COGNITION AND INSTRUCTION, 1990, 7(1), 1-39


# Structuring Effective Worked Examples

Mark Ward and John Sweller
*University of New South Wales*

Under some conditions, substituting worked examples for problems or exercises enhances learning and subsequent problem solving. Under other conditions, worked examples are no more effective, and possibly less effective, than solving problems. Using cognitive load theory, we hypothesize that the critical factors for enhanced learning are whether the worked examples can direct attention appropriately and reduce cognitive load. It is suggested that worked examples requiring students to mentally integrate multiple sources of information are not effective because they fail with respect to both of these factors. The results of five experiments using geometric optics and kinematics under classroom conditions provided evidence for these hypotheses. Worked examples, formatted to reduce the need for students to mentally integrate multiple sources of information, resulted in test performance superior to either conventional problems or to worked examples requiring students to split their attention between, for example, text and equations or text and diagrams. We conclude that because traditional worked example formats frequently are random with respect to cognitive factors, they may be ineffective in some areas and require restructuring.

Considerable evidence has accumulated recently, showing that learning and problem solving can be facilitated more by students studying many worked examples rather than solving many problems. Using algebra transformation problems, Cooper and Sweller (1987) and Sweller and Cooper (1985) demonstrated enhanced performance on subsequent problems after students studied worked examples. Solving the same problems rather than studying them as worked examples reduced the facilitative effects. Zhu and H. A. Simon (1987) provided evidence of substantial performance gains when worked examples were used as a substitute for lectures and other conventional classroom activities. In one long-term study, they found that a conventional 3-year mathematics course was completed in 2 years with a slightly higher level of performance than conventionally taught students.

Requests for reprints should be sent to John Sweller, School of Education, University of New South Wales, P.O. Box 1, Kensington, New South Wales 2033, Australia.

A detailed theoretical structure is available to explain these results and suggests the conditions under which they can be obtained. Schema acquisition and rule automation are assumed to be basic components of skilled problem solving performance. A *schema* is defined as a cognitive construct that permits problem solvers to categorize problems according to solution moves (see Chi, Glaser, & Rees, 1982; Schoenfeld & Herrmann, 1982). Possession of appropriate schemas permits expert problem solvers to recognize problems and problem states and to use a schema to generate moves. Schemas facilitate performance on problems structurally similar to ones seen previously.

Rule automation allows rules (e.g., rules of mathematics or science) to be employed correctly with little or no conscious control (see Kotovsky, Hayes, & H. A. Simon, 1985). Automated rules free problem solvers of the need to consider consciously the validity of rule use, and this in turn allows cognitive resources to be employed fully in searching the problem space. Consider, for example, a student who must carefully evaluate an equation such as $s = vt$ before use, to ensure that it is being used correctly. That student is likely to have fewer cognitive resources available for problem search than a student who does not have to think about the equation because it is fully automated. By *automated*, we mean that the rule can be retrieved and applied without first considering, for example, whether $v$ stands for average or final velocity.

Rule automation should facilitate performance on all problems, but especially on transfer problems that are sufficiently different from previously encountered examples to reduce the utility of schemas. A transfer problem that requires use of the same rules as previously solved problems, but in a different context, should be solved more readily if a problem-solving search can proceed while using automated rules.

Cooper and Sweller (1987) made several suggestions concerning the interaction of schema acquisition and rule automation. First, schemas are more likely to facilitate solutions to problems reasonably similar to previously seen problems encompassed by the orbit of the relevant schema. Second, rule automation should facilitate problem solution on all problems requiring the use of that rule. Third, schema acquisition can occur relatively quickly, whereas rule automation is a slow process requiring considerable practice. Fourth, and as a consequence of the first three points, a procedure such as worked examples that can facilitate learning will first improve performance on problems similar to those seen previously, due to schema acquisition. Only later will transfer improve due to rule automation.

Experts in a given area can use schemas to generate moves, but novices who do not possess schemas must use alternative move-generating strategies, with the most common strategy being means-ends analysis (Larkin, McDermott, D. Simon, & H. A. Simon, 1980). A problem solver using this strategy attempts to reduce differences between each problem state encountered and the goal state using allowable problem solving operators. In science and mathematics, the operators are the rules of mathematics and science. In the case of mathematics and science students, these rules are normally not automated.

If novices tend to solve problems by means-ends analysis using nonautomated rules, and if experts solve problems using schemas and automated rules, the consequence of a means-ends strategy for schema acquisition and rule automation becomes an important consideration. Substantial evidence now is available, using a wide variety of problems, suggesting that a means-ends strategy interferes with learning (see Owen & Sweller, 1985; Sweller & Levine, 1982; Sweller, Mawer, & Howe, 1982; Sweller, Mawer, & Ward, 1983). Sweller (1988) theorized that means-ends analysis interferes with learning because, although it is an efficient strategy for achieving a problem goal, with respect to schema acquisition, it inappropriately focuses attention and imposes a heavy cognitive load. Attention is directed to differences between problem states rather than to each state and its associated moves. A heavy cognitive load is imposed because of the need to simultaneously consider and make decisions about the current problem state, the goal state, differences between states, and problem solving operators that can be used to reduce such differences. When nonautomated operators are being used, the process becomes even more difficult.

If a student, when learning the structure of problems in a new area, tends to use a strategy that facilitates problem solution but interferes with schema acquisition and rule automation, then more appropriate alternatives to conventional problem solving may be more effective. These alternatives should focus attention on problem states and reduce cognitive load. Worked examples can focus attention on problem states and their associated moves. They can also reduce cognitive load. Consequently, they should facilitate learning and subsequent problem solving to a greater extent than actually engaging oneself in the solution process. As previously indicated, Cooper and Sweller (1987) and Sweller and Cooper (1985) obtained this result using algebra transformation problems.

The current series of experiments was designed with two issues in mind. First, although there is evidence that a heavier than normal emphasis on worked examples can have beneficial effects compared with a similar emphasis on problems, much of this evidence has been obtained under laboratory conditions using a restricted range of materials—algebra problems. There is a clear need to widen the range of problems and to demonstrate the effect under conventional classroom conditions. Zhu and H. A. Simon's (1987) finding that mathematics worked examples can act as a very effective substitute for general classroom teaching is encouraging in this respect. The experiments of the present article used a variety of physics problems with comparisons between worked examples and conventional problems carried out as part of routine high school programs.

The second and more important issue concerns the format of the worked examples used. Sweller (1988) suggested that the manner in which cognitive resources must be distributed while engaged in a task (cognitive load theory) is critical to learning and problem solving. According to this theory, alternatives to conventional problems can be effective if they appropriately direct attention and reduce cognitive load. Although conventional algebra worked examples have this effect (Cooper & Sweller, 1987; Sweller & Cooper, 1985), it is not a necessary consequence of worked examples. According to cognitive load theory, the critical factors are not whether worked examples are used, but rather whether the techniques used appropriately direct attention and impose a relatively light cognitive load.

Tarmizi and Sweller (1988), using circle geometry problems, provided evidence that, for geometry worked examples to be effective, their format had to be substantially altered in order to reduce cognitive load. Worked examples that required students to split their attention between multiple sources of information and mentally integrate those multiple sources were ineffective. Such split-attention worked examples are conventional in geometry, in which students must consider both a diagram and a set of statements. They occur also in many areas of physics. In these areas we might expect conventional worked examples to be ineffective unless appropriately modified. In the current series of experiments, the first two experiments presented unified geometric optics worked examples; the third used a variety of dynamics worked examples that required students to split their attention; the fourth used dynamics worked examples that, in one group, eliminated the need to split attention; and the fifth tested the consequences of requiring students to split their attention on optics worked examples.

## EXPERIMENT 1

This experiment compared worked examples with conventional problems using lens and mirror ray diagram problems. Material was presented in class under conventional conditions, followed by homework consisting of either a mixture of worked examples and problems or identical material presented as conventional problems alone. On the next day, the homework was followed by a conventional classroom test. The number of problems solved correctly served as the primary dependent variable.

Table 1 summarizes the experimental design. Two classes taught by the same teacher were used. For both classes, the classroom presentation, homework, and test associated with mirrors were completed before commencing work on lenses. For one class, the initial work on mirrors was associated with worked examples as homework. The subsequent material on lenses was associated with conventional problems for homework. The other class was given conventional problems on mirrors, followed by worked examples on lenses. This counterbalancing was used as an alternative to splitting each class into worked example and conventional problem groups. Splitting classes into groups was administratively difficult in the school used. Counterbalancing was used to reduce the likelihood that any differences between worked examples and conventional problems could be directly attributed to differences in ability between classes or differences in difficulty between the mirror and lens material.

## Method

*Subjects.* The subjects were 42 Year 10 students from the second-level (Class A, composed of 21 students) and fourth-level (Class B, composed of 21 students) science classes of a Sydney high school with eight Year 10 science classes. Students had been placed in these classes (or levels) according to their examination performances in the preceding year. Students in the second-level class were ranked 30–50 in the year and in the fourth-level class were ranked 130–150 in the year. These classes were chosen to avoid possible asymptotic effects. At this school, higher ability classes tended to have many students able to solve the problems very rapidly, whereas lower classes tended to have many students unable to solve the problems.

*Materials.* The problems used in this experiment came from two topics. Topic 1 required students to locate the position of an image formed in a mirror (concave or convex) by conventional ray diagrams. In these diagrams, three incident rays were drawn as given from the object to the mirror, and the path of the reflected rays then had to be drawn using learned rules (the rules are listed in Figure 1). The intersection or apparent intersection of these reflected rays provided the location of the image formed in the mirror. Figure 1 shows an example of the mirror problems used in the experiment.

Topic 2 required students to locate the position of an image formed in a lens (convex or concave) by conventional ray diagrams. As was the case for mirrors, in these diagrams, three incident rays were drawn from the object to the lens, and the path of the refracted rays then had to be drawn using learned rules.

*Procedure.* The experiment was conducted as part of the normal teaching program of the school, and the test was considered a school test by the students. This experiment and all subsequent experiments were conducted by the usual teacher of the students concerned. As indicated in the Materials section, the physics problems used in this experiment were geometric optics problems involving the use of rules (or ray constructions) to

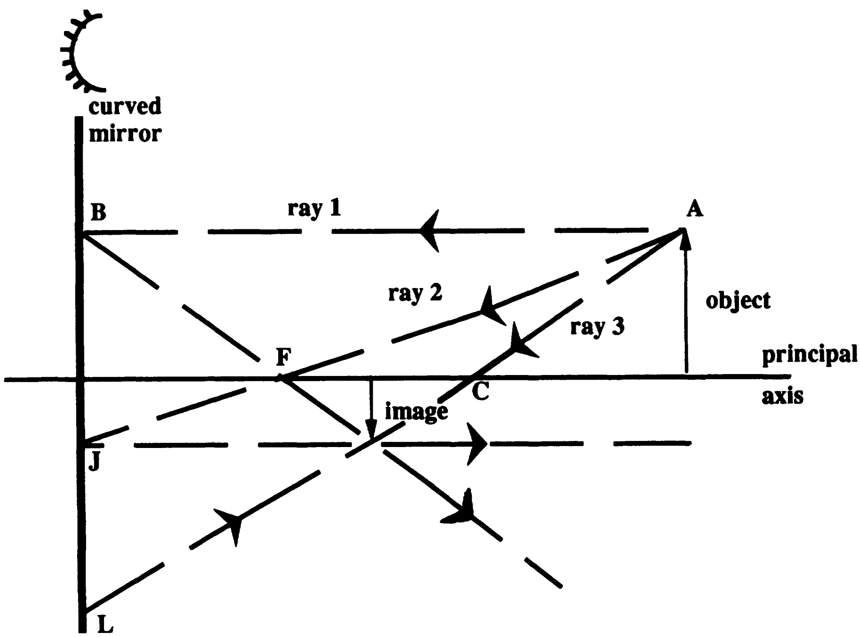


FIGURE 1 An example of the mirror problems and solutions used in Experiments 1, 2, and 5. Assume that the problem goal required students to find the image formed by the mirror. For Experiments 1, 2, and 5 in the homework phase, the worked-examples group contained only the three rays and the letters *C* and *F*. The conventional-problem group contained only the object's position and the letters *C* and *F*. The rays were absent. For Experiment 5, the split-attention group contained all of the information in the diagram plus the following rules needed for solution: (a) ray AB, drawn parallel to the principal axis, passes through the principal focus, F, after being reflected; (b) ray AJ, drawn through the principal focus, F, is parallel to the principal axis after being reflected; (c) ray AL, drawn through the center of curvature C, is incident normally at L and retraces its path after being reflected; (d) the image is located at the intersection of the rays.

locate geometrically the image formed by the optical components (i.e., mirrors and lenses). These two topics—mirrors and lenses—provided a very small subset from the domain of geometric optics but were typical of what is taught about this topic in a junior high school science course.

Before the study, the only relevant knowledge that both classes had was work done just prior to the study on the formation of an image in a plane mirror by ray constructions. This material gave students the necessary background knowledge to complete the problem sets used, but it allowed for the potential observation of substantial improvement during the study. The background knowledge included information on (a) how to represent the path of light by use of a mathematical construct (i.e., the ray) and (b) how to describe the image formed by defining such terms as real/virtual, upright/inverted, and magnified/same size/diminished.

During the experiment, three phases were administered (see Table 1). These phases conformed to a conventional teaching sequence. The first phase, the teaching phase, was presented as a normal classroom lesson to the class as a whole with no distinctions between students to be presented with either worked examples or conventional problems. The teacher began with an explanation of the three rules that were necessary to locate the image formed by an optical component. These explanations were presented on an overhead transparency, and students were given a copy of the transparency in the form of an information sheet. The rules were presented in words, and each rule was reinforced by a ray diagram (Figure 1 shows the four rules for locating a mirror image). Then students were shown a worked example and were given two problems to solve covering rule use with concave and convex mirrors or concave and convex lenses (see Figure 1 for a concave mirror example). Students were asked to solve each problem by using the three rules that had been explained previously. Further verbal instructions were given in class to those students who were unable to solve the problems. The solutions to the problems were then presented to the class on overhead transparencies, allowing students to check their answers. The information sheet was available at all times. Students were asked if they had any questions concerning these solutions, and all questions were answered.

The teaching phase was followed by an "acquisition" or homework phase during which each group was presented with a series of homework problems. In this phase (and only in this phase), the critical distinction between worked examples and conventional problems occurred, with some of the problems presented as worked examples. The 10 homework-acquisition problems were presented as five pairs, with two identical format problems in each pair. When the experimental design required worked examples, the first member of each pair was presented as a worked example, and the second was a conventional problem. The homework problems had to be solved (or, in the case of worked examples, solutions were presented) using the pre-

TABLE 1
Experimental Design Used in Experiments 1 and 2

| | *Teaching Phase* | *Homework Phase* | | *Test Phase* |
|---|---|---|---|---|
| | | *Mirrors* | *Lenses* | |
| Class A | Mirrors followed by lenses and identical for both classes | Worked examples | Conventional problems | Mirrors followed by lenses and identical for both classes |
| Class B | | Conventional problems | Worked examples | |

viously taught rules. Students were informed that the homework would be collected and marked the following day. They were further instructed that they would be given a test on the homework immediately following the collection of these problems.

Four of the five homework pairs dealt with one optical component (in Topic 1, concave mirrors; in Topic 2, convex lenses) and were similar to the first of the problems presented on the problem sheet used in the teaching phase (see Figure 1). Each pair of problems within a topic differed from the other pairs in terms of the possible positions of the image in front of the optical component. Within each pair, the problems contained objects at the same approximate distance from the optical component or at the same point but with an altered orientation. The remaining fifth pair consisted of two problems on either convex mirrors for Topic 1 or concave lenses for Topic 2. Both pairs were similar to the second problem on the example sheet used in the teaching phase and will be referred to as low practice problems. All problems had to be solved by finding the image formed using the three rules. Throughout the homework phase, students had free access to materials such as the information sheet containing the rules and the applicable ray diagrams.

The test phase occurred on the following day. As was the case for the initial teaching phase, distinctions between worked examples and conventional problems did not occur in this phase. For the mirror test, nine conventional problems were presented on paper. Test Problems, 1, 2, 5, 6, and 8 were identical in structure to the homework problems requiring the use of the three rules to find the image formed in a concave mirror. Test Problems 4 and 9 were identical in structure to the homework problems requiring the use of the rules to find the image formed in a convex mirror (low practice problems). Test Problems 3 and 7 were different in structure from the homework set and from each other. In the first of these problems, students were required to construct a ray diagram from a description of requirements: "Draw a ray diagram in the space provided to show how a virtual image can be produced by a concave lens." In the second problem, an image was given, and students were asked to find the correct positioning of an object that would give this image by constructing a ray diagram. These two problems were used to test for transfer.

For the lens test, eight conventional problems were presented on paper. Test Problems 1, 4, 6, 7, and 8 were similar in structure to the homework problems requiring the use of the three rules for finding the image formed in a convex lens. Test Problem 3 was similar to the concave lens problems used in the homework phase (low practice problems), whereas Problems 2 and 5 (transfer problems) were slightly different in structure from the homework set and from each other. Similar to the mirror topic, these two lens problems required a ray diagram to be constructed from a description

TABLE 2
Mean Percentage (and Standard Deviations) of Problems Correctly Solved During the Test Phase of Experiment 1

| | *Topic* | |
|---|---|---|
| | *Mirrors* | *Lenses* |
| Class A | | |
| Including transfer problems | 53.3 (16.9)[a] | 71.4 (16.4)[b] |
| Excluding transfer problems | 61.2 (17.7)[a] | 80.2 (15.1)[b] |
| Class B | | |
| Including transfer problems | 67.9 (14.3)[b] | 51.5 (16.5)[a] |
| Excluding transfer problems | 78.9 (12.1)[b] | 56.8 (19.9)[a] |

[a]Conventional problems used in the homework phase. [b]Worked-example problems used in the homework phase.

and, in the second problem, the object position to be found given the image position.

During the test phase, students did not have access to previous work conducted in the teaching and homework phases. Fifteen minutes were allowed for each test. This procedure was used on both mirror and lens topics.

The distinction between worked examples and conventional problems occurred in the homework-acquisition phase. The conventional problem group was simply required to solve the 10 problems using pencil and paper. The worked-example group was given the same problems, except that the first problem of each pair of identical format problems had the solution written out (see Figure 1). Students were informed that they should study each worked example carefully until they were sure they understood it, because the following problem would be similar.

*Experimental design.* Both classes were taught the two topics—mirrors and lenses—using an alternation procedure in the homework phase as shown in Table 1. In Class A, mirror problems in the homework phase were given as worked example types, and the second topic, lenses, was presented as conventional problems. The presentation method was reversed for Class B, with the mirror problems presented as conventional problems and the following lens problems presented as worked examples. Such counterbalancing ensured that any differences between worked examples and conventional problems were not likely to be contaminated by differences between the two classes or the two topics.

## Results and Discussion

This experiment tests the hypothesis that knowledge of a problem domain will be relatively enhanced by use of worked examples instead of conventional problems during the homework phase. This difference should be revealed by the performance of each group on both topics during the test phase. Table 2 shows the mean percentage of problems correctly solved by each group on both mirror and lens tasks during the test phase. A problem was scored as correctly solved if the rules were applied correctly. The path of all rays drawn in the answer had to accurately follow the rules displayed in the teaching phase. Showing rays without arrows or, alternatively, showing the correct intersection of rays without then drawing the image were not counted as errors. Drawing an incorrect pathway for a ray, treating a concave mirror or lens as a convex mirror or lens, or vice versa, are examples of errors.

An analysis of variance (ANOVA) on these data indicated that there was no significant difference between classes, $F(1, 36) = 0.3$, $MSE = 139.4$ (the .05 level of significance is used throughout this article), or between topics, $F(1, 36) = 0.104$, $MSE = 13.5$. The interaction between the two classes and topics was highly significant, $F(1, 36) = 43.1$, $MSE = 5581.1$. These results indicate enhanced test performance after using worked-example problems in the homework phase.

A more detailed analysis may reveal some of the characteristics of the superior performance following worked examples. Differential performance of similar and transfer problems needs to be considered, for example. Table 2 also indicates the mean percentage of problems, excluding transfer problems, that was solved correctly during the test phase for both classes on both types of problems. An ANOVA on these data indicated that there was again no significant difference between classes, $F(1, 36) = 0.3$, $MSE = 142.8$, or between topics, $F(1, 36) = 0.3$, $MSE = 46.6$. The interaction between the two classes and topics was again highly significant, $F(1, 36) = 57.680$, $MSE = 7928.8$. The worked-example group was clearly more proficient at solving test problems similar to the homework problems.

The results also suggest that the effect of worked examples can become apparent after very few trials. In the homework phase for mirrors, there were only 2 problems using convex mirrors but 8 problems using concave mirrors. In the homework phase for lenses, there were only 2 problems using concave lenses but 8 problems using convex lenses. Therefore, both topics included low practice categories of problems. The experimental design ensured that students who were presented with worked examples during homework on one topic were presented with conventional problems on the other topic. It follows that each student had worked examples on the low practice problems from one topic and conventional problems from the other topic. Furthermore, half of the students had worked examples on mirrors followed by conventional lens problems, whereas the other half had conventional mirror problems followed by worked examples on lenses.

In the test phase for both topics, each student was scored correct or incorrect on problems similar to both the low practice worked example and conventional problems from the homework phase. Students could correctly answer the problem presented previously as a worked example but fail to

TABLE 3
Frequencies of Correct Solutions on the Low Practice and Transfer Problems of Experiment 1

| *Worked Example* | *Conventional* | *Low Practice Problems on Both Topics* | *First Transfer Problem on Both Topics* | *Second Transfer Problem on Both Topics* |
|---|---|---|---|---|
| Solved | Not solved | 12 | 6 | 11 |
| Not solved | Solved | 2 | 3 | 5 |
| Solved | Solved | 20 | 6 | 6 |
| Not solved | Not solved | 4 | 23 | 16 |

*Note.* Each frequency refers to the performance on two problems, following worked examples in one case and conventional problems in the other case.

correctly answer the equivalent problem presented as a conventional problem. Table 3 indicates this frequency (which is 12 in Table 3) and the frequencies of the three other possible combinations. A Sign test comparing worked examples (solved), conventional (not solved) with worked examples (not solved), conventional (solved), indicated the difference to be significant.

From these results, it is clear that, when given limited exposure to a conventional problem in the homework phase, relatively few students (compared with worked-example conditions) were able to solve the relevant test problems correctly. In contrast, similarly limited exposure to worked examples resulted in many more students (compared with conventional problem conditions) correctly solving the same test problem. Worked examples allow students to use a rule more competently, even after only two homework problems.

Further differences between groups on the test problems are also indicated by the total number of problems that each group failed to attempt. In the mirror topic, the median number of unattempted test problems per student on the conventional and worked-example groups was 3.0 and 1.0, respectively; whereas in the second topic, lenses, the median number of unattempted test problems per student was 2.0 and 1.0, respectively. In both topics, the worked-example groups had significantly fewer nonattempts than the conventional groups under the Mann–Whitney $U$ test, $U(17, 21) = 43.5$, $z = 3.96$ (mirrors), and, $U(17, 21) = 93.5$, $z = 2.50$ (lenses), respectively. (Throughout this article nonparametric tests are used on nonattempt scores because of the distorted [skewed] distributions.)

We had predicted that heavier use of worked examples should reduce cognitive load by switching attention away from goal-directed search, thereby assisting in both schema acquisition (i.e., allowing students to recognize problem types and their associated moves) and rule automation (using the rules listed in Figure 1 with reduced conscious effort). A reduced cognitive load should permit resources that are no longer required for problem solving search to be directed to learning. Students should be better able to recognize relevant problem states and use the rules with reduced conscious effort.

Rule automation should contribute to superior performance on the test transfer problems to which specific schemas may be less applicable. There were two transfer problems for both mirrors and lenses, resulting in each student being presented with four transfer problems. For purposes of counterbalancing, the initial transfer problems for both mirrors and lenses were analyzed together, with a separate analysis for the second problem. A similar analysis to that carried out on the low frequency problems can be carried out on the transfer problems. Table 3 indicates the frequency of solution of the first transfer problem (following worked examples) on one of the topics, combined with solution (following conventional problems) of the first transfer problem of the alternate topic. Although in the expected direction, a Sign test indicated that a worked example (solved), conventional (not solved) combination was not significantly more frequent than a worked example (not solved), conventional (solved) combination. A similar result was obtained on the second transfer problem.

It should be noted that the lack of effects on the transfer problems could be due to many students choosing to attempt the easier similar problems first and then running out of test time. The results provide some support for this suggestion. In all groups there was considerable evidence of nonattempts on the transfer problems, yet subsequent problems in the test sequence were being answered.

We suggest that the differences between groups are due to an inappropriate focus of attention and heavy cognitive load associated with a means-ends strategy. A possible alternative reason for the results could be due purely to the increased guidance given to the worked-example group in the homework phase. The conventional problem-solving group did not obtain feedback in the homework phase and could, therefore, make errors that carried through into the test phase. Such an occurrence could result in differential test scores due to differential error rates caused by uncorrected errors in the conventional group. Although this result would preserve the educational significance of the findings, it suggests that the theoretical rationale provided is inappropriate. Overall, the significance in test scores could be attributed, not to the use of worked examples facilitating schema acquisition, but rather to the increased feedback in the homework phase for the worked-example group causing fewer errors.

In fact, few errors were made by either group on repeat presentations during homework. Only the repeat problem for each pair in the homework phase was used in the following data, because there was no opportunity for errors by the worked-example groups on the initial problem presentations. Regardless of presentation type, 33 students did not make any errors on the repeat presentation problems. For the remaining five students, a total of four errors were made after both worked examples and conventional problems. Clearly, there were few errors made on the repeat problems (and a nondifferential error rate). Furthermore, if a subject made an error in both the homework and test phases, the type of error was different. Under these circumstances, it is unlikely that uncorrected errors have significantly affected our results.

A comparison of errors, made in the first and second presentation homework problems for students in the conventional problem-solving groups, gives some evidence of a reduction of errors between the first and second presentation problems. The conventional problem groups for both topics made significantly fewer errors on the repeat problems than on the initial problems, using a Wilcoxon matched-pairs signed ranks test, $T(9) = 5$. (The Wilcoxon matched-pairs signed ranks test has an adjusted $N$ according to the presence of any "zero" differences. Only nonzero differences influence the statistic. This accounts for the $N$ of 9 in the preceding statistic.) Eight subjects made a total of 11 errors on the first presentation problems and three errors on the repeat problems. Only one subject made more errors on the repeat problems than on the conventional problems, making zero and one error, respectively. In the test phase, there again were few errors, with no significant difference between the worked-examples and conventional groups, $T(10) = 27.5$.

We can conclude, in this topic, that conventionally worked examples can facilitate learning substantially under normal classroom conditions. The lack of significant transfer effects is, nevertheless, serious. Cooper and Sweller (1987), in a series of experiments designed to investigate the lack of transfer effects found by Sweller and Cooper (1985), found that differences between worked examples and conventional problems were obtainable on transfer problems after extended practice. They concluded that this was because rule automation was required for transfer and that rule automation developed slowly. It follows that lack of transfer in the current experiment may have been caused by insufficient practice during the homework phase. Alternatively, as previously indicated, the result may have been due to too few students attempting the transfer problems. For this reason, the next experiment was designed to ensure that all students attempted these problems.

## EXPERIMENT 2

Although Experiment 1 yielded a clear advantage for worked examples over problems during homework, the design did not permit independent assessment of the effects on transfer problems. It is important to provide evidence of transfer effects for both practical and theoretical reasons. From a practical point of view, if worked examples only facilitate performance on similar problems, this may justify their use but would leave open the possibility that they restrict problem-solving adaptability and flexibility. With respect to theory, rule automation is a primary ingredient of transfer performance. The reduced cognitive load associated with worked examples should not only facilitate schema acquisition and performance on similar problems that a particular schema can incorporate, but it should also facilitate rule automation and thus improve performance on problems not covered by a previously acquired schema (see Cooper & Sweller, 1987). Experiment 2 allowed a test of this hypothesis by ensuring that all subjects attempted the transfer problems.

### Method

*Subjects.* The subjects were 32 Year 10 students from the fourth-level (Class A, composed of 16 students) and sixth-level (Class B, composed of 16 students) science classes of a Sydney high school with seven Year 10 science classes. Students were placed in these classes according to their examination performances from the preceding year. Students in the fourth-level class were ranked 90–120 in the year, and students in the sixth-level class were ranked 150–180 in the year. As was the case in Experiment 1, these classes were chosen to avoid possible asymptotic effects.

*Procedure.* The general procedure, topics, and materials were identical to those used in Experiment 1, except in the test phase, the transfer problems were removed from the main test and examined independently. This was done to ensure that all students attempted the transfer problems and, in contrast to Experiment 1, to provide clearer evidence of possible transfer effects. The two transfer problems for both of the topics of Experiment 1 were replaced in the main test with problems similar to the homework problems. These transfer problems were then presented to the students separately, with a limit of 5 min to solve each problem. The first transfer problem was collected from students before they were permitted to attempt the second.

### Results and Discussion

Table 4 contains the mean percentage of problems correctly solved by each group on both mirror and lens topics during the test phase. An ANOVA on these data, which excluded transfer problems, indicated that there was a significant effect due to class, $F(1, 30) = 7.79$, $MSE = 4115.22$, whereas the difference between the two topics was not significant, $F(1, 30) = 1.21$, $MSE = 3985.36$. The interaction between the two classes and topics was significant, $F(1, 30) = 93.21$, $MSE = 12382.13$.

Further differences between groups on the test problems are also indicated by the total number of problems that each group failed to attempt. In

TABLE 4
Mean Percentage (and Standard Deviations) of Problems Correctly Solved During the Test Phase of Experiment 2

| | *Topic* | |
|---|---|---|
| | *Mirrors* | *Lenses* |
| Class A | 63.2 | 87.9 |
| | (16.6)[a] | (13.8)[b] |
| Class B | 75.0 | 44.0 |
| | (20.2)[b] | (19.1)[a] |

[a]Conventional problems used in the homework phase. [b]Worked-example problems used in the homework phase.

TABLE 5
Frequencies of Correct Solutions on the Low Practice and Transfer Problems of Experiment 2

| *Worked Example* | *Conventional* | *Low Practice Problems on Both Topics* | *First Transfer Problem on Both Topics* | *Second Transfer Problem on Both Topics* |
|---|---|---|---|---|
| Solved | Not solved | 13 | 17 | 18 |
| Not solved | Solved | 4 | 3 | 2 |
| Solved | Solved | 11 | 6 | 5 |
| Not solved | Not solved | 4 | 6 | 7 |

*Note.* Each frequency refers to the performance on two problems, following worked examples in one case and conventional problems in the other case.

the mirror topic, the median number of unattempted problems per test problems presented to the conventional and worked-example groups was 0.38 and 0.11, respectively; whereas in the second topic, lenses, the median number of unattempted problems per test problems attempted was 0.25 and 0, respectively. Replicating the results of Experiment 1, the worked-example groups in both topics had significantly fewer nonattempts than the conventional groups under the Mann–Whitney $U$ test, $U(16, 16) = 67.0$, $z = 2.30$ (mirrors), and, $U(16, 16) = 24.0$, $z = 3.92$ (lenses), respectively.

As in Experiment 1, the results of Experiment 2 suggest that the effect of worked examples can become apparent after very few trials. In the homework phase for both topics, convex mirrors and concave lenses were again treated as low practice categories of problems. In the test phase for both topics, each student was scored correct or incorrect on the low practice problems. The relevant frequencies are indicated in Table 5. A Sign test comparing worked examples (solved), conventional (not solved) with worked examples (not solved), conventional (solved) indicated the difference to be significant. It seems that, when using a conventional problem in the homework phase, far fewer subjects were able to solve the relevant test problems correctly. Worked examples allowed students to solve a similar problem more competently, even after a small number of homework problems.

Further evidence for the effects of learning can be seen most clearly by comparing performance on the transfer problems. Schemas acquired during the homework phase can be used to directly solve the similar test problems but are probably of limited use on the transfer problems. These problems, identical to those used in Experiment 1, were the first problems seen by students in which either a diagram of the components had to be drawn before using the rules or, given the image position, the correct positioning of the object creating the image had to be worked out. It is suggested that, because these problems differed sufficiently from those given for homework, acquired schemas would be of limited use. Students may be more likely to use a search strategy, such as means-ends analysis, when solving the transfer problems. Because the same problem-solving operators (the rules of geometric optics) are required in both the homework and the transfer problems, if differential automation of these operators occurs, then differential transfer should also occur.

In the test phase, each student was scored correct or incorrect on both the first and second transfer problems for mirrors and lenses. A similar analysis to that carried out for the transfer problems in Experiment 1 was used (see Table 5). A Sign test indicated that a worked example (solved), conventional (not solved) combination was significantly more frequent than a worked example (not solved), conventional (solved) combination on both the first and second transfer problems. These results strongly suggest that, under conventional classroom conditions, the use of worked examples can substantially facilitate performance. Furthermore, enhanced facilitation was obtained not only on problems similar to those seen previously, but also on novel problems requiring considerable transformation of previously learned material.

Very few errors were made by either group on repeat presentation problems (problems were presented in pairs of very similar problems—an initial and repeat presentation) during the homework phase, and in no case did an error carry through from the homework to the test phase. Regardless of presentation type, 25 students did not make any errors on the repeat presentation problems. The remaining 7 students made a total of 10 errors after worked examples and 2 errors after conventional problems. These error rates were too small to permit meaningful statistical analyses.

Unlike Experiment 1, a comparison of errors made in the first and second presentation homework problems for students in the conventional problem-solving groups provided little evidence of a reduction of errors between the first and second presentation problems. Too few errors were made to permit a meaningful reduction. A high percentage of subjects made no errors (26 of 32). Five subjects made a total of eight errors on the first presentation problems, and 3 subjects made a total of seven errors on the repeat problems.

In the test phase, the worked-example group had significantly fewer test errors using a Wilcoxon test, $T(20) = 41.1$. Median error rates for the worked example and conventional problem groups were 0.12 and 0.20, respectively. Unlike the previous experiment, the use of worked examples caused a significant reduction in the number of test errors.

In all cases, if a subject made an error in both the homework and test phases, the type of error was different. These results are similar to those of Experiment 1 and unambiguously indicate that the differences between groups were not due to the persistence of uncorrected errors on the part of students solving conventional problems during homework.

These results confirm the advantages of relying heavily on worked examples in a classroom context. Worked examples improve performance on both similar and transfer problems, suggesting a general and robust effect. We suggest the simplest explanation of these and previous findings is that: (a) schema acquisition and rule automation are essential to problem-solving skill; (b) efficient problem solving via a means-ends strategy requires a search for operators (rules) to reduce differences between problem states; (c) this search, although efficient from a problem-solving perspective, is an inefficient learning device, because it inappropriately directs attention and imposes a heavy cognitive load that interferes with learning; (d) with respect to schema acquisition and rule automation, worked examples can appropriately direct attention and reduce cognitive load.

## EXPERIMENT 3

Experiments 1 and 2 demonstrated the efficacy of worked examples over conventional problems using geometric optics under standard classroom conditions. Should we expect conventionally worked examples to be equally effective on all or most topics presented to students? As indicated in the introduction, both cognitive load theory and empirical evidence suggest that, under some specified conditions, worked examples are quite ineffective.

Tarmizi and Sweller (1988) suggested that if a worked example required students to split their attention between multiple sources of information that had to be integrated mentally, the cognitive load generated was sufficiently high to eliminate any advantage of worked examples over conventional problems. Empirical evidence, using circle geometry, supported this suggestion. In contrast, if the worked examples were reformulated to eliminate a split-attention effect, worked examples regained their advantage over conventional problems.

The worked optics examples used in Experiments 1 and 2 did not require students to split their attention among, for example, textual material, sets of equations, and diagrams. As predicted by cognitive load theory, worked examples were effective. In contrast, other areas of physics conventionally use worked-example formats that require students to split their attention among multiple sources of information. Kinematics is one such area. Worked examples consist normally of a problem statement and sets of equations. The equations may or may not be accompanied by textual material. In order to understand the worked example, attention must be directed to mentally integrate its various facets. This process of integration will require cognitive resources to be devoted to activities that, at best, are marginal to schema acquisition and rule automation. It might be predicted that these conventionally worked examples will not prove more effective than solving problems. Experiment 3 tested this hypothesis using kinematics problems.

### Method

***Subjects.*** In the first year of this experiment, the subjects consisted of 17 Year 11 students from the third-level physics class (Class A) of a Sydney high school with five Year 11 physics classes. In the following year, subjects were 17 Year 11 students from the fourth-level physics class (Class B) of the same school. There were six Year 11 physics classes in that year. Students were placed in these classes according to their examination performances from the preceding year. Students in the third-level class were ranked 50–75 in the year, whereas students in the fourth-level were ranked 74–94 in the year. As was the case in the preceding experiments, these classes were chosen to avoid possible asymptotic effects.

***Materials.*** The problems used in this experiment came from four subareas of physics. Topic 1 required students to solve simple linear motion problems in the horizontal direction using one of the following three equations of motion; $v = u + at$, $s = ut + \frac{1}{2}at^2$, and, $v^2 = u^2 + 2as$, in which $v$ = final velocity, $u$ = initial velocity, $a$ = acceleration, $t$ = time, and $s$ = distance. In order to solve each problem correctly, it was necessary for students to choose the equation that included the required variable ($v$, $u$, $a$, $t$, or $s$) and no other unknowns. The following is an example of this type of problem:

> An object with a speed of 12.2 $ms^{-1}$ travels with an acceleration of 8.0 $ms^{-2}$. What is the final speed at the end of 5 seconds?

Topic 2 required students to solve simple multistep linear motion problems in the vertical direction using one of the three aforementioned linear-motion equations. As was the case for Topic 1, problems contained sufficient information to allow students to choose the equation containing the required answer ($v$, $u$, $t$, or $s$) as the only unknown. These problems, unlike the Topic 1 problems, required students to always use a definite value for acceleration (the acceleration due to gravity = 9.8 $ms^{-2}$) and to take the direction of motion into account by including a sign convention. An example of this type of problem is:

> A ball is thrown vertically into the air with a velocity of 100.0 $ms^{-1}$. Taking the acceleration due to gravity = 10.0 $ms^{-2}$, find: (a) how long it took to reach the topmost point of its flight; (b) the greatest height reached; (c) the total time of flight; (d) the velocity at the end of 4.0 seconds; (e) how far the ball is above the ground after 4.0 seconds.

Topic 3 required students to solve projectile motion problems. As was the case for the previous two topics, students chose an appropriate linear motion equation that included the goal as the only unknown. Unlike the previous topics, these equations were applied to either the horizontal or vertical components of the projectile motion. Each projectile motion problem consisted of a number of one-step subgoals, each requiring a different equation for solution. An example of this type of problem is:

> A stone of mass 2.0 kg is thrown with a velocity of 20.0 $ms^{-1}$ at an angle of 40° above the horizontal. Take the acceleration due to gravity as 10.0 $ms^{-2}$. Calculate: (a) the initial horizontal component of velocity; (b) the initial vertical component of velocity; (c) the greatest vertical height reached; (d) the range of the stone.

Topic 4 required students to solve two-dimensional collision problems by correctly analyzing the motion of two objects before and after collision. Students were presented with a stroboscopic photograph of the entire motion from which information could be extracted to calculate the velocity (including direction) of each object. They then had to check for the conservation of momentum before and after collision. The only equations required in this topic, $v = s$(distance)/$t$(time), and $p$(momentum) = $m \times v$, were different from those used in the previous topics.

***Procedure.*** The general format of the experiment followed the design of Experiment 2. The experiment was conducted under conventional classroom conditions as part of the normal teaching program of the school, and the tests were considered school tests by the students. Each topic was completed separately before proceeding to the next topic.

During the experiment, three phases similar to those of the previous experiments were effected. In the first phase, a teaching phase, all topics were taught in detail with emphasis on the necessary equations. Problems were worked through by the teacher, and the students also tried solving a set of exercises themselves. In addition, in Topic 1, they were also asked to solve one 2-move problem to give them some experience with this type of problem. All problems were corrected, and any difficulties answered.

The teaching phase was followed by the homework phase, during which students were presented with a set of homework problems. These were presented in pairs with every second problem being similar to the immediately preceding one, as in Experiment 2. Students were instructed to solve these problems as accurately as possible by using the previously taught theory and were told that the homework would be collected and marked the following day. They were further instructed that they would be given a test on the homework immediately following the collection of these problems.

The number of homework problems given varied according to the difficulty of the topic taught. In Topic 1, "linear motion in the horizontal direction," eight homework problems were presented as four pairs, with two identical format problems in each pair. Three of the four pairs were one-move problems. On these three, problems within a pair required the same equation, but a different motion equation was used on each pair of problems. The last two problems were two-move problems requiring the use of one equation to find an unknown that could then be used in a second equation to solve the problem.

In Topic 2, "linear motion in the vertical direction," six multistep homework problems were presented as three pairs with two identical format problems in each pair. Each pair varied according to the initial firing position and subsequent direction of travel. Within a problem, there were a number of subquestions covering aspects of the motion.

In Topic 3, "projectile motion," six homework problems were presented as three pairs with two identical format problems in each pair. Similar to the previous topic, problems differed according to the firing position and angle of firing.

In the final topic, "collisions," two similar format homework problems were given to students. In all topics, the homework problems were similar to those used in the previous teaching phase. Students had access to the teaching phase information during the homework phase.

The next day, a test phase followed for each topic. The format of the main test used as well as length of time given to attempt the test varied with the topic. Some topics required more reading than others. In Topic 1, "horizontal linear motion," 12 conventional problems were presented on paper. Test Problems 1 to 12, excluding Test Problem 2, were identical in structure to the one-move homework problems requiring the use of one of the equations for solution. Test Problem 2 was identical in structure to the last pair of homework problems requiring the use of two moves for solution. Students were allowed 15 min to complete the test.

In the second topic, "vertical motion," six conventional problems identical in structure to the homework problems were used, with a total of 26 subquestions. In order to ensure that students' attention focused solely on solving the current problems, they were presented in pairs, on paper, with all previous workings and problems being collected before students attempted succeeding pairs of problems. For example, the first pair of problems was presented to the students with instructions to solve the first of the pair and, if time permitted, to continue with the second of the pair. These problems were then collected and the second pair distributed. This procedure was continued until the completion of all test problems, with students being allowed a maximum of 8 min for completion of each pair of problems.

In the third topic, "projectile motion," a similar procedure to the previous topic was used. Six conventional problems identical in structure to the homework problems were used, with a total of 24 subquestions. Problems were presented in pairs on paper. The first pair of problems was collected before the second pair was distributed. This procedure was repeated until the completion of the test. Students were allowed 12 min to complete each pair of problems.

In the final topic, "collisions," two conventional problems were presented, one at a time. Both problems were similar in structure to the homework problems. Subjects were allowed 15 min on each of the problems.

To observe possible transfer effects, a second test was given, after the main test, on each topic. These problems differed in structure from the problems used in the homework phase. In Topic 1, "horizontal linear motion," students were required to construct a problem requiring two moves to attain a solution. In Topic 2, "vertical motion," the format varied slightly from those seen in the homework phase, with information presented in tabular form. Nevertheless, the problem could be solved using the same manipulations used in the solution of the homework problems. In Topic 3, "projectile motion," although the transfer problem was somewhat different in format, with information presented in pictorial and word form, the same calculations used in the homework phase applied. In the final topic, "collisions," the transfer problem contained information that allowed the momentum of both objects before collision to be calculated but gave insufficient information to calculate the momentum of one object after collision. Problem solution required using all available information and the conservation of momentum principle to calculate the unknown momentum.

In both the main test and the test for transfer, students did not have access to previous work that had been conducted in the teaching and homework phases. As was the case in the preceding experiments, the three phases (teaching, homework, and test) were used on all topics, with the distinction between worked examples and conventional problems occurring in the homework phase. The worked-example group had the first of each problem pair presented as a conventionally worked example in which the problem statement was given with the problem's solution path immediately beneath. The conventional-problem group was required to solve each of the problems that were presented in conventional fashion.

The requirement for students to split their attention between and mentally integrate the problem statement and the problem's solution path on worked examples of all four topics should be noted. In each case, the solution path had little meaning without referring back to the problem statement.

*Experimental design.* Both classes were taught the four topics using an alternation procedure in the homework phase for counterbalancing. In Class A, which was taught in the first year of the experiment, Topic 1 problems in the homework phase were given as worked-example types, Topic 2 problems as conventional problems, Topic 3 as worked-example types, and Topic 4 as conventional problems. The presentation method was reversed for Class B, which was taught in the second year of the experiment. Topic 1 was presented as conventional problems, Topic 2 as worked examples, Topic 3 as conventional problems, and Topic 4 as worked examples. Such counterbalancing ensured that any overall differences between worked examples and conventional problems were not likely to be contaminated by differences between the two classes or the four topics.

## Results and Discussion

The performances of each group during both the homework and test phases were analyzed. Any differences between groups should have been revealed by the performance of each group on all topics during the test phase. Table 6 shows the mean percentage of problems correctly solved by each group (Group A and Group B) on the four topics during the test phase. A problem was scored as being correctly solved if the correct solution paths were followed and the correct numerical answer was obtained. (Although some students followed the correct solution path but made numerical errors, no students obtained the correct numerical answer following an incorrect path.)

There was no significant difference between the conventional and worked-example groups in the number of test problems correctly solved on the horizontal, linear motion topic, $F(1, 32) = 0.04$, $MSE = 0.1176$, or on the second topic, vertical motion, $F(1, 32) = 3.55$, $MSE = 36.0294$, $.05 < p < .10$. In contrast, on the remaining topics, projectile motion and collisions, the conventional groups solved significantly more test problems than the worked-example groups, $F(1, 33) = 7.57$, $MSE = 79.53$, and, $F(1, 33) = 8.70$, $MSE = 2.9412$, respectively.

These results indicate that topic areas for which conventionally worked examples require students to split their attention between multiple sources

TABLE 6
Data From Experiment 3

| | *Topic* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Linear Motion (Horizontal)* | | *Linear Motion (Vertical)* | | *Projectile Motion* | | *Collisions* | |
| | *A* | *B* | *A* | *B* | *A* | *B* | *A* | *B* |
| | *Mean Percentage (and Standard Deviations) of Test Problems Solved* | | | | | | | |
| | *WE* | *CP* | *CP* | *WE* | *WE* | *CP* | *CP* | *WE* |
| | 50.0 (11.8) | 48.3 (15.5) | 45.8 (12.8) | 53.8 (9.6) | 42.2 (12.1) | 54.9 (14.8) | 88.2 (12.3) | 66.7 (16.4) |
| | *Median Number of Nonattempted Problems* | | | | | | | |
| | 5 | 4 | 12 | 7 | 8 | 5 | 0 | 0 |
| | *Number of Subjects Correctly Solving the Transfer Problems* | | | | | | | |
| | 3 | 5 | 7 | 5 | 6 | 8 | 4 | 3 |
| | *Median Number of Errors Per Homework Problems Attempted* | | | | | | | |
| Initial | NA | 0 | 0 | NA | NA | 0.5 | 0 | NA |
| Final | 0 | 0 | 0.17 | 0 | 0.17 | 0.25 | 0 | 0 |
| | *Median Number of Test Errors Per Problems Attempted* | | | | | | | |
| | 0.14 | 0.11 | 0.08 | 0.12 | 0.29 | 0.20 | 0 | 0.5 |

*Note.* WE refers to worked examples used in the homework phase. CP refers to conventional problems used in the homework phase. NA means not applicable.

of information are no more advantageous with respect to schema acquisition than conventional problems. Worked examples may, in fact, have adverse effects. In two of the four topics, worked examples interfered with test performance compared with conventional problems. On projectile motion, Class B was given conventional problems and was superior, whereas, on collisions, Class A was given conventional problems and was superior.

The first topic, linear motion, did not yield a difference between groups on the test problems, but this may have been because the problems were too easy, resulting in asymptotic effects. (The approximated 50% success rate on these problems was more a function of the excessive number of problems presented during the limited testing sessions than of problem difficulty.) More difficult problems may allow differences to appear. This hypothesis can be tested by the use of a more difficult, two-move test problem. The test phase of Topic 1 included such a problem. In one class (the second class), 15 students correctly solved the two-move problem after being presented with conventional problems in the homework phase, whereas two students incorrectly solved the problem. In the other class (the first class), only eight students solved the two-move problem after being presented with worked example problems in the homework phase, whereas nine students incorrectly solved the problem. A chi-square test performed on these frequencies indicated that there was a significant effect on this two-move problem, $\chi^2(1, N = 34) = 4.84$. Inspection of the frequencies indicated that the difference was due to more students from the conventional problem-solving group, as opposed to the worked-example group, correctly solving the problem. This result suggests that the same mechanisms favoring conventional homework problems, as opposed to worked examples, were operating in Topic 1 as in Topics 3 and 4. The only difference may be in the difficulty of the test problems used.

Further differences between groups on the test problems may also be indicated by the total number of problems that each group failed to attempt. Table 6 shows the median number of nonattempts for both groups on each topic. In the horizontal linear motion topic, the conventional and worked-example groups did not differ significantly on unattempted problems, $U(17, 17) = 118.5$, $z = .89$. In the second topic, vertical motion, the worked-example group had significantly fewer unattempted problems than the conventional group, $U(17, 17) = 73.5$, $z = 2.45$. In the third topic, projectile motion, the conventional group had significantly fewer nonattempts than the worked-example group, $U(17, 17) = 89.5$, yielding a normal approximation adjusted for ties of $Z = 2.02$ (a correction for ties was used only when its use could alter conclusions). In the remaining topic, collisions, all test problems were attempted, regardless of treatment in the homework phase. It should be noted that lower nonattempt scores were obtained only by the second class, irrespective of homework format. This result suggests that the class was superior to the other class and that differences were due largely to this factor rather than to differences in homework problems.

Analysis of the transfer problems given in the test phase indicated a lack of difference between groups on all topics. Table 6 indicates the frequencies of transfer problems correctly solved for the two groups on each topic. Chi-square tests were performed on these frequencies, and the following results were obtained: in Topic 1, horizontal linear motion, $\chi^2(1, N = 34) = 0.18$; in Topic 2, vertical motion, $\chi^2(1, N = 34) = 0.13$; in Topic 3, projectile motion, $\chi^2(1, N = 34) = 0.14$; and in the final topic, momentum, $\chi^2(1, N = 34) = 0.08$.

Table 6 includes the median error rates on homework problems (with the homework period separated into initial and repeat problem presentations) attempted for conventional and worked-example groups on all topics. As can be seen, few errors were made. Mann–Whitney $U$ tests were performed on each topic comparing errors made on the repeat problems in both groups. In all topics, there was no difference between the conventional and worked-examples group in errors for the second (repeat) presentation problems: $U(17, 17) = 136.8$, $z = 0.26$, for the linear motion topic; $U(17, 17) = 138$, $z = 0.22$, for the vertical motion topic; and for the projectile motion topic, $U(17, 17) = 153.9$, $z = 0.32$. In the remaining topic, collisions, an extremely small error rate was recorded in both groups; only two subjects in each group made a total of two errors each.

Comparing the number of errors made on the initial and repeat presentation problems, the conventional-problem group, for the projectile motion topic, made significantly more errors on the initial problems (median error rate = 0.5) than on the repeat problems (median error rate = 0.25) using a Wilcoxon matched-pairs signed ranks test, $T(12) = 3.5$, but a similar difference was not obtained in the conventional-problem groups for the other topics: $T(7) = 5.0$ for linear motion; $T(8) = 13.0$ for vertical motion; and, for the collisions topic, $T(7) = 12$. The results obtained from the last three topics are biased due to the large number of subjects who made no errors on either presentation.

Table 6 also shows the median error rates in the test phase for both groups on all topics. There was no difference between groups for errors on the linear motion topic, $U(17, 17) = 115.1$, $z = 1.01$. In the projectile motion topic, the conventional group made significantly fewer errors than the worked-example group, $U(17, 17) = 84.5$, $z = 2.07$; whereas, in the vertical motion topic, the conventional group made significantly fewer errors, $U(17, 17) = 91$, $z = 1.52$, $p = 0.06$. Similarly, in the collisions topic, the difference in errors between the conventional (mean errors per problems attempted = 0.29) and worked-example groups (mean errors per problems attempted = 0.94) may represent a real effect favoring the conventional group, $U(17, 17) = 99$, $z = 1.57$, $p = 0.06$.

The results provide evidence that using conventionally worked examples on the topics of this experiment does not facilitate learning. This finding contradicts the findings of the first two experiments and of Cooper and Sweller (1987) and Sweller and Cooper (1985) using algebra worked examples but confirms the results of Tarmizi and Sweller (1988) using circle geometry. In the current experiment using physics problems, guidance provided by conventionally worked examples has not only failed to facilitate subsequent performance but may well have retarded learning. As suggested by the preceding hypothesis, these results may be due to conventionally worked examples in the areas studied requiring students to mentally integrate disparate sources of information. The resultant cognitive load may interfere with learning. From an educational perspective, the results suggest that the substitution of conventional, kinematics worked examples for problems will be either ineffective or even counterproductive.

Experiment 4 tests alternative formats of worked examples designed to accord more closely with our hypothesized cognitive processes.

## EXPERIMENT 4

Unlike the preceding experiments, this experiment was designed after Tarmizi and Sweller's (1988) finding that a requirement to mentally integrate multiple sources of information could interfere with learning from worked examples. The intention of the present experiment was to explain the failure of Experiment 3 to yield an advantage to worked examples using kinematics problems.

As indicated previously, Tarmizi and Sweller (1988) found that conventional, geometry worked examples failed to facilitate subsequent problem-solving performance due to students having to split their attention between the geometrical diagram and the equations and theorems that referred to the diagram. By integrating the diagram and the textual material, worked examples proved superior to conventional problem-solving activity. It is suggested that kinematics worked examples could be similarly advantageous, but only if they are presented in a manner that eliminates the need for students to divide their attention among multiple sources of information. Although kinematics worked examples normally are presented without diagrams, students still must split their attention between the problem statement and the sets of equations. These must be integrated by the student in order to make sense of the example. It seems plausible to suggest, therefore, that the requirement to mentally integrate a problem statement and a set of equations eliminated any facilitatory effects due to worked examples in Experiment 3.

Experiment 4 tested this hypothesis by presenting kinematics worked examples in a format that obviated the need for students to divide their attention among multiple sources of information. The elimination of the usual requirement to divide attention between the problem statement and the solution equations can be accomplished by incorporating the solution equations into the problem statement rather than separating them. Using a modified format should reduce the need for students to map the solution equations onto the problem statement; the presentation format should provide the necessary integration. Because a requirement to map equations onto a statement is likely to impose a heavy cognitive load and to be an irrelevant activity for schema acquisition and rule automation, learning should be facilitated by the elimination of such a requirement.

### Method

*Subjects.* The subjects were 45 Year 10 students from two fourth-level science classes of a Sydney high school with seven Year 10 science classes.

The experiment was conducted over 2 years to allow sufficient subjects of similar ability to be tested from the same school. Subjects were allocated to classes according to their examination performances in the preceding year. Students in the fourth-level class were ranked 90–120 in the year. The fourth-level science class was chosen to avoid possible asymptotic effects. Although the material used in this experiment is similar to that of Topic 1, "horizontal linear motion" of Experiment 3, the one-step problems used in this experiment were easier because they only required students to choose between two motion equations for solution (see next section), instead of the three in the preceding experiment. Simpler problems allowed us to use Year 10 rather than the Year 11 students used in Experiment 3.

*Procedure.* The general format of the experiment followed the design of the previous experiment, except that random allocation of students into groups was used. Randomization eliminated the need to counterbalance topics and classes. As was the case in the preceding experiments, the same three phases (initial teaching, homework, and test) were effected. The teaching phase began with a description and definition of two of the equations of motion that can be used when dealing with uniform acceleration. One-move problems were demonstrated by the teacher using each equation, and students also tried solving a set of one-move exercises themselves. In addition, students were asked to solve one 2-move problem to give them some experience with this type of problem. All other aspects of the teaching phase were similar to those used in previous experiments.

The 18 homework problems were presented as nine pairs, with two identical format problems in each pair. Four of the pairs could be solved using the equation $v = u + at$; four pairs needed the second equation, $s = ut + 0.5at^2$; and the last pair was a two-move problem requiring both equations for solution. All homework problems were similar to the problems presented in the teaching phase. The difference between one pair of problems and another pair using the same equation for solution was in the information given and the goal required. For example, four different pairs of problems using the equation $v = u + at$ could be used where the goal was one of the variables $v$, $u$, $a$, or $t$. The information presented in the teaching phase was available throughout the homework phase.

Students were assigned randomly to one of three 15-member groups. The conventional group was required to solve 16 problems using pencil and paper. The conventionally worked-examples group was given the same problems, except that the first problem of each pair of identical format problems had the solution written out in a conventional manner. The following provides an example:

> A car moving from rest reaches a speed of 20 m/s after 10 seconds. What is the acceleration of the car?

$u = 0$ m/s
$v = 20$ m/s
$t = 10$ s

$v = u + at$
$a = (v - u)/t$
$a = (20 - 0)/10$
$a = 2$ m/s$^2$

Students in this group had to split their attention between the problem information and the worked solution.

The third group, the integrated worked-examples group, was given the same problems, except that the first problem of each pair of identical format problems had the solution written out by incorporating the solution into the problem statement. The following provides an example:

> A car moving from rest ($u$) reaches a speed of 20 m/s ($v$) after 10 seconds ($t$): [$v = u + at$, $a = (v - u)/t = (20 - 0)/10 = 2$ m/s$^2$]. What is the acceleration of the car?

Subjects in this group were given general instructions identical to the conventionally worked-examples group.

The next day a test phase followed, consisting of three class tests. In the first test, a total of 12, one-move, conventional problems was presented on paper, and students were given 7 min to complete this test. The problems given were similar in structure to the homework problems requiring the use of one of the equations to find an unknown. In the second test, students were instructed to "construct a problem in which you first calculate $t$, assuming that $v$, $u$, and $a$ are given, and then calculate $s$." This test investigated possible transfer effects by requiring students to construct the two-move word problem. The final test consisted of a single two-move problem requiring both equations for solution and was similar to a pair of problems used in the homework phase.

A maximum of 4 min was allowed on both the transfer and the two-move problems, with each problem being collected before the next one was attempted. The students did not have access to their answers to previous problems attempted or the solutions to unsolved problems.

## Results and Discussion

The performances of each group during both the homework and test phases were compared. Data from the test phase are summarized in Table 7, which provides the mean number of problems correctly solved by each group. The integrated worked-examples group solved significantly more test problems

TABLE 7
Data From Experiment 4

| | *Conventional Problem* | *Integrated Worked Example* | *Conventional Worked Example* |
|---|---|---|---|
| Mean number (and standard deviations) of problems solved | 4.33 (2.09) | 6.67 (1.50) | 3.4 (2.69) |
| Number of subjects solving the two-move problem[a] | 6 | 11 | 5 |
| Number of subjects solving the transfer problem[a] | 5 | 10 | 7 |
| Median error rate | 0.23 | 0 | 0.29 |

[a]$n = 15$ per group.

than the other two groups, $F(1, 42) = 16.9$, $MSE = 41.6$. The conventionally worked examples and the conventional-problem groups were not significantly different from each other, $F(1, 42) = 1.4$, $MSE = 13.87$. These results indicate that, by eliminating the need for students to split their attention between multiple sources of information, integrated worked examples of the type used in this experiment have facilitated learning. Although there was no significant difference between the conventionally worked examples and conventional-problem groups, the mean number of problems solved indicates a trend similar to that found in Experiment 3; more problems were solved by the conventional-problem group. (It might be noted that no students wrote their answers using an integrated format. Insufficient space was left on the problem statement, thereby preventing the use of such a format.)

As was the case in Experiment 3, the test phase included a single two-move problem. Table 7 indicates the frequencies of correct solutions for the three groups on this problem. A chi-square test was performed on these frequencies. The test indicated that there may have been a real effect, $\chi^2(2, N = 45) = 5.52$, $.05 < p < .1$. Inspection of the frequencies indicates this possible difference is due primarily to an increased number of students in the integrated worked-examples group that solved the two-move problem.

Possible differences between the groups on the test problems are also indicated by the total number of problems that each group failed to attempt. Analysis of nonattempts on the test problems indicated that the integrated worked examples, the conventional worked examples, and the conventional groups had medians of 6.0, 9.0, and 8.0 nonattempts, respectively. The integrated worked-example group made significantly fewer nonattempts than the other two groups using a Mann–Whitney $U$ test, $U(15, 30) = 76$,

$z = 3.59$. The conventional problem and the conventionally worked-example groups did not differ significantly from each other, $U(15, 15) = 94.5$, $z = .74$.

The test phase transfer problem was used to indicate possible transfer effects. Table 7 indicates the frequency with which the transfer problem was correctly solved by the three groups. A chi-square test was performed on these frequencies. The test indicated that there was no significant difference between groups on this transfer problem, $\chi^2(2, N = 45) = 3.39$. There were, nevertheless, twice as many subjects from the integrated worked-examples group than the conventional-problem group who solved the transfer problem.

The median error rate for the integrated worked examples, the conventionally worked examples, and the conventional groups on the repeat homework problems was 0.1, 0.2, and 0.1, respectively. There was no difference in errors made between the integrated worked-examples group and the other two groups, $U(15, 30) = 224.5$, $z = .02$, or between the conventionally worked example and conventional-problem groups, $U(15, 15) = 92$, $z = .85$. In the conventional-problem group, there was no significant difference in errors between the initial and repeat presentation problems using a Wilcoxon matched-pairs signed ranks test, $T(8) = 8.5$.

Table 7 shows the median error rate for the three groups during the test phase. The integrated worked-examples group made significantly fewer errors than the other two groups, $U(15, 15) = 71.5$, $z = 1.70$, whereas the conventional and conventionally worked-example groups did not differ significantly, $U(15, 15) = 135$, $z = .93$. These results further support the suggestion that presenting worked examples in the conventional, split-source format may have imposed a heavy cognitive load on the students and, in effect, hindered development of problem-solving skills. In the integrated worked-examples group, the reduction in split attention appears to have enhanced learning and reduced error rates.

The results of this experiment, in conjunction with those of Experiment 3, confirm that, although studying worked examples can be a highly effective mode of learning, the presentation format used is critical. Using kinematics problems, guidance provided by conventionally worked examples failed to facilitate subsequent performance. In contrast, the increased number of problems solved and the reduced number of errors and nonattempts in the integrated worked-example group provide strong evidence of enhanced performance.

We can conclude that conventional presentation formats are likely to be effective only insofar as they minimize the need for students to integrate disparate sources of information. As Experiments 1 and 2 indicated, presentation formats in some areas are traditionally structured in a manner that facilitates learning. Kinematics worked examples are not so structured. Nevertheless, they can be reformatted in order to eliminate a split-attention effect. The resultant reduction in cognitive load facilitates learning, thus allowing the superiority of worked examples over conventional problems to manifest itself.

## EXPERIMENT 5

The manner in which worked examples are presented normally is arbitrary with respect to split attention and cognitive load. As we have seen from the previous experiments, conventionally worked examples are appropriately structured with respect to cognitive factors in some areas but not in others. The standard geometric optics worked examples of Experiments 1 and 2 were appropriately structured in that they did not require students to split their attention. As a consequence, they were effective. The more detail and assistance a worked example provides, the more difficult it is to format the problem with a unitary structure. Additional information intended to be helpful to students but not strictly necessary may be difficult to integrate physically with essential, core information, leaving students to accomplish the integration mentally, with the deleterious effects we have seen in the previous experiments.

Experiment 5 tests the hypothesis that geometric optics worked examples will lose their effectiveness if they include a textual explanation of the ray diagrams in association with, but not incorporated within, the diagram. According to our hypothesis, students, in integrating the text and diagram, are engaging in a cognitively demanding activity irrelevant to schema acquisition and rule automation. We can thus make the counterintuitive prediction that the more explanatory material we give students in a worked example, the less effective it will be. Mirror and lens problems identical to those of Experiments 1 and 2 were used.

### Method

*Subjects.* The subjects consisted of the following: (a) in the mirror topic, 48 Year 10 students (18 from the third-level, 18 from the fourth-level, and 12 from the seventh-level science classes of a Sydney high school with eight Year 10 science classes) and (b) in the lens topic, 45 Year 10 students from the same level classes (18 subjects from the third-level, 18 from the fourth-level, and 9 subjects from the seventh-level science classes at the same school). The same students participated in both topics. It should be noted that three students in the lens topic failed to complete the homework task and were, therefore, removed from the experiment. Students were placed in these classes according to their examination performances from the preceding year. Students in the third-level class were ranked 40–60 in the year, in the fourth-level class 61–80, and in the seventh-level class 140–160.

*Procedure.* The problem sets in this experiment consisted of the mirror and lens problems used in Experiments 1 and 2. In the homework phase, both topics contained 10 homework problems that were presented as five pairs, with every second problem being similar to the immediately preceding one. Four of the five pairs dealt with locating the image formed by one optical component (in Topic 1, concave mirrors; in Topic 2, convex lenses), and the remaining fifth pair consisted of two problems on either convex mirrors for Topic 1 or concave lenses for Topic 2. The rules used to construct the ray diagrams for both optical components were identical to Experiments 1 and 2. Unlike Experiments 1 and 2, counterbalancing was not used in this experiment, although both topics (mirrors and lenses) were used to see if the results were replicated on both topics.

In the mirror topic, students were divided randomly into three groups, with 16 subjects per group; in the lens topic, students from the same classes were randomly assigned to three groups, with 15 subjects per group. Care was taken to ensure that each experimental group had the same number of subjects from a particular level science class. Regardless of topic, subjects from the conventional-problem group were required to construct ray diagrams to solve all homework problems that were presented in conventional fashion (see Figure 1). The split-attention group had the first of each pair of homework problems presented as a worked example, for which considerable information concerning the rules was included in the answer (see Figure 1). Students had to split their attention between this information and the ray diagram. Students were required to solve the second problem of each pair presented conventionally. This problem was similar to the first problem in the location of the object but differed in the size and orientation of the object. The conventionally worked-examples group was similar to the worked-example groups of Experiments 1 and 2. They followed the same procedure as the split-attention group, except that they were shown the worked answers required for solution without the extra information given in the split-attention group (see Figure 1).

A test phase followed for each topic on the next day. Three class tests were used on each topic. In the mirror test, nine conventional problems were presented on paper, and students were allowed 10 min to complete this test. These test problems were similar in structure to the first four pairs of homework problems, requiring use of the four rules to find an image in a concave mirror. The second test consisted of a single problem, similar in structure to the last pair of homework problems, requiring the use of four rules to find an image in a convex mirror. Three min were allowed on this problem. The third test consisted of two transfer problems. They were identical to those used in Experiments 1 and 2 and were presented separately, with a limit of 5 min to solve each problem.

In the lens test, eight conventional problems were presented on paper, and students were allowed 10 min to complete the test. These test problems

TABLE 8
Test Phase Data for Experiment 5

| | *Split Attention Worked Example* | *Conventionally Worked Example* | *Conventional Problem* |
|---|---|---|---|
| | *Mean Percentage of Problems Solved (and Standard Deviations)* | | |
| Mirrors | 59.66 (14.62) | 83.98 (17.13) | 65.21 (11.01) |
| Lenses | 54.17 (29.46) | 84.17 (16.12) | 60.0 (18.37) |
| | *Median Number of Nonattempted Problems* | | |
| Mirrors | 0.0 | 1.5 | 1.5 |
| Lenses | 3.0 | 0.0 | 2.0 |
| | *Number of Subjects Solving Low Practice Problems*[a] | | |
| Mirrors | 5.0 | 12.0 | 7.0 |
| Lenses | 4.0 | 11.0 | 6.0 |
| | *Number of Subjects Solving First (and Second) Transfer Problems*[a] | | |
| Mirrors | 5 (5) | 12 (12) | 7 (6) |
| Lenses | 3 (5) | 11 (11) | 5 (5) |

[a]$n$ = 16 per group for mirrors and 15 per group for lenses.

were identical in structure to the first four pairs of homework problems, requiring the use of four rules to find an image in a convex lens. The second test consisted of a single problem, similar in structure to the last pair of homework problems, requiring the use of four rules to find an image in a concave lens. Three min were allowed on this test. The third test consisted of two transfer problems identical to those used in Experiments 1 and 2.

## Results and Discussion

The performances of each group during both the homework and test phases were compared. Data from the test phase are summarized in Table 8, which provides the mean percentage of problems correctly solved by each group on both mirror and lens problems. In both topics, mirrors and lenses, the conventionally worked-examples group solved significantly more test problems than the other two groups, $F(1, 45) = 22.2$, $MSE = 4950.754$, and, $F(1, 42) = 14.02$, $MSE = 7335.069$, respectively. The split-attention and the conventional problem-solving groups were not significantly different from each other on mirrors or lenses, $F(1, 45) = 1.103$, $MSE = 246.420$, and, $F(1, 42) = 0.49$, $MSE = 255.208$, respectively.

These results indicate that, unlike Experiment 4, conventionally worked examples of the type used in this experiment integrate the material into a single source, reduce cognitive load compared with conventional problems, and so facilitate learning. When the examples were modified to introduce multiple sources of information, the worked example advantage was eliminated; there was no significant difference between the split-attention and conventional problem-solving groups.

Differences between the groups on the test problems also can be indicated by the total number of problems that each group failed to attempt. Table 8 shows the median number of nonattempts in both topics in the test phase. In the mirror topic, the conventionally worked-example group did not differ significantly from the other two groups in the median number of nonattempts, $U(16, 32) = 220.5$, $z = .78$. The conventional and the split-attention groups also did not differ significantly from each other, $U(16, 16) = 122$, $z = .23$. In contrast, in the lens topic, the conventionally worked-example group differed significantly from the other two groups in the median number of nonattempts, $U(15, 30) = 103.8$, $z = 2.9$. The split-attention and the conventional problem-solving groups were significantly different also, $U(15, 15) = 58.5$, $z = 2.24$, favoring fewer nonattempts by the conventional problem-solving group. These results indicate that, depending on the topic, the conventionally worked-example group either showed no significant difference in unattempted problems or had fewer nonattempts than the other two groups.

As was the case in Experiments 1 and 2, the effect of a reduced cognitive load in the worked-example group became apparent after very few trials. In the homework phase for both topics, convex mirrors and concave lenses were treated as low practice problems. Unlike Experiments 1 and 2, a single low practice problem for each topic was given to subjects separately after the main test. This procedure ensured that all students attempted the low practice problems given to them. Table 8 indicates the frequency with which the low practice problems were correctly solved by the three groups on both mirror and lens topics. Chi-square tests were performed on each of these sets of frequencies. The tests indicated that there was a significant effect on both the low practice mirror problem, $\chi^2(2, N = 48) = 6.5$, and on the low practice lens problem, $\chi^2(2, N = 45) = 6.97$. Inspection of the frequencies indicated that the differences were due primarily to more conventional worked-example-group students solving these problems.

Table 8 indicates the frequencies with which the two transfer problems in each topic were correctly solved by the three groups. Chi-square tests were performed on these frequencies. The tests indicated that there were significant differences between groups on all transfer problems, with, $\chi^2(2, N = 48) = 6.5$, and, $\chi^2(2, N = 48) = 7.2$, for the first and second presentation mirror transfer problems, respectively, and, $\chi^2(2, N = 45) = 7.14$, and, $\chi^2(2, N = 45) = 6.97$, for the first and second presentation lens transfer problems, respectively. Inspection of the frequencies indicated that all differences were due primarily to the conventionally worked-example group solving more transfer problems.

As was the case in the previous experiments, very few errors were made by each group in the homework phase on the repeat presentation problems. No errors were made on the repeat problems by the three groups on the mirror topic, and a total of three errors, one from each group, was made on the lens topic.

The conventional-problem group for mirrors made more errors on the initial problems than on the repeat presentation problems using a Wilcoxon matched-pairs signed ranks test, $T(6) = 0$. Median error rates for both presentations were 0.0. It should be noted that the medians are influenced by the 10 students who scored no errors on either presentation. Only 6 subjects made a total of nine errors. A similar effect also occurred in the conventional-problem group for lenses, $T(7) = 0$, with 9 subjects making no errors on either presentation. In this topic, 7 subjects made a total of 10 errors.

Analysis of errors on the test problems indicated that, in the mirror topic, the conventionally worked-example, the split-attention, and the conventional-problem groups, median error rates were 0.0, 0.15, and 0.19 per test problems attempted, respectively. Equivalent figures for the lens topic were 0.0, 0.25, and 0.17, respectively. The 0.0 median error rate scores in both topics for the conventionally worked-example group are due to the large number of students making no errors (14 of 16 students in both topics). Analysis of these results indicated that in both topics, mirrors and lenses, the conventionally worked-example group made significantly fewer test errors than the other two groups, $U(16, 32) = 50.5$, $z = 4.50$, and, $U(15, 30) = 100$, $z = 3.01$, respectively. There was no significant difference between the split-attention and the conventional problem-solving groups, $U(16, 16) = 98.5$, $z = 1.11$, for mirrors, and, $U(15, 15) = 103.0$, $z = .39$, for the lens topic.

The results of this experiment, in conjunction with those of the previous experiments, confirm that, although the use of worked examples in the classroom can facilitate learning, the method of presenting these problems is critical. Using the geometrical optics topic, split-attention worked-example problems have not only failed to facilitate subsequent performance but may have retarded learning. This effect occurred despite the considerable amount of additional information given to the split-attention group. In contrast, the conventionally worked-example group, by solving more test problems with fewer errors, provided strong evidence that this group learned more than the other two groups. As was the case in Experiment 2, increased learning occurred not only on problems similar to those seen previously, but also on novel problems requiring considerable transfer of previously learned material.

We can conclude that conventional presentation formats are likely to be effective only insofar as they minimize the need for students to integrate disparate sources of information. This experiment has shown how geometric optics problems can be reformatted in order to increase the split-attention effect. In effect, we have demonstrated that providing students with additional, seemingly useful information can have negative consequences. The resultant increase in cognitive load imposed by the need to integrate this information with core material impedes learning, thus allowing the superiority of conventional over split-attention worked examples to manifest itself.

## GENERAL DISCUSSION

The experiments discussed in this article provide evidence that, under conventional classroom conditions, a heavy use of worked examples, if appropriately structured, can facilitate subsequent problem solving in a variety of physics subareas. Improved test performance can occur on both similar and transfer problems but will not occur following worked examples that require students to split their attention between, and mentally integrate, two or more sources of information.

These results are of interest from several perspectives. First, unlike most prior studies, they were not obtained under laboratory conditions, but rather within a routine teaching context. Second, worked examples have not previously been shown to be effective using physics problems. The ability to obtain similar results under a wide variety of conditions suggests robust findings that are applicable under a variety of circumstances. Third, and most important, the results demonstrate the conditions under which we can expect worked examples to be effective, and they show us how to make the necessary modifications required to transform ineffective worked examples into effective ones.

In essence, our findings indicate that the use of worked examples is not in itself a critical point. Rather, the cognitive consequences of presentation formats need to be addressed. Recently, evidence has begun to accumulate concerning the importance of solution of superficial aspects of problem formats (see Ross, 1984, 1987, 1989a, 1989b). Our current work suggests that formats must be designed to facilitate schema acquisition and rule automation. A presentation format that directs attention to problem states and their associated moves and reduces cognitive load appears to be the critical consideration. Conventional problem solving using a means-ends strategy, or conventionally worked examples requiring the integration of multiple sources of information, are both ineffective vehicles for learning because of their inappropriate attentional and cognitive load characteristics. Because a worked example will not necessarily reduce cognitive load or appropriately direct attention, it cannot be guaranteed to facilitate learning. It appears probable, nevertheless, that most worked examples can be reformatted in a manner that is appropriate to our cognitive machinery.

The current series of experiments provided little data concerning students' use of worked examples during the homework period. For example, the effectiveness of worked examples could conceivably result from students spending far more time studying worked examples than solving problems during homework. Unfortunately, the use of normal school procedures prevented the gathering of data relevant to such questions. Nevertheless, Cooper and Sweller (1987), Sweller and Cooper (1985), and Tarmizi and Sweller (1988) demonstrated repeatedly over a very large number of laboratory-based experiments that appropriately structured worked examples are processed far more rapidly than conventional problems. If measured, there seems to be no reason to suppose that a different result would have been obtained in the current experiments.

For similar reasons, the current series of experiments did not include direct probes of cognitive processes. The use of intact classes in which experimental manipulations were part of classroom teaching procedures prevented such measures. Other work has used verbal protocols obtained from students studying physics worked examples. Chi, Bassok, Lewis, Reimann, and Glaser (1989) found that more-able students were more likely than less-able students to generate detailed explanations of worked examples and were more aware of their own failures to comprehend. From a cognitive load theory framework, these findings can be explained by assuming that better students have a greater cognitive capacity that permits a more complex manipulation and consideration of a worked example.

It should be noted that the manipulations used in the current series of experiments flowed directly from cognitive load theory, and we believe that the results can be interpreted sensibly only in light of this theory. Our finding that worked examples are effective only if they are presented in an integrated format can be explained readily if one assumes that requiring students to split their attention between multiple sources of information that must be mentally integrated imposes a heavy cognitive load. Consequently, cognitive resources so used are unavailable for schema acquisition and rule automation. Therefore, it may be difficult to find alternative explanations for the current and prior findings.

Our results and associated theorizing also provide an explanation for some of the detailed findings of Zhu and H. A. Simon (1987). These authors found strong advantages for worked examples over direct instruction using algebra. Using geometry, effects appeared to be far smaller. Different conventions in the presentation of worked examples in algebra and geometry may explain these results. Conventional algebra worked examples do not incorporate multiple sources of information that need to be integrated. As a consequence, they are highly effective (see Cooper & Sweller, 1987; Sweller & Cooper, 1985). In contrast, conventional geometry worked examples require students to integrate the diagram with sets of equations and theorems and, therefore, are far less effective. In order to be equally as effective as algebra worked examples, geometry examples must be modified to reduce multiple sources of information (see Tarmizi & Sweller, 1988). The present series of experiments demonstrates graphically the operation of these principles.

The concepts used in this article also can be used to explain the results obtained by Charney and Reder (1986). In teaching the use of a computer application, they compared worked examples with guided solutions and problems. Their results indicated that pure worked examples with no interspersed problems (a technique rarely used) were comparatively ineffective. Interspersed problems were likely to be required for motivational reasons. More important, a group presented with guided solutions and interspersed problems did not perform any better than a group presented with problems alone. This result may be explained by the concepts used in our present article. Worked examples, or any other form of guidance, that is not or cannot be presented using an integrated format probably will not be effective. Charney and Reder's (1986) guided solutions appear not to have been integrated. As indicated previously, it is clear that the important point is not whether worked examples are presented, but rather whether cognitive load is reduced and attention is appropriately directed.

Worked examples are a very common instructional technique. The following points summarize suggestions for their use that flow from the findings reported in this article. First, we have confirmed that practice with worked examples generally is superior to practice with conventional problems. Second, and notwithstanding the first point, not all worked examples are effective. Third, ineffective worked examples can be converted to effective ones by eliminating, as far as possible, the need for students to split their attention among multiple sources of information. Fourth, excessive explanatory material associated with worked examples may be not merely redundant, but detrimental as well. The temptation to provide unnecessary, additional explanations should be resisted, especially if these explanations result in a split-source format. The best worked examples have all information integrated into a single unit.

If these points are valid, then the presentation of technical material requires substantial alteration in a variety of areas. Not only should there be a very heavy emphasis on worked examples, but in many areas, the structure of those examples should also be quite different from the structures currently used. Consideration needs to be given to the integration of unnecessarily separated sources of information. Modifications to present procedures along the lines suggested by the work outlined in this article contain the promise of substantially easing the burden of students in technical areas.

## ACKNOWLEDGMENT

The work reported in this article was supported by a grant (No. A78716120) from the Australian Research Council to John Sweller.

## REFERENCES


Charney, D. H., & Reder, L. M. (1986). Designing interactive tutorials for computer users: Effects of the form and spacing of practice on skill learning. *Human-Computer Interaction, 2*, 297-317.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in problem solving. *Cognitive Science, 13*, 145-182.

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: Lawrence Erlbaum Associates, Inc.

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

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

Larkin, J., McDermott, J., Simon, D., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science, 4*, 317-348.

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology, 77*, 272-284.

Ross, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16*, 371-416.

Ross, B. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 629-639.

Ross, B. (1989a). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456-468.

Ross, B. (1989b). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438-469). Cambridge, England: Cambridge University Press.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 484-494.

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*, 59-89.

Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. *Journal of Experimental Psychology: Learning , Memory, and Cognition, 8*, 463-474.

Sweller, J., Mawer, R., & Howe, W. (1982). Consequences of history-cued and means-ends strategies in problem solving. *American Journal of Psychology, 95*, 455-483.

Sweller, J., Mawer, R., & Ward, M. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 112*, 634-656.

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424-436.

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