia, 3842. Electronic mail may be sent to juhani.tuovinen@celts.monash.edu.au

Second, the characteristics of the instructional material are important. Some material can be learned element by element without relating one element to another. Learning a vocabulary provides an example. Each vocabulary item can be learned without reference to any other item. Such material is low in element interactivity and low in intrinsic cognitive load. It imposes minimal demands on working memory. Alternatively, situations where a number of elements must be considered simultaneously for the successful execution of a task are called high element interactivity tasks. Learning the order of words in English provides an example. Word order cannot be learned without considering several words simultaneously. Under these circumstances, intrinsic cognitive load is high because of high element interactivity. These situations occur often in mathematics, computer programming, design development, etc., where no individual component can be considered in isolation because any action on a given component will have complex and far-reaching effects on the task outcome.

Third, the characteristics of the learner and the material to be learned interact. Material that imposes a heavy cognitive load for some people because they must deal with large numbers of interacting elements may impose less of a cognitive load for other people because they have acquired automated schemas that incorporate the individual elements. An expert in elementary algebra will treat the equation $(a + b)/c = d$ as a single, automated schema requiring limited working-memory resources. A novice who has just commenced learning algebra may need to treat each symbol and relations between symbols as individual, interacting elements, resulting in a working-memory overload.

This theory has proved beneficial for the improvement of the planning, organization, and implementation of learning in many fields. It is argued that in the process of dealing with information, working memory has only a limited processing capacity available to deal with distinct items at any given time and that the capacity of working memory is often overloaded because of inappropriate presentation of material or inappropriate learner activities, leading to a reduction in learning and the capacity to solve problems. Thus, new material is learned most effectively and efficiently if the unnecessary cognitive load is reduced to a minimum.

The cognitive load associated with any task consists of two parts. There is the intrinsic or natural cognitive load, that is, the inherent aspects of the mental task that must be understood for the learner to be able to carry out the task. Intrinsic cognitive load is determined by levels of element interactivity. However, in addition, there is usually a range of extraneous matters associated with the way the instructional material is taught that may add to the inherent nucleus of the intrinsic cognitive load (Sweller et al., 1998). This category of cognitive load is classed as extraneous cognitive load.

In an effort to reduce the extraneous cognitive load associated with the use of problem-solving search strategies, several researchers have tested the effects of using worked examples. Worked examples require the reader to attend to problem states and their associated moves rather than *searching for the* right moves involved in conventional problem solving. In situations where an extraneous cognitive load existed because of problem-solving search, worked examples were found to effectively reduce that load and enhance learning (e.g., Carroll, 1994; Cooper & Sweller, 1987; Paas, 1992; Paas & van Merriënboer, 1994; Sweller & Cooper, 1985; Trafton & Reiser, 1993; Zhu & Simon, 1987).

This article deals with a comparison of two approaches in learning to use a computer database program. The first method is based on exploratory or discovery learning principles that, as indicated above, have been widely promoted since the 1960s at various levels of education. The second method is based on cognitive load theory with specific emphasis on worked examples. Discovery learning is contrasted with worked examples because the two procedures require differential problem-solving search. Cognitive load theory suggests that search imposes an extraneous cognitive load that interferes with learning (e.g., Sweller, 1988). Discovery learning assumes that search is beneficial with any guidance incorporated merely to assist the search process. Worked examples are designed to reduce search in order to reduce cognitive load (e.g., Cooper & Sweller, 1987).

Recent Work Comparing Direct Instruction and Discovery-Based Techniques in Learning to Use a Computer Application

Charney et al., (1990) described differences between tutorials, problem solving, and learner exploration in learning a computer application. The tutorials condition consisted of hands-on, example-based instructional situations. Participants were given basic information and then were required to follow through examples as presented in the tutorial manual by entering the information in the computer program. In the problem-solving practice condition, the participants were presented with new information on how to achieve certain computer tasks and then were provided with specific problems to solve. The learner exploration practice condition also involved the presentation of information as in the above two approaches and practice with the new skills, but the problems in this mode were set by the participants themselves rather than being externally imposed. Thus, in terms of guidance or structure, the tutorial condition provided the most structure in that participants were presented problems along with their solutions, that is, worked examples; the problem-solving condition had less structure and more discovery because participants also were presented with problems but had to discover the solutions to the problems themselves; the exploration condition had the least structure and most discovery in that participants had to devise their own problems.

Charney et al. used these three instructional approaches to teach three groups of University undergraduates to use VisiCalc (a spreadsheet program with a command line interface). The problem-solving group took the longest amount of training time. The least amount of time was taken by the tutorial group, which had detailed steps provided in the user manual to be entered on the computer. The exploration group's mean training time was only slightly longer than the tutorial group.

The problem-solving group had the fastest mean time for reaching the correct solution on a subsequent test, followed by the exploration group and the tutorial group. The problem-solving group also had the highest mean percentage of correct solutions, with the exploration and the tutorial group mean percentages for correct solutions being almost identical. The exploration group tended to practice and become more competent in a more limited range of spreadsheet features than the interactive tutorial or problem-solving groups.

The poor showing of the exploration approach contrasts with the support discovery or exploration learning has enjoyed in the pedagogical circles over many years. The poor showing of the tutorial group contradicts expectations based on cognitive load theory and the results of experiments on worked examples, which suggest that showing learners a procedure reduces cognitive load and so is superior to having them work it out themselves by problem solving.

Charney et al.'s (1990) results were confounded by differences in training times, with the groups requiring more training time performing better during the test. A series of regression analyses were carried out demonstrating that training times did not have an appreciable effect on test performance scores. Nevertheless, the statistical procedures may not have eliminated the effect of training times on test scores because two factors, instructional procedure and student ability, determine training times, and these two factors have contrary effects on performance: giving students more time to learn because of instructional procedure should increase test performance, but in contrast, if students require more time to learn because they are less able, that should decrease test performance. Thus, the first factor (instructional procedure) may tend to improve test performance with increased training times, but the second factor (individual differences) may tend to decrease test performance with increased training times. Tests between group means will not reflect individual differences in ability, only instructional procedures. Regression analyses on individual scores will reflect both instructional procedures and individual differences with the two factors tending to cancel each other out. The net result could be differences between instructional procedures because of training times but no regression effects because of training times. These results were obtained by Charney et al.

The current experiment compared a tutorial/problem-solving approach with problem-solving/exploration. Both conditions had identical training times. Cognitive load theory predicts that the greatest cognitive load is placed on the participants in the problem-solving/exploration group. The worked-examples group's cognitive load should be reduced because learners do not have to use mental resources in defining the task parameters.

An exploration and a worked-example group studied a database program with a graphical user interface. The same learning time was allowed for the two groups with each of the participants having an equal time to read the computer-aided instruction lessons and given the same amount of time for practice. Testing consisted of database construction, modification, and operational tasks. The questions were provided on paper to test the synthesis of learned skills to see how well the learned items are available when required for constructive tasks, especially when the tasks involved integration of multiple, high interactivity items. Students in the two groups were expected to use the database skills learned in this program in subsequent work. They were expected to apply the database concepts in a variety of subject teaching contexts in secondary schools.

Method

Participants

The participants were 32 Diploma of Education students at Charles Sturt University, Wagga Wagga, Australia. All students had completed at least a first degree at a University and were studying to become secondary teachers. All of them had previous experience with the Macintosh computers used in this study during the 3 previous weeks in this subject. Some students had not used computers before studying this subject, whereas others had used computers previously, either through casual contact or introductory computing courses. Fifteen students had used database programs previously. None of the participants had used the FileMaker Pro database program previously.

Procedure and Materials

The experiment was conducted over 3 weeks. During their normal 2-hour class in the first week, the students were informed of the experiment and invited to participate. After completing a questionnaire about their computing experiences, they were introduced to the construction of FileMaker Pro databases, the modification of database layouts, and searching and finding items on databases. FileMaker Pro is a graphical user interface database, where a new database can be constructed by making choices from available selection tables or windows and some typing. Items that need to be typed during the database construction consist mainly of field names, rather than commands for creating fields or specifying them as text, number, etc., types.

A classroom instruction approach was used initially, followed by database creation and layout modification tasks to be executed on the computer. Flowchart summaries of the operation steps were introduced at this stage. All the work during this session was common to all the participants.

After the first lesson, the students were randomly assigned to either the worked-examples or exploration group for the subsequent instruction and practice phase. There were 32 students, 16 in the worked-examples group and 16 in the exploration group. For instructional consistency between three separate class times during the second week, the instruction was delivered via HyperCard lessons, and the students were allowed to practice for the same amounts of time. The times allocated to the lessons and practice sessions were

Calculation fields HyperCard lesson	10 min
Calculation fields practice	40 min
Summary/totals fields HyperCard lesson	15 min
Summary/totals fields practice	40 min

In the calculation-fields lesson, the students were shown how to construct composite fields using existing field values as variables, for example, to compute the total carbon pollution by adding carbon dioxide and carbon monoxide masses recorded in two fields. In the summary-fields lesson, they were shown how to compute averages of all the records in a given field, for example, the averages of gas pollution for different gases throughout the world.

First Instructional and Acquisition Phases

In the first instructional phase, all the students read the common calculation-field HyperCard lessons on the computers for 10 min. In the subsequent acquisition phase, the exploration group was provided with printed instructions to explore the construction and use of the fields they had just read about, that is, "Try out the functions in each of the lessons in situations you create yourself, saving your files on the floppy disk provided. You may use any of the databases on the floppy disk if you wish. You will be asked to solve problems similar to the one shown in the lessons, in the test on this work. So direct your exploration towards gaining adequate mastery of the program to deal with such questions." The exploration group was expected to use the FileMaker Pro program and the provided data files on the disk to try out the calculation-field construction techniques shown in the lesson, with the aid of the flowchart of steps involved.

The worked-examples group was directed to work through the problems provided by the printed practice instructions. They were asked to read the first problem and its worked-out solution, then use FileMaker Pro, a database provided on disk, and the flowchart summary of steps required to carry out the second calculation-field construction task on the computer.

Similarly, after reading the third problem and its comprehensive solution, they were expected to attempt the subsequent calculation-field tasks on the computer. Each worked example consisted of a problem statement related to calculation or field construction or use and then an annotated step-by-step example of the way the problem could be solved with computer-screen views seen by the operator working to obtain the solution. Throughout their work on tasks three to six, the students had available printed task descriptions, flowcharts summarizing steps required, and databases on disk.

The first calculation-field question required the students to prepare a calculation field involving a division of one existing field by another. Questions 2 and 3 required the participants to develop calculation fields to add the values in the database fields with seven field values to be added in Question 2 and four in Question 3. Questions 4, 5, and 6 all required averages to be computed.

Second Instructional and Acquisition Phases

A similar sequence of HyperCard instruction and practice by exploration or by reading and working on the computer were also provided for the summary/totals field work in the second half of the same 2-hour session. All the students were provided with the same HyperCard instruction stack, databases on disks, flowchart step summaries, and FileMaker Pro database program but with different printed instructions for practice. The worked-examples group had a total of six practice tasks. The first and the third problems had worked-out solutions in the same format as for calculation-field practice, which these students were asked to read before attempting to carry out the subsequent tasks. The worked examples occupied from five to eleven pages each. Questions 1 to 4 required the calculation of averages of two items across the databases. Question 5 required computations of averages and maximum and minimum amounts and an interpretation of the results. Question 6 required a computation of an average amount and an interpretation of the results.

Test Phase

A common test, provided on paper, with similar questions to the problems discussed in the lessons and the problems provided to the worked-examples group, was conducted during the third week. The test was conducted in three distinct stages. Each stage was 20 min long. Separate question booklets were provided to test the preliminary database work, calculation, and the summary/totals work. Stage 1 of the test covered the material taught and practiced during the first week, familiarizing the students with the content and structure of the databases on the disks. It tested the common first lesson and practice content.

The following is a sample question:

> Earth Action Co. want you to analyse United Nations environmental pollution figures relating to cigarette consumption in the world.
>
> Develop a database to accept the information about cigarette consumption and save it on your floppy disk. You need to record the region of the world (mostly the name of the continent), country's name, the number of manufactured cigarettes smoked during the year per each adult for the years 1965, 1970, 1975, 1980, 1985, and the adult smoking prevalence for males and females, ie. what percentage of the adults in the population smoke.
>
> Enter the following data into the database:
>
> For Morocco, the region is Africa. The number of cigarettes consumed per person: In 1965 it was 570, in 1970 it was 620, in 1975 it was 690, in 1980 it was 994 and in 1985 it was 1070. Male smoking prevalence was 75% and female smoking prevalence was 0%.
>
> For Denmark, the region is Europe. The number of cigarettes consumed: In 1965 it was 1500, in 1970 it was 1690, in 1975 it was 2210, in 1980 it was 1970 and in 1985 it was 2110. Male smoking prevalence was 49% and female smoking prevalence was 38%.
>
> Close the database.

To answer this question, students had to create and name a new database file, generate and name two text database fields and seven number fields, and enter the data into each field for two database records.

When the time had expired, Stage 1 materials were removed, and the students were asked to work on Stage 2. This test stage was composed of five questions requiring the construction of calculation fields.

The following is an example:

> Open a comprehensive database called CIGARETTE2 from your floppy disk, which has the same structure and all of the information you entered on the Cigarette Consumption database in Question 3, as well as information from many other countries around the world.
>
> Make new computation fields for the database to show:
>
> a. actual change in cigarettes consumed per person between 1965 and 1985 (What is the numerical change in cigarette consumption for Denmark from 1965 to 1985: ____)
>
> b. the percentage change in cigarettes consumed per person between 1965 and 1985. (What is the percentage change in cigarette consumption for Ghana from 1965 to 1985:____)

Part a of this question had to be answered by naming and creating a computation field that computed the difference between the 1985 and the 1965 cigarette consumptions. This procedure involved selecting the 1985 cigarette-consumption field, the subtraction operator, and the 1965 cigarette-consumption field in the calculation-field construction window, that is, three distinct steps in particular order giving the formula 'Cig cons 85' − 'Cig cons 65'. After the new field to calculate the difference was constructed, the student could simply read the value for the specified country, Denmark, from this field to answer the last part of a.

Part b of this question also required the construction and naming of a difference field as above but which needed to be converted to a percentage by dividing by the 1965 cigarette consumption and multiplying by 100, that is, more steps were needed than in part a above. The formula to be created was ('Cig cons 85' − 'Cig cons 65') × 100/'Cig cons 65', where the items in single quotes were available as field names to be accessed during the formula construction. The result for the last part of the question could again be read from the newly constructed field for the desired country, Ghana.

After the time had expired for Stage 2, the materials were removed, and the students were supplied with Stage 3 of the test. This test stage consisted of six questions on summary/totals fields. The students were able to use the flowchart summaries of steps for constructing the two kinds of fields during the test, as used during practice sessions by everyone, and they were provided prepared databases on disks to modify according to the test instructions.

The following is an example:

> Use the database called CIGARETTE2 from your floppy disk, already used in Question 6, to develop the following summary fields and place them in a Totals Part at the end of the columns of figures:
>
> a. average male prevalence of smoking in all the countries recorded on this database (What is the average male prevalence in these 47 countries: ____)
>
> b. maximum number of cigarettes smoked per person in 1985. (What is the maximum number of cigarettes smoked per person: ____)

An appropriate answer to this question required the student to create and name a summary/totals field, where they had to choose one option from each of two menus presented in the summary/totals construction window, the relevant database field on which the computation was based, and the type of computation, which in part a was average and in part b was maximum. Then the results requested for the second part in a and b above could be read from the newly created fields in the database.

Throughout the lessons, practice sessions, and tests, the students were asked to record the mental effort required to complete the tasks using a 9-point Likert scale (Paas & Van Merriënboer, 1993).

Results

The variables under analysis were total test scores, mental effort ratings, and the relative efficiencies of the different learning conditions. The individually correct question parts were added together for each question to derive the question score, and the question scores were added together to obtain the total test scores.

In the initial survey of the students' computing skills at the beginning of the subject, they were asked to indicate how often they had used databases, using options *never, seldom, sometimes, often,* and *very often,* and to indicate the specific database software they had used. Overall, of the 32 students, 4 indicated they had used databases sometimes, and 11 seldom. The 4 sometimes and 4 of the seldom choices came from the worked-examples group, and 7 seldom replies were from the exploration group. None of the students indicated they had used databases often or very often. Four students in the worked-examples group and 3 of the exploration group were able to name the databases they had used. Thus, it appears the database levels of familiarity were very similar for the two groups.

The students in the seldom and sometimes categories were grouped together, allowing us to compare total test score means by using a 2 (higher or lower levels of experience) × 2 (worked-example or exploration group) analysis. A 2 (groups) × 2 (levels of experience) analysis of variance (ANOVA) on the total test scores gave a nonsignificant main effect with respect to groups, $F(1, 28) = 1.00$, $MSE = 180$. (The .05 level is used throughout this article.) The main effect with respect to the database level of experience was significant, $F(1, 28) = 5.32$, $MSE = 180$. A statistically significant interaction between these variables also was indicated, $F(1, 28) = 4.18$, $MSE = 180$. As can be seen from Figure 1, where the test means are plotted, the worked-example group obtained a much higher score than the exploration group when students had no prior exposure to databases, but a slightly lower score when both groups had some degree of experience.

When we compared the test results of the two groups of students without previous database exposure by using a simple effects analysis, we found a significant effect, $t(15) = 2.30$, in favor of the worked-examples group. The comparison of the two groups with previous database exposure was in the reverse direction but indicated no significant difference, $t(13) = 0.69$, $p = .50$.

We obtained the overall practice mental effort ratings by computing the means of the students' recorded responses on the 9-point mental effort Likert scales during either worked examples or exploration practice. (See Figure 2. Higher values reflect a greater mental effort.) We compared the practice mean mental effort by using a 2 × 2 factorial ANOVA, resulting in a significant main effect on the experimental group factor, in favor of the worked-examples group, $F(1, 28) = 8.19$, $MSE = 1.28$. The main effect with

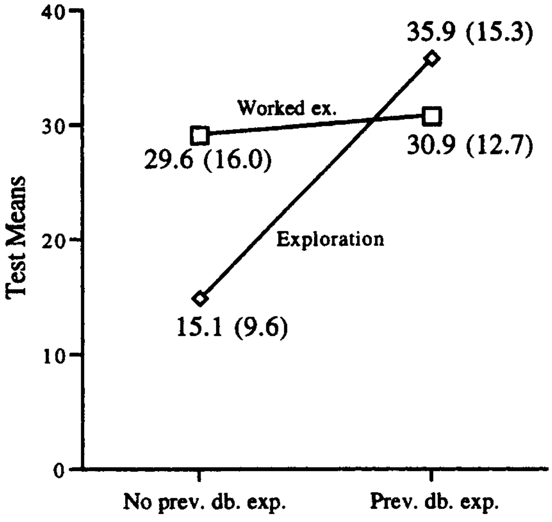

Figure 1. Test means of performance scores. Standard deviations are shown in parentheses. No prev. db. exp. = no previous database exposure: Worked ex. = worked-example group, $N = 8$; exploration group, $N = 9$. Prev. db. exp. = previous database exposure: Worked ex. group, $N = 8$; exploration group, $N = 7$.

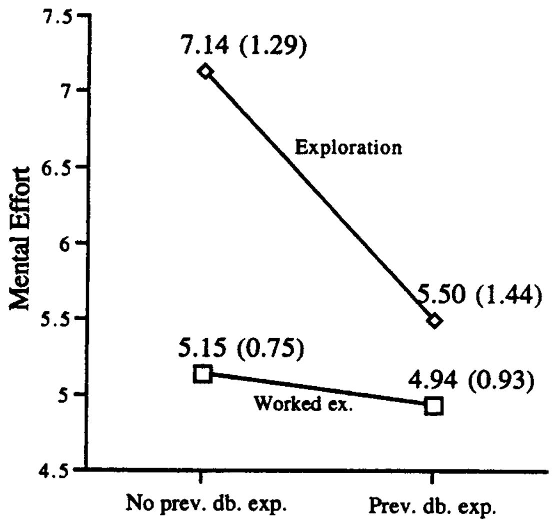

Figure 2. Practice mental effort means for experimental groups. Standard deviations are shown in parentheses. Higher mental effort score indicates greater effort. No prev. db. exp. = no previous database exposure. Prev. db. exp. = previous database exposure. Worked ex. = worked-example group.

respect to the database level of experience also was significant, in favor of previous database experience, $F(1, 28) = 5.28$, $MSE = 1.28$. The interaction only was significant at the .1 level and so may reflect a real effect, $F(1, 28) = 3.18$, $MSE = 1.28$. Students with the least previous database exposure in the exploration practice recorded the highest mental effort ratings, as shown in Figure 2.

Although the interaction was not significant, it might be noted that when we compared the practice mental effort of the two groups of students without previous database exposure by using a simple effects analysis, we found a highly significant result, $t(15) = 3.81$, in favor of the worked-examples group. The comparison of the two groups with previous database exposure indicated no significant difference between the groups, $t(13) = 0.91$. These results suggest that the cognitive load associated with exploration was considerably higher than for worked examples, but only for students who had less experience in dealing with the curriculum area.

Paas and Van Merriënboer (1993, 1994) developed a statistical technique for assessing the effectiveness or efficiency of an instructional technique. The procedure is based on a joint analysis of performance and mental effort ratings. It assumes that the most effective instructional procedures decrease cognitive load (measured by mental ratings) but increase performance scores. The method combines test performance and perceived mental task effort measures by using standardized test scores and mental effort ratings. Paas and Van Merriënboer called their measure *relative condition efficiency.* The formula for efficiency is

$$\text{Efficiency} = \frac{\text{Mean } z \text{ score of mental effort rating} - \text{Mean } z \text{ score of performance scores}}{\sqrt{2}}.$$

We computed the test relative efficiencies for each student and compared them with regard to previous database experience and type of practice. A 2 × 2 factorial ANOVA indicated significant main effects with respect to groups, in favor of the worked-examples group, $F(1, 28) = 7.20$, $MSE = 0.82$, and with respect to the level of experience in favor of previous database experience, $F(1, 28) = 9.25$, $MSE = 0.82$. The interaction was also significant, $F(1, 28) = 6.43$, $MSE = 0.82$.

Figure 3 illustrates the difference between the efficiencies. The line E = 0 indicates a neutral efficiency condition. The values above and to the left of the line are more efficient, and the values below and to the right of this line are less efficient. The perpendicular distance of a point from the E = 0 line is a measure of the size of the practice format efficiency.

When we compared the relative efficiency of the two groups of students without previous database exposure by using a simple effects analysis, we found a significant difference, $t(15) = 3.59$, in favor of the worked-examples group. The comparison of the two groups with previous database exposure indicated no significant difference between the groups, $t(13) = 0.11$, $p = .92$. Worked examples were a much more efficient instructional technique than exploration for less experienced students, but this advantage disappeared when we used more experienced students.

All of these measures indicate that worked-examples practice benefited the students without previous database experience when compared with exploration practice. The experimental measures were often in a ratio of 2:1 or better, in favor of worked-examples practice.

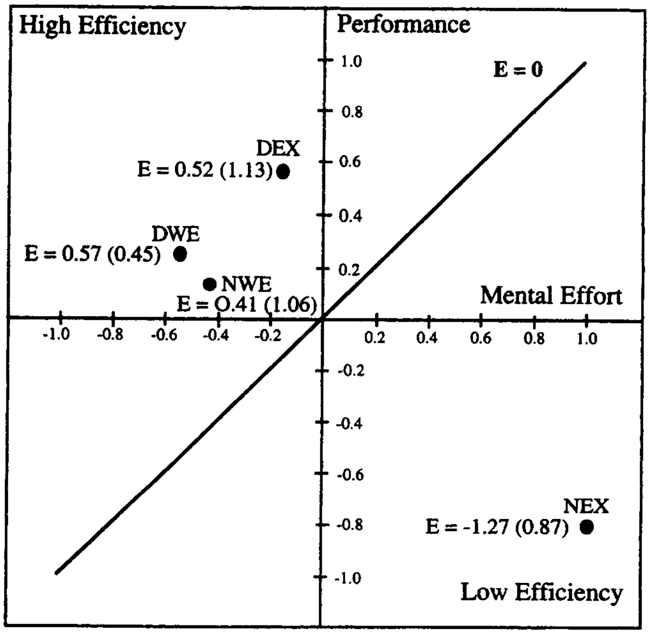

Figure 3. Practice relative efficiency means. Standard deviations are shown in parentheses. NWE = no database experience, worked-examples practice. DWE = database experience, worked-examples practice. NEX = no database experience, exploration practice. DEX = database experience, exploration practice.

Discussion

The results of the experiment support the suggestion that in comparison to exploration, presenting students with worked examples assists them in learning to use a database program. The advantage occurred only for students with no prior experience with database programs. For students with some experience in the area, the advantage of worked examples was eliminated. No differences between examples and exploration were found for students with previous experience.

Mental effort ratings suggested that the advantage of worked examples was due to differential cognitive load. Inexperienced students in the exploration condition reported high mental effort levels compared with students in the worked-example condition. If exploration generated high mental effort levels that assisted learning, it would of course be advantageous. Evidence that the heavy cognitive load associated with exploration was extraneous to learning relevant aspects of the database came from the efficiency measures. Not only did exploration result in a heavy cognitive load, but that load interfered with rather than facilitated learning. There was a very large difference in efficiency favoring the worked-examples group for students with no experience using the database.

The effectiveness of worked examples clearly depends on the previous domain knowledge of the students. If they have sufficient domain knowledge, the format of practice is irrelevant, and discovery or exploration practice is at least as good, or may even be better, than worked-examples practice. However, if the students' previous domain knowledge is restricted, then worked-examples practice can be more beneficial than exploration practice. The reasons may be associated with schema acquisition. If a student has no schemas or very limited schemas associated with an area of study, it may be difficult for the student to generate suitable aspects of the area to explore. Deciding what to explore may be cognitively demanding (hence the high mental effort rating scores by our inexperienced students), and even when decisions are made, the student may have insufficient knowledge to make choices that are important in providing experience of substantial aspects of the area. The acquisition of additional schemas through experience may provide exploration decisions with firmer foundations. It is possible that with sufficient experience in an area, exploration may be superior to worked examples, with worked examples only being beneficial for novices who first encounter new material.

These results both contradict and refine those obtained by Charney et al. (1990), who interpreted their results to indicate problem solving led to superior learning outcomes compared with tutorial and exploration approaches. They specifically excluded from their experiment previous domain knowledge, which appeared to be a very powerful factor in the Carroll, Mack, Lewis, Grischowsky, and Robertson (1985) study. Where a good knowledge level existed, Carroll et al. (1985) found that exploration practice was at least as effective, if not more effective, than other practice formats. We have partly replicated this finding, having found worked-examples and exploration approaches provide about equal benefit for students with some domain experience. However, when the preexisting knowledge level was limited, exploration practice clearly caused a much larger cognitive load and led to poorer learning than worked-examples practice. In our experiment, worked-examples practice could be regarded as an integrated tutorial/problem-solving practice approach using Charney et al.'s terms. Thus, in this experiment, we found that combining worked examples and problem solving produced better learning for students totally unfamiliar with a new domain, but exploration practice was just as good as this combined approach for students with some domain experience. We obtained evidence that these results were due to cognitive load factors rather than other factors by recording mental effort ratings.

It can, of course, be argued that exploration practice may be superior to worked examples, even for novices, if measures other than those of the present experiment are used. For example, exploration may favor long-term retention. Although this question must remain open until tested, it should be noted that in the present case, students with no previous database experience who learned by exploration, achieved such low test scores that minimal knowledge was available for long-term retention. Long-term retention requires that a reasonable amount be learned in the first instance.

In conclusion, on the evidence of the current experiment, providing students with more rather than less structure is beneficial. This conclusion applies only to students with very little knowledge of a subject area. With more knowledge, the advantage of additional structure may disappear.

References

Anthony, W. S. (1977). Learning to discover rules by discovery. In M. C. Wittrock (Ed.), *Learning and instruction* (pp. 491–495). Berkeley, CA: McCutchan.

Ausubel, D. P. (1961). Learning by discovery: Rationale and mystique. *Bulletin of the National Association of Secondary School Principals, 45,* 18–58.

Bell, B. (1991). Implications for curriculum. In J. Northfield & D. Symington (Eds.), *Learning in science viewed as personal construction. An Australian perspective* (pp. 34–51). Perth, WA: Key Centre for School Science and Mathematics, Curtin University of Technology.

Blanchette, M., & Brouard, F. (1995). Learning by discovering: Using a case in an introduction to accounting course. In H. E. Klein (Ed.), *The art of interactive teaching with cases, simulations, games and other interactive methods* (pp. 121–130). Boston: WACRA.

Bruner, J. S. (1959). Learning and thinking. *Harvard Educational Review, 29,* 184–192.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31,* 21–32.

Bruner, J. S. (1962). *On knowing. Essays for the left hand.* Cambridge, MA: Harvard University Press.

Bruner, J. S. (1966). Some elements of discovery. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal* (pp. 101–113). Chicago: Rand McNally.

Carlson, R. A., Lundy, D. H., & Schneider, W. (1992). Strategy guidance and memory aiding in learning a problem-solving skill. *Human Factors, 34*(2), 129–145.

Carroll, J., Mack, R., Lewis, C., Grischowsky, N., & Robertson, S.

(1985). Exploring a word processor. *Human-Computer Interaction, 1,* 283–307.

Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 86,* 360–367.

Charney, D., Reder, L., & Kusbit, G. W. (1990). Goal setting and procedure selection in acquiring computer skills: A comparison of tutorials, problem solving, and learner exploration. *Cognition and Instruction, 7*(4), 323–342.

Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Hillsdale, NJ: Erlbaum.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*(4), 347–362.

Crawford, K. (1995). Designing, exploring and interacting: Novice activities in the Boxer computational medium. In A. A. diSessa, C. Hoyles, R. Noss, & L. Edwards (Eds.), *Computers and exploratory learning* (pp. 443–462). Berlin, Germany: Springer-Verlag.

Cronbach, L. J. (1966). The logic of experiments on discovery. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal* (pp. 76–92). Chicago: Rand McNally.

diSessa, A. A., Hoyles, C., Noss, R., & Edwards, L. (1995). Computers and exploratory learning: Setting the scene. In A. A. diSessa, C. Hoyles, R. Noss, & L. Edwards (Eds.), *Computers and exploratory learning* (pp. 1–12). Berlin, Germany: Springer-Verlag.

Gagné, R. M. (1966). Varieties of learning and the concept of discovery. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal* (pp. 135–150). Chicago: Rand McNally.

Gagné, R. M., & Brown, L. T. (1961). Some factors in the programming of conceptual learning. *Journal of Experimental Psychology, 62,* 313–321.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15–46). New York: Macmillan.

Holzman, T. G., Pellagrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. *Journal of Educational Psychology, 75*(4), 603–618.

Hoyles, C. (1995). Thematic chapter: Exploratory software, exploratory cultures? In A. A. diSessa, C. Hoyles, R. Noss, & L. Edwards (Eds.), *Computers and exploratory learning* (pp. 199–219). Berlin, Germany: Springer-Verlag.

Hsu, J.-F. J., Chapelle, C. A., & Thompson, A. D. (1993). Exploratory learning environments: What are they and do students explore? *Journal of Educational Computing Research, 9*(1), 1–15.

Kamouri, A., Kamouri, J., & Smith, K. (1986). Training by exploration: Facilitating the transfer of procedural knowledge through analogical reasoning. *International Journal of Man-Machine Studies, 24,* 171–192.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from the Tower of Hanoi. *Cognitive Psychology, 17,* 248–294.

Maor, D., & Taylor, P. C. (1995). Teacher epistemology and scientific inquiry in computerized classroom environments. *Journal of Research in Science Teaching, 32*(8), 839–854.

Mayes, R. L. (1992). The effects of using software tools on mathematical problem solving in secondary schools. *School Science and Mathematics, 92*(5), 243–248.

Miller, G. A. (1956). The magical number, seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Njoo, M., & de Jong, T. (1993). Exploratory learning with computer simulation for control theory: Learning processes and instructional support. *Journal of Research in Science Teaching, 30*(8), 821–844.

Nuthall, G., & Snook, I. (1973). Contemporary models of teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp. 47–76). Chicago: Rand McNally.

Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434.

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737–743.

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122–133.

Rieber, L. P., & Parmley, M. W. (1995). To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. *Journal of Educational Computing Research, 13*(4), 359–374.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66.

Sfondilias, J. S., & Siegel, M. A. (1990). Combining discovery and direct instruction strategies in computer-based teaching of mathematical problem solving. *Journal of Computer-Based Instruction, 17*(4), 130–134.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127–190.

Shulman, L. S., & Tamir, P. (1973). Research on teaching in the natural sciences. In R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp. 1098–1148). Chicago: Rand McNally.

Simon, H., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5,* 29–46.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59–89.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296.

Trafton, J. G., & Reiser, B. J. (1993). The contributions of studying examples and solving problems to skill acquisition. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1017–1022). Boulder, CO: Erlbaum.

Vitale, B. (1995). From local to global: Programming and the unfolding of local models in the exploratory learning of mathematics and science. In A. A. diSessa, C. Hoyles, R. Noss, & L. Edwards (Eds.), *Computers and exploratory learning* (pp. 45–58). Berlin, Germany: Springer-Verlag.

Whitehead, G. (1978). *Enquiry learning in social studies.* Hawthorn, Victoria: Australian Council for Educational Research.

Wittrock, M. (1966). The learning-by-discovery hypothesis. In L. Shulman & E. Keisler (Eds.), *Learning by discovery: A critical appraisal* (pp. 33–75). Chicago: Rand McNally.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4,* 137–166.

Received February 11, 1998
Revision received May 10, 1998
Accepted July 9, 1998 ■