Using a General Problem-Solving Strategy to Promote Transfer

Amina Youssef-Shalala  
Australian Catholic University

Paul Ayres  
University of New South Wales

Carina Schubert  
Knowledge Media Research Center, Tübingen, Germany

John Sweller  
University of New South Wales

Cognitive load theory was used to hypothesize that a general problem-solving strategy based on a make-as-many-moves-as-possible heuristic could facilitate problem solutions for transfer problems. In four experiments, school students were required to learn about a topic through practice with a general problem-solving strategy, through a conventional problem-solving strategy or by studying worked examples. In Experiments 1 and 2 using junior high school students learning geometry, low knowledge students in the general problem-solving group scored significantly higher on near or far transfer tests than the conventional problem-solving group. In Experiment 3, an advantage for a general problem-solving group over a group presented worked examples was obtained on far transfer tests using the same curriculum materials, again presented to junior high school students. No differences between conditions were found in Experiments 1, 2, or 3 using test problems similar to the acquisition problems. Experiment 4 used senior high school students studying economics and found the general problem-solving group scored significantly higher than the conventional problem-solving group on both similar and transfer tests. It was concluded that the general problem-solving strategy was helpful for novices, but not for students that had access to domain-specific knowledge.

Keywords: general problem-solving strategies, cognitive load theory, random move heuristics, problem-solving transfer

The search for general problem-solving strategies that transcend specific domains has been a goal of researchers in the field of cognitive processes and instructional design for many decades (Newell & Simon, 1972). The aim has been to find teachable/learnable strategies that students could acquire and then apply to all or most of the subject areas that they encounter. It is a search that generally has been elusive. Attempts to identify such strategies have included investigations into simple heuristics, multistep problem-solving routines, and metacognitive strategies (see Perkins & Salomon, 1989; Schoenfeld, 1982). Many of these studies have lacked replication and have not used randomized, controlled experiments. Furthermore, as Perkins and Salomon (1989) argued, the field has been hampered by trying to identify general cognitive skills without integrating domain-specific knowledge with general strategic knowledge. Despite decades of effort, arguably there are few agreed upon general problem-solving strategies that can be imparted to learners.

Closely aligned to research into general problem-solving strategies have been investigations into transfer. By its very nature, an effective general problem-solving strategy will be usable in different domains, and hence promote transfer. The research literature has provided many examples of failures to find transfer effects (see Perkins & Salomon, 1989), but also some successes (see Bransford & Schwartz, 1999). Analogical reasoning for example has been widely researched, but effective applications are usually associated with domains that share common features (Ross, 1989) and/or where the transfer links are specifically flagged (see Bassok, 1990; Gick & Holyoak, 1980). For transfer to occur, Perkins and Salomon (1989) argued that learners must either have a well-automated skill that has been practiced extensively and that can be applied to a similar situation, or have abstracted a principle that can be applied to a new situation. Consequently, for transfer to occur, considerable prior knowledge needs to be present.

The difficulty in obtaining transfer effects and identifying general problem-solving strategies has led some theorists such as those associated with cognitive load theory (e.g., Paas, Renkl, & Sweller, 2003, 2004; Sweller, Ayres, & Kalyuga, 2011) to assume that teachable general problem-solving strategies will never be found, and to instead, concentrate on teaching specific strategies that apply to particular domains (Tricot & Sweller, 2014). On this assumption, problem-solving skill consists primarily of learning to recognize categories of problems and their appropriate moves. Thus, learners faced with a problem such as, \((a + b)/c = d\), need to solve for \(a\), need to learn that the appropriate initial move for some categories of algebra problems is to multiply out the denominator on both sides of the equation. This problem-solving strategy is highly effective in solving some categories of algebra problems.
but useless in unrelated areas and so is domain specific rather than
general. Problem solving skill is assumed to derive from the
acquisition of large numbers of such domain specific strategies.
Evidence for this assumption flows directly from the pioneering
work of De Groot (1965) who found that the skill of chess masters
derived from a huge knowledge base of chess board configurations
rather than from general problem-solving strategies. Cognitive
load theory has generated many instructional procedures based on
this and other assumptions (Sweller, 2003, 2004, 2011, 2012;
Sweller et al., 2011).

However, through the recent alignment of cognitive load theory
with biological evolutionary principles (Paas & Sweller, 2012;
Sweller, 2011, 2012; Sweller et al., 2011; Sweller & Sweller,
2006) along with experimental evidence (see Youssef, Ayres &
Sweller, 2012), it may be possible to explain the difficulty the field
has had in devising teachable/learnable general problem-solving
strategies. Furthermore, based on our knowledge of human cogni-
tive architecture, it may be possible to devise a teachable general
problem-solving strategy, and identify the conditions, under which
it will be effective. This article investigates such a strategy. We
will begin by considering the cognitive architecture used by cog-
nitive load theory.

Cognitive Load Theory

Cognitive load theory is based on a particular view of human
cognitive architecture and recently has begun to incorporate evo-
lutionary principles into that architecture. Human cognition is
assumed to be a natural information processing system generated
by another natural information processing system, biological evo-
lution. The view that human cognition functions analogously to
evolution by natural selection has a very long history going back
to Darwin (1871/2003). In more recent times, the view has been espoused by Popper (1979). As far as we are aware, its first
appearance in psychology was Campbell’s (1960) important but
neglected paper. There are many ways of specifying the analogy
(e.g., Sweller, 2003) but we will use the five natural information
processing principles described by Sweller and Sweller (2006).

Information Store Principle

Natural information processing systems store massive amounts
of information. A genome provides that information in the case of
biological evolution. Long-term memory has a similar function in
the case of human cognition. The centrality of long-term memory
to problem-solving has been apparent since De Groot (1965) found
that memory for chess board configurations taken from real games
provided the major distinction between more and less expert chess
players. Based on information in long-term memory, a good prob-
lem solver has detailed, domain-specific information concerning a
huge number of problem states and the best moves associated with
each state.

Borrowing and Reorganizing Principle

Most of the information held in an information store is borrowed
from other stores. Asexual and sexual reproduction provides the
machinery in the case of biological evolution. In the case of human
cognition, we obtain information by imitating other people, (Ban-
dura, 1986), listening to what they have to say and reading what
they write. Information obtained in this manner is reorganized
depending on information already held in long-term memory. The
bulk of cognitive load theory has been concerned with techniques
for presenting written and spoken information (Sweller, 2003,
2004, 2011; Sweller et al., 2011) intended to assist in the storage
of domain-specific information in long-term memory.

Randomness as Genesis Principle

Although most information held in an information store is
obtained from other stores, machinery is required to initially gen-
erate information. For biological evolution, random mutation pro-
vides the ultimate source of all biological diversity. Random
generation and test during problem solving provides an identical
function using identical machinery in the case of human cognition.
Random generate and test is assumed by many cognitive load
theory researchers to be the main identifiable source of novelty in
human cognition. This principle is central to human creativity
(Sweller, 2009) and provides the primary theoretical base for the
hypotheses and experiments of this article. In the absence of
sufficient knowledge to guide a solution path, a random generation
and test process is required.

It needs to be emphasized that it is not being argued that novices
who lack knowledge must rely solely on a random generate and
test procedure to generate moves. It is being argued that once
knowledge is exhausted, to make any further moves, there is no
alternative but to randomly choose a possible move and test its
effectiveness. It is for this reason that as far as we are aware, all
problem-solving programs using artificial intelligence have ran-
dom generation as the ultimate breaker of impasses. The experi-
ments of this article test the potential advantages of alerting
learners to using random generate and test through a general
problem-solving strategy.

Narrow Limits of Change Principle

Although knowledge can be used to limit combinatorial explo-
sions, because knowledge frequently is unavailable, alternative
means must be made available to limit the number of combinations
considered when dealing with novelty. Human cognition achieves
this objective via a limited capacity (Miller, 1956), limited dura-
tion (Peterson & Peterson, 1959) working memory. The epigenetic
system plays a similar role in biological evolution.

Environmental Linking and Organizing Principle

This principle provides the purpose and culmination of natural
information processing principles. Information is collected in the
information store to guide action appropriate to a particular envi-
ronment. As was the case for the narrow limits of change principle,
working memory again provides a mediator, in this case dealing
with information from long-term memory used to coordinate ac-
tion appropriate to the environment. Unlike the case of novel
information from the environment, there are no capacity or tem-
poral limits when dealing with familiar information. Unlimited
amounts of information from long-term memory can be used for an
unlimited duration. With knowledge stored in long-term memory,
working memory transforms from a limited capacity, limited du-

ration structure to one with no known capacity or duration limits (Ericsson & Kintsch, 1995). Similarly, in the case of biological evolution, unlimited amounts of genomic information can be used by the epigenetic system to determine phenotypic features.

Cognitive load theory has used versions of this architecture, especially the changing characteristics of working memory as knowledge levels alter, to generate instructional procedures (see Paas, Renkl, & Sweller, 2003, 2004). In line with the above cognitive architecture, most previous cognitive load theory based instructional procedures assumed that the purpose of instruction is to alter the domain-specific contents of long-term memory and that the characteristics of working memory are critical when designing instruction. Furthermore, if the randomness as genesis principle provides the primary procedure of problem solving when dealing with novel problems, it seemed unlikely that teaching general problem-solving skills would prove productive. There seemed little point in teaching learners a random generate and test strategy, a conclusion reached much earlier by Newell and Simon (1972).

The failure to identify teachable/learnable general problem-solving strategies strengthened this assumption. Accordingly, all cognitive load theory derived instructional effects to this point have been based on domain specific knowledge as suggested by the above cognitive architecture. Nevertheless, as indicated in our discussion of the randomness as genesis principle, if teachable/learnable general problem-solving strategies do exist, they should be found in some version of that principle.

Problem Solving and Biologically Primary Knowledge

Geary (2008, 2012) has identified two categories of knowledge based on biological evolution. Our ability to acquire biologically primary knowledge, has evolved over countless generations with examples including learning to listen to and speak one’s own language, and recognizing faces. Biologically primary knowledge is modular with each skill likely to have evolved independently from other primary skills. In contrast, our ability and need to acquire particular examples of biologically secondary knowledge is culturally based. Examples include learning mathematics, reading and the bulk of other topics taught in education and training institutions. Although we have evolved the general ability to acquire biologically secondary skills, unlike biologically primary knowledge, we have not evolved to acquire any particular secondary skill. The machinery we use to acquire the knowledge of two different subjects at school is likely to be identical. In contrast, the machinery used to learn to speak a native language and the machinery used to recognize faces is likely to be very different.

The previously described cognitive architecture applies to secondary, not primary knowledge. From a working memory perspective, primary knowledge occurs effortlessly without explicit instruction, because humans over a long period of time have developed a need for such knowledge. Secondary knowledge, on the other hand, does not have the same evolutionary advantage and requires significantly more effort to acquire. This type of knowledge is most commonly acquired through the borrowing and reorganizing principle, and requires explicit instruction. Most of the research into cognitive load theory has focused on secondary knowledge (Sweller et al., 2011).

Although the above cognitive architecture indicates how humans acquire, process and use biologically secondary information, the principles themselves constitute biologically primary processes and so do not need to be taught because they are learned automatically, without instruction. The current work is based on the assumption that although we do not have to be taught how to randomly generate information because random generate and test is biologically primary, we do need to be taught which biologically secondary problems lend themselves to solution using the randomness as genesis principle as a problem-solving device. The current experiments are concerned with testing the effect of informing learners to use random generate and test as a general problem-solving strategy, to solve novel problems. As indicated above, random generate and test is likely to be a biologically primary skill that should not need to be taught, based on the assumption that it is acquired automatically without explicit instruction. Accordingly, in the experiments below, we did not teach learners how to engage in random generation of problem-solving moves. Rather, with a single statement, we alerted them that for the (biologically secondary) problems with which they were faced, a familiar, make as many moves strategy, including random generation when necessary was likely to be effective. It was hypothesized that alerting learners to the fact that a random generate and test strategy embedded in a general heuristic could be used to solve the problem being presented would facilitate problem-solving transfer to novel problems.

Goal-Free Strategies and the Random Generation Heuristic

Alerting learners that a random generation and test strategy might be effective in solving the problems with which they are faced bears a resemblance to a goal-free strategy. This strategy was the first generated using cognitive load theory. It is an alternative to learning from conventional, means-ends problem-solving (see Sweller, 1988; Sweller et al., 2011). Substantial research has shown that learning through conventional problem solving is detrimental because novices will have to rely on a means-ends strategy (Sweller et al., 2011). This strategy requires problem solvers to simultaneously consider their current problem state and the goal state, extract differences between the two states and find a problem-solving operator to reduce those differences creating an extraneous cognitive load (see Sweller, van Merrienboer, & Paas, 1998). A problem may be solved using means-ends analysis but learning is compromised because working memory resources are devoted to solving the problem rather than to key relations that need to be acquired for knowledge formation. The goal-free strategy was based on the assumption that if the goal was removed, then learners could not use means-ends analysis.

A goal-free strategy occurs when a conventional problem with a specific goal is replaced by a problem with a nonspecific goal. As well as removing the goal, learners were required to adopt the general heuristic—“find as many unknowns as you can.” Tailored to meet the features of the domain, learners have been asked to find as many distance or velocity variables in physics problems (Sweller, Mawer, & Ward, 1983), distances in trigonometry problems (Owen & Sweller, 1985), and angles in geometry problems (Ayres, 1993). In all the goal-free studies, this strategy has been shown to be superior to a conventional problem-solving approach, where problem solvers were not given any problem-solving guidance. However, as Youssef et al. (2012) pointed out, a goal-free
strategy not only removes the goal, but also introduces a general problem-solving heuristic. This heuristic has not been directly specified in any of the goal-free studies because it was assumed that students already knew how to use it. It can also be considered as a random generate and test heuristic outlined above.

The evidence suggests that problem solvers could use this strategy within the given context without direction, and it was very helpful in acquiring new knowledge. From this perspective, a primary skill is being used to help build secondary knowledge. A great advantage in tapping into a primary knowledge skill is that it requires few working memory resources, and hence unhelpful extraneous cognitive load is reduced, whereas more resources are available for learning (germane cognitive load). It is important to investigate whether this heuristic can be used independently of a goal-free environment. In other words, does asking learners to make as many moves as possible facilitate problem solving in contexts other than a goal-free context?

**Introduction to the Present Study**

This study tests the general hypothesis, based on the randomness as a principle, that when novice learners are guided to use a random generate and test heuristic, problem-solving is enhanced. However, it is possible that many learners do not automatically use this strategy when faced with biologically secondary tasks. It may be necessary for instructional designers to explicitly indicate such connections. This study sought to investigate this hypothesis.

It was predicted that for novices, with little domain-specific knowledge, a random generate and test problem-solving strategy embedded in a general heuristic will substitute for missing domain knowledge and enable solutions to be generated. However, for learners who have relevant domain-specific knowledge, it was expected that such a strategy would be redundant, as prior knowledge would guide problem-solving strategies without any need for the heuristic. We assume that domain-specific prior knowledge is a far more effective problem-solving device than a domain-general strategy. A domain-general strategy is only likely to be effective where domain-specific knowledge is missing. There may be a window of opportunity for the strategy to work, before domain-specific expertise develops. Furthermore, it was also expected that given sufficient practice on the general problem-solving strategy in a specific domain, faced with novel transfer problems for which domain-specific knowledge was absent, the general problem-solving strategy would be used. Two specific predictions flow from this argument. It was hypothesized that

**Hypothesis 1:** For novices in the domain, the general problem-solving strategy would result in more problem solutions than a conventional, means-ends problem-solving strategy

**Hypothesis 2:** For more expert learners in the domain the general problem-solving strategy would be no more effective, or even less effective than a conventional means-ends problem-solving strategy

In this study, two variations of the general problem solving strategy were investigated in a mathematical and economics learning domain.

**Experiment 1**

The aim of Experiment 1 was to test the broad hypotheses advanced above, by comparing a general problem-solving strategy with conventional problem solving during learning. The curriculum domain was geometry. The overall design of the experiment consisted of an acquisition phase, where students solved basic geometry problems followed by two tests (similar and transfer). The similar test contained problems almost identical to those practiced during acquisition, while those in the transfer test required use of the same geometry theorems but in quite different problem settings. It was predicted that the general problem-solving strategy would most likely be effective on the transfer problems only, where less domain-specific knowledge was present. For acquisition and similar test problems, it was expected that more prior knowledge would be available. However, it was also expected that there would be individual differences between participants, which potentially could interact with performance on any of the tasks. To investigate such prior knowledge effects, students were classified as having high or low prior knowledge based on whether they were in a high or low ranked mathematics class.

**Method**

**Participants.** Fifty-four Grade 9 (mean age of 14.8 years) students from two Sydney high schools participated in the study. In both schools, students had been placed in streamed mathematics classes according to general mathematical ability as measured by school-based assessments. To ensure that students had sufficient but not excessive prior knowledge to be matched appropriately with the tasks, participants were chosen from classes ranked 2 and 3 (from a total of 4) from one school, and classes ranked 3 and 4 (from a total of 6) from the second school. To create two groups with different mathematical abilities, the students from the highest ranked class in each school were combined and called the Higher Ability group. Similarly, the students from the lowest ranked class in each school were combined and called the Lower Ability group. Each student was randomly assigned to one of two experimental conditions, conventional problem-solving and general problem-solving before the study, creating a 2 (ability) × 2 (strategy) design. For the Higher Ability group, there were 16 students in the conventional problem-solving group and 14 in the general problem-solving group. For the Lower Ability group, there were 12 students in the conventional problem-solving group and 12 in the general problem-solving group. Group numbers varied according to the attendance of student volunteers on the day of testing. All students were familiar with the basic mathematical content (geometry theorems) used in the study, having previously studied this content at school in their regular classes, but were not expected to be familiar with many of the test problems, particularly those used to test transfer.

**Materials and Procedure**

**Problem tasks.** There were three sets of problems, one set for each of a learning phase (acquisition), a test of similar geometry problems (similar), and a test of unfamiliar geometry problems (transfer). The acquisition materials consisted of 16 problems presented in diagrammatic form, where each problem required
knowledge of at least two elementary geometry theorems taken from a set of eight to find a solution. The set of theorems contained three parallel-line properties (alternate, co-interior, and corresponding), three straight-line properties (adjacent, reflex, and 180°-straight line), and the angle sum of a triangle and quadrilateral. The two groups received identical problems, but with one important difference. To facilitate different problem-solving strategies, the instructions given differed accordingly. For the conventional problem-solving group, a conventional goal-specified instruction was given—“For each question find the value of angle x.” For the general problem-solving group the instruction was—“For each question find as many angles as you can in any order you like.” The reference “in any order you like” was included to allow students to start generating answers without an initial focus on a specific goal. It is important to highlight that the conventional problem-solving and the general problem-solving groups, both had an angle specified on the diagrams as “x.” It was assumed that all learners understood the goal to be “x” as they had previously learnt this in their geometry classes.

**Acquisition phase.** The acquisition set contained 14 problems that could be solved in a minimum of two steps (two consecutive calculations using different geometry theorems). In the case of the problem (conventional problem-solving format) depicted in Figure 1a, angle BCD must be calculated first (subgoal) before the goal (“x”) can be calculated. The remaining two problems could be solved in three steps, consisting of two subgoals and a goal (“x”). The full set of problems was printed back-to-back on A4 sheets of paper with each page containing four problems and the written instruction printed at the top of each page. Sufficient space was provided to enable all answers to be written on the paper. It should be noted that no direct instructions on how to solve these problems were given. Any further learning was assumed to have taken place by practice gained through the problem-solving activities on this set of problems. No feedback was given. For the conventional problem-solving group, a minimum of 34 angles had to be found to successfully complete the problem set. It was expected that the general problem-solving group would find more angles. For example, for the problem depicted in Figure 1a, a total of five angles could be calculated (angles BCD, CDF, CDG, GDE, and FDE), whereas only two are required for the solution (angles BCD and FDE). Overall, the general problem-solving group could calculate as many as 54 angles, in the absence of any additional constructions. In the postacquisition phases, identical problems were given to both groups. The problems were presented in a conventional problem-solving format (find “x”).

**Similar test phase.** The first test phase (similar problems) contained eight problems identical to those in the learning phase, except the size of the angles was changed. Students could complete seven of these problems in a minimum of two steps and the remaining problem in three steps.

**Transfer test phase.** In the transfer phase (eight problems), the first four problems were similar to those in the learning phase except while students had to apply the same combinations of theorems, the order in which the angles could be calculated by the conventional problem-solving group, were reversed. The remaining four problems were constructed from very differently configured diagrams to those shown previously in acquisition and similar test phases, although similar combinations of the same geometry theorems were required (see Figure 1b). In a previous study, Ayres and Sweller (1990) found that students found it difficult to solve problems with unusual configurations even though they could solve equivalent problems when the diagrammatic form was familiar. Unfamiliar configurations led to more solution path searches and raised cognitive load.

Overall, students could solve three of the problems in a minimum of two steps, and the remaining five problems in three steps. Both sets of test materials were printed back-to-back on A4 sheets of paper. For both problem sets, each page showed 4 problems and the instruction was printed on top of each page. For all four groups, the instruction read “for each question find the value of angle x.”

The experiment was conducted in the students’ regular mathematics classes, under strict supervision. A brief introduction concerning the format of the study was given, and students were told that it was a test of their geometry abilities. They were also instructed that all their answers should be written on the question papers provided. The whole experiment consisted of three phases: learning phase (15 min), similar test phase (10 min), and transfer test (10 min). For each phase, the materials were distributed and then collected after the given time had expired.

**Scoring of tasks.** As the conventional problem-solving group was asked to find a specific goal (x) during acquisition, it was possible that only the required goal and subgoal(s) would be calculated. In contrast, the general problem-solving group was expected to find these angles as well as other angles. Therefore, only performance on the critical angles (goal and subgoals) was

---

**Figure 1.** (a) Example of a two-step geometry problem. (b) Example of unfamiliar geometry problem.
used to compare the two groups during acquisition. For two-step problems, if the nominated goal angle was calculated correctly then two marks were assigned. If the required subgoal was calculated correctly, but the goal incorrectly, then one mark was assigned. One mark was also awarded for an incorrect subgoal calculation but correct goal solution. All other incorrect responses were awarded no marks. For the three-step problems, two marks were assigned if the nominated goal was calculated correctly. One mark only was assigned if both subgoals were calculated correctly but the goal incorrectly. All other incorrect responses earned no marks.

For the test sets, where both groups were required to find a goal, an identical scoring pattern to the acquisition phase was used. Consequently, maximum scores for the three tests were 32 (acquisition), 16 (similar), and 16 marks (transfer). Two raters scored each test and any discrepancy was resolved giving 100% agreement. This strategy to achieve interrater reliability was adopted in all experiments in the study. As it was expected that the general problem-solving group would find more angles than the conventional problem-solving group, the number of angles calculated by each student was counted on each of the three tasks involving the geometry problems.

Results and Discussion

Performance scores. Mean scores per group on each of the three performance tests are shown in Table 1. On each of the measures a 2 × 2 analysis of variance (ANOVA) was conducted.

Acquisition scores. There was an ability main effect showing that the Higher Ability group performed significantly better than the Lower Ability group: F(1, 50) = 51.1, MSE = 38.5, p < .001, ηp2 = 0.51. There was no significant main effect for strategy, or an interaction (both F < 1, ns).

Test scores. There was a significant ability effect, showing that the Higher Ability group performed significantly better than the Lower Ability group: F(1, 50) = 12.5, MSE = 8.3, p = .001, ηp2 = 0.20. There was no significant effect for strategy or an interaction (both F < 1, ns). It is notable that the combined group success rates were 90% (Higher Ability group) and 73% (Lower Ability group), which may suggest that this test was fairly easy and too similar to the acquisition problems to elicit significant strategy effects.

Transfer scores. There was a significant ability effect, showing that the Higher Ability group performed significantly better than the Lower Ability group: F(1, 50) = 36.3, MSE = 10.1, p < .001, ηp2 = 0.42. There was a significant effect for strategy, F(1, 50) = 7.4, p < .01, ηp2 = 0.13, where the general problem-solving group (M = 11.5, SD = 4.26) scored higher than the conventional problem-solving group (M = 9.86, SD = 4.55). There was a marginally significant interaction, F(1, 50) = 2.69, p = .08, ηp2 = 0.06. As it was predicted that students with different levels of prior knowledge would perform differently according to the acquisition strategy followed, simple effects tests were completed. For the Higher Ability group, there was no significant difference between the conventional problem-solving and general problem-solving groups (t < 1). However, for the Lower Ability group, the general problem-solving group outperformed the conventional problem-solving group, t(22) = 2.0, p = .05, Cohen’s d = 0.79.

These performance test results indicated, as would be expected, that the Higher Ability group outperformed the Lower Ability group on each test, with very large effect sizes indicating that the method adopted to create groups with differing levels of expertise was effective. By examining the scores for the conventional problem-solving groups only, who received no direct instructional guidance other than problem-solving practice, an indication of prior knowledge on this topic was gained. Mean scores for the Higher Ability group were 15.4 (88.1% correct) and 13.9 (83.1%) for the similar and transfer tests, respectively, suggesting that this cohort had a very high level of prior knowledge for these problems, and may explain why students in the Higher Ability groups were not affected by the different strategies. Whereas the Lower Ability conventional problem-solving group had mean scores of 11.4 (71.3% correct) and 6.5 (40.6%) for the similar and transfer tests, respectively, indicating lower levels of prior knowledge, particularly for the transfer test. There was an overall strategy effect in favor of the general problem-solving strategy on the transfer test, which was greatly influenced by performance of the Lower Ability group.

Angles Calculated

The means and SDs for the total number of angles calculated per group are shown in Table 2. Individual 2 × 2 ANOVAs were conducted on each data set.

Acquisition phase. During the learning phase, the Higher Ability group found significantly more angles than the Lower Ability group: F(1, 50) = 10.5, MSE = 163.6, p < .001, ηp2 = 0.43. There were also a significant strategy effect, F(1, 50) = 10.5, p < .01, ηp2 = 0.17. The general problem-solving group calculated significantly more angles than the conventional problem-solving group.

Test phase. The Higher Ability group calculated significantly more angles than the Lower Ability group: F(1, 50) = 4.41,

<table>
<thead>
<tr>
<th>Ability</th>
<th>Strategy</th>
<th>Acquisition (max = 32)</th>
<th>Similar test (max = 16)</th>
<th>Transfer test (max = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Conventional problem-solving</td>
<td>28.1 (3.8)</td>
<td>14.8 (1.7)</td>
<td>13.3 (2.5)</td>
</tr>
<tr>
<td></td>
<td>General problem-solving</td>
<td>26.9 (5.2)</td>
<td>14.1 (2.3)</td>
<td>14.1 (2.3)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>27.5 (4.5)</td>
<td>14.4 (2.0)</td>
<td>13.7 (2.3)</td>
</tr>
<tr>
<td>Low</td>
<td>Conventional problem-solving</td>
<td>14.5 (7.6)</td>
<td>11.4 (3.7)</td>
<td>6.5 (3.6)</td>
</tr>
<tr>
<td></td>
<td>General problem-solving</td>
<td>16.2 (8.1)</td>
<td>11.8 (3.7)</td>
<td>10.4 (4.4)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>15.3 (7.7)</td>
<td>11.6 (3.6)</td>
<td>8.4 (4.4)</td>
</tr>
</tbody>
</table>
was a significant interaction, effect for strategy or an interaction (both a far-transfer test, which featured problems combining geometric acquisition, test and transfer phases. Following these three phases, specific hypotheses of Experiment 1 were again tested in the current experiment, which involved the same problem tasks for two experimental conditions conventional problem-solving and general problem-solving before the study commencing, creating a 2 (ability) × 2 (strategy) design. For the Higher Ability group, there were 13 students in the conventional problem-solving group and 14 in the general problem-solving group. For the Lower Ability group there were 9 students in the conventional problem-solving group and 11 in the general problem-solving group.

Experiment 2

The aim of Experiment 2 was to build on the findings of Experiment 1 by introducing a far-transfer test phase. The two specific hypotheses of Experiment 1 were again tested in the current experiment, which involved the same problem tasks for acquisition, test and transfer phases. Following these three phases, a far-transfer test, which featured problems combining geometric theorems with applications of Pythagoras’ Theorem and the properties of plane shapes, along with algebraic word and sequence problems was presented. Based on the results of Experiment 1, it was predicted that the general problem-solving strategy would be effective on the transfer and far-transfer problems only, where learners were likely to have acquired less domain-specific knowledge.

Method

Participants. Forty-three Year 9 (mean age of 14.2 years) students from one Sydney high school for girls participated in the study. Students were invited to participate from the top three (of four) mathematics classes of the school. To create two groups with different mathematical abilities, school-based tests completed at the end of Year 8 were used to find a median split. Students on or above the median score were classified as Higher Mathematical Ability, and students below the median were classified as Lower Mathematical Ability. Each student was randomly assigned to one of two experimental conditions conventional problem-solving and general problem-solving before the study commencing, creating a 2 (ability) × 2 (strategy) design. For the Higher Ability group, there were 13 students in the conventional problem-solving group and 14 in the general problem-solving group. For the Lower Ability group there were 9 students in the conventional problem-solving group and 11 in the general problem-solving group.

Materials and Procedure

Problem tasks and materials. There were four phases (acquisition, similar test, transfer test, and far-transfer test) in this experiment requiring participants to complete a set of problems in each phase. The same set of problems used in the acquisition, test, and transfer phases of Experiment 1 was used in the current experiment. All of the acquisition, test, and transfer phase procedures in this experiment were identical to the procedures of Experiment 1.

In the far-transfer phase (seven problems, see Appendix), the problems were designed to determine whether students could transfer their general problem-solving strategy to different domains. A number of multistep problems were constructed using Pythagoras’ Theorem (Question 1), a combination of Pythagoras’ Theorem and equilateral triangle properties (Question 2), a combination of Pythagoras’ Theorem and parallelogram properties (Question 3), circle geometry properties (Question 4), solving three simultaneous equations by substitution (Question 5), a word problem (Question 6), and a numerical sequence problem (Question 7).
The far-transfer test materials were printed back-to-back on A4 sheets of paper. Each page showed two problems on each side apart from the last page, which had one problem. Instructions were printed at the top of each question. In all of the questions used in the far-transfer test, students were expected to know the basic theorems and properties underlying the problems. The challenge was to know where to start and find a successful path through the problem space.

**Procedure.** The experimental procedures were identical to Experiment 1 except that there was an additional post acquisition far-transfer test. The experiment ran for a total of 50 min in a single school period. Time allotted to each phase was 15 min (acquisition), 10 min (similar test), 10 min (transfer test), and 15 min (far-transfer test).

**Scoring of tasks.** The same scoring method used in the acquisition, test, and transfer test phases of Experiment 1 was used in the current experiment. For the far-transfer phase all the problems could be solved in a minimum of three steps. For each correct answer three marks were awarded, making a maximum score of 18 for the set. For incorrect answers, one mark was awarded for one incorrect subgoal, or equivalent working, and two marks for two correct subgoals, or equivalent working. If no correct steps or equivalent working was found, 0 was awarded. Hence scores of 0, 1, 2, or 3 were possible for each problem.

**Results and Discussion**

**Performance scores.** Mean scores per group on each of the four performance tests are shown in Table 3. On each of the measures a $2 \times 2$ ANOVA was conducted.

**Acquisition, similar test and transfer test scores.** There was a significant ability effect, showing that the Higher Ability group performed significantly better than the Lower Ability group on acquisition, $F(1, 43) = 46.8, MSe = 48.3, p < .001, \eta^2_p = .52$; similar test, $F(1, 43) = 39.7, MSe = 820, p < .001, \eta^2_p = .48$; and transfer test, $F(1, 43) = 47.8, MSe = 14.2, p < .001, \eta^2_p = .49$. There was no significant effect for strategy or an interaction on the acquisition, similar, and transfer tests (all $F < 1, ns$).

**Far-transfer test scores.** There was a significant ability effect, showing that the Higher Ability group performed significantly better than the Lower Ability group: $F(1, 43) = 10.4, MSe = 5.10, p < .01, \eta^2_p = .19$, but no significant effect for strategy ($F < 1 ns$). However, there was a significant interaction, $F(1, 43) = 10.9, p < .01, \eta^2_p = .20$. Simple effects tests revealed that for high ability students the conventional problem-solving group outperformed the general problem-solving group, $t(25) = 2.56, p < .05$, Cohen’s $d = .89$. For the lower ability students the reverse effect was found as the general problem-solving group outperformed the conventional problem-solving group, $t(18) = 2.15, p < .05, Cohen’s d = .90$.

Overall, as would be expected, the Higher Ability group outperformed the Lower Ability group on each test. This finding suggests that the method adopted to create groups with differing levels of expertise was successful. An indication of prior knowledge was gained by examining the scores for the conventional problem-solving groups only, who received no instructional guidance other than problem-solving practice. Mean scores for the Higher Ability group were 15.4 (88.1% correct), 13.9 (83.1%), and 7.6 (36.25%) for the similar, transfer, and far-transfer tests, respectively, suggesting that this cohort had a very high level of prior knowledge for the geometry problems, but not the far-transfer problems. In contrast, the Lower Ability conventional problem-solving group had mean scores of 10.0 (62.5% correct), 8.3 (51.9%), and 3.2 (15.2%) for the similar, transfer, and far-transfer tests, respectively, indicating that this cohort had lower levels of prior knowledge than the Higher Ability group on all three tasks.

Clearly, the far-transfer tasks were found to be difficult by both ability groupings. It was only on this task that a strategy difference occurred. In a classic, disordinal interaction, the general problem-solving strategy was more helpful for the lower ability students but less helpful for the higher ability students, demonstrating an expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003).

**Angles Calculated**

The means and SDs for the total number of angles calculated per group on the first three pure geometry tasks are shown in Table 4. Individual $2 \times 2$ ANOVAs were conducted on each data set. This information was not collected on the far-transfer tasks, as they were not all geometric in nature.

**Acquisition phase.** During the learning phase, the Higher Ability group found significantly more angles than the Lower Ability group, $F(1, 43) = 13.0, MSe = 230.7, p < .001, \eta^2_p = .23$. There was also a significant strategy effect, $F(1, 43) = 6.37, p < .05, \eta^2_p = .13$. The general problem-solving groups calculated significantly more angles than the conventional problem-solving groups. There was no interaction ($F < 1, ns$).

**Similar test phase.** The Higher Ability group calculated significantly more angles than the Lower Ability group, $F(1, 43) = 11.7, MSe = 84.4, p = .001, \eta^2_p = .21$. However, there was no significant effect for strategy, $F(1, 43) = 1.25, p > .10$, nor an interaction ($F < 1, ns$).

<table>
<thead>
<tr>
<th>Ability</th>
<th>Strategy</th>
<th>Acquisition (max = 32)</th>
<th>Similar test (max = 16)</th>
<th>Transfer test (max = 16)</th>
<th>Far-transfer test (max = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Conventional problem-solving</td>
<td>26.1 (6.2)</td>
<td>15.4 (1.0)</td>
<td>13.9 (1.9)</td>
<td>7.6 (2.2)</td>
</tr>
<tr>
<td></td>
<td>General problem-solving</td>
<td>26.8 (4.6)</td>
<td>15.3 (1.0)</td>
<td>13.9 (2.3)</td>
<td>5.4 (2.3)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>26.5 (5.3)</td>
<td>15.3 (1.0)</td>
<td>13.9 (2.1)</td>
<td>6.4 (2.5)</td>
</tr>
<tr>
<td>Low</td>
<td>Conventional problem-solving</td>
<td>14.1 (11.0)</td>
<td>10.0 (6.2)</td>
<td>8.3 (4.9)</td>
<td>3.2 (2.5)</td>
</tr>
<tr>
<td></td>
<td>General problem-solving</td>
<td>10.6 (6.1)</td>
<td>10.0 (3.4)</td>
<td>7.8 (3.2)</td>
<td>5.4 (2.0)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>12.2 (8.6)</td>
<td>10.0 (4.2)</td>
<td>8.0 (3.9)</td>
<td>4.4 (2.5)</td>
</tr>
</tbody>
</table>
Transfer test phase. The Higher Ability group calculated significantly more angles than Lower Ability group, $F(1, 43) = 7.63, MSe = 45.7, p = .001, \eta^2_p = .15$. There was no significant effect for strategy, $F(1, 43) = 1.97, p > .10$, $\eta^2_p = .04$, nor an interaction ($F < 1, n.s.$).

Overall, the Higher Ability group found significantly more angles than the Lower Ability group on all three tasks. Furthermore, consistent with the different instructions, the general problem-solving group found more angles than the conventional problem-solving group during the acquisition task. No significant differences were found on either of the test tasks, indicating that the strategy to calculate more angles demonstrated during acquisition by the general problem-solving groups did not transfer across tasks to the same extent.

Two hypotheses were tested in this experiment. The first hypothesis predicted that the general problem-solving strategy would be more effective compared with a conventional problem-solving strategy on problems where students had low levels of prior knowledge. This hypothesis was supported for far-transfer only, where an interaction was found. Students with the lowest level of mathematical knowledge benefited from the strategy, whereas students with higher prior knowledge were disadvantaged by the general problem-solving strategy. It is feasible that on these tasks, the students classified as high ability students possessed domain-specific knowledge that they could use with the general problem-solving strategy merely serving to interfere with the use of that knowledge.

On the similar and transfer tests, no group differences were found in support of Hypothesis 2. Scores by the conventional problems solving groups suggested that students had a high degree of prior knowledge, especially in the Higher Ability group. If prior knowledge reduced the need for a general problem-solving strategy, then no significant differences would be expected between the two strategies on these tasks.

Further insights were gained into the extent that the general problem-solving strategy was applied by examining the number of angles calculated during completion of the various tasks. As expected, consistent with how they were instructed, more angles were calculated by the general problem-solving groups during the acquisition tasks.

Overall this experiment showed that the general problem-solving strategy could be extended to far-transfer problems for students with low levels of prior knowledge. In comparison to the Experiment 1 findings, a significant effect was not found on transfer geometry problems in this experiment, and therefore, the main finding of the earlier study was not replicated. One potential explanation for this variation may be a difference in prior knowledge between the student samples. The Lower Ability group of students in the present study, who had conventional problem-solving training, had a higher score (51.9%) than the equivalent group in Experiment 1 (40.6%). This difference in prior knowledge may have reduced the impact of the general problem-solving strategy.

### Experiment 3

Evidence from far-transfer tasks in the second experiment suggested that students with low prior knowledge could benefit from a general problem-solving approach. In contrast, students with higher levels of prior knowledge did not find the general problem-solving approach useful. The most likely explanation is in terms of differential knowledge. General problem-solving strategies such as the heuristic used in the earlier experiments may be relatively effective only in the absence of relevant domain-specific knowledge. Learning to solve problems means acquiring the domain-specific knowledge required for particular classes of problems. General problem-solving techniques are only likely to be used in the absence of relevant domain-specific knowledge.

Experiment 3 explores this hypothesis further by adding a worked examples contrast. Worked examples have been shown to be a very effective way of generating domain specific knowledge, especially for novice learners (see Sweller et al., 2011). They have also been shown to facilitate transfer effects in a mathematical domain after extended practice (see Cooper & Sweller, 1987). However, no research has examined their ability to generate far-transfer such as that illustrated in Experiment 2. The main aim of this experiment was to investigate whether worked examples could induce transfer as effectively as the given general problem-solving strategy.

Because previous research has shown that worked examples are very effective in generating domain-specific knowledge and transfer in that domain, two predictions were made.

**Hypothesis 3:** On tasks most closely related to those practiced during acquisition, the worked example strategy will be superior to a general problem-solving strategy

**Hypothesis 4:** On transfer tasks outside the domain, the general problem-solving strategy will be more effective than a worked-example approach.

In addition, because of the impact of expertise in the domain, it was expected that these effects would be moderated by prior knowledge. In the previous experiments, high and low ability groups were created by using school-based tests of general math-
emathematical ability. Although this method was found to be effective in differentiating between groups, a more precise measurement was adopted in this experiment, by pretesting all participants on their prior knowledge of basic geometry, the main learning domain of the experiment. As a result of collecting a continuous variable measure of prior knowledge, regression analysis of the data was conducted.

Method

Participants. Forty Grade 9 students (mean age of 14.4 years) from two Sydney high schools participated in the study (19 females and 21 males). Each student was randomly assigned to a group representing one of two experimental conditions: the worked examples group ($N = 21$) or the general problem-solving group ($N = 19$).

Materials and Procedure

Problem tasks and materials. There were five phases (prior knowledge test, acquisition, similar test, transfer test, and far-transfer test) in this experiment. To test prior knowledge in the geometry domain, a set of eight problems similar to those used in the similar test phase of Experiments 1 and 2 was constructed. The problems were the same diagrammatically as those used before; the only difference was the angle sizes and the order in which they had to be solved.

In the acquisition phase, the two groups followed different learning strategies. To ensure a controlled experiment each group received the same problem set of 18 geometry problems. To comply with the "study one–solve one" alternation strategy used in the previous experiments, the two problems in each pair were identical in structure, the only difference being that the general problem-solving group students were told—“Find as many angles during acquisition.”

Written instructions informed students of the procedure for each group. For the worked example group, on the first problem in each pair, it was stated—“Study the solution to this problem,” which required angle x to be found. On the second problem students were simply told—“Find the value of angle x.” For the general problem-solving group students were told—“Find as many angles as you can in any order you like,” as outlined in the previous experiments.

The nine problem pairs were constructed from the acquisition materials used in the previous experiment. The two problems in each pair were identical in structure, the only difference being that the angles were changed. Of the nine problem sets to be solved by the general problem-solving group, which was asked to calculate more angles, could find as many as 64 angles on the corresponding set of nine problems. Hence, it was expected as in the previous experiments that the general problem-solving group would find many more angles during acquisition.

The set of problems was printed back-to-back on A4 sheets of paper with each page containing two problem pairs. All other acquisition phase procedures, including the awarding of marks was identical to the previous experiment. However, a maximum score of 18 (nine problems worth two marks each) could be achieved.

For the similar test phase following the acquisition phase, the same set of problems as used in the prior knowledge test was given to the two groups again. For the transfer phase, a similar set of problems to the transfer test problems of Experiment 2 were used in the current experiment. The eight problems were the same diagrammatically to those used previously with the only difference being in the angle sizes and the order in which the problems had to be solved.

The problem set for the far-transfer phase was different from the far-transfer set used in Experiment 2, as only algebra related problems were used. Two questions were based on the substitution problem (Question 5 in Experiment 2), three questions on the word problems (Question 6 in Experiment 2), and one number sequence problem (Question 7 in Experiment 2). All of the problems could be solved in a minimum of three steps. For each correct answer three marks were awarded, making a maximum score of 18 for the set. For incorrect answers, one mark was awarded for one correct subgoal, or equivalent working, and two marks for two correct subgoals, or equivalent working. Hence on each problem, scores of 0, 1, 2, or 3 were possible. The set of problems were printed back-to-back on A4 sheets of paper. Each page contained two problems.

Procedure

The experimental procedures were identical to Experiment 2 except that there was an additional prior knowledge test. The experiment ran for a total of 65 min in a single school period. Time allotted to each phase was 15 min (prior knowledge test), 15 min (acquisition), 10 min (similar test), 10 min (transfer test), and 15 min (far-transfer test).

Results and Discussion

Analysis of data. Stepwise linear regression was used to analyze the data for each dependent variable measured in the study. Data was entered into the regression program in the following way. For Step 1 (Model 1), the prior knowledge variable, centered by subtracting the overall mean from each raw score as recommended by West, Aiken, and Krull (1996), was entered first as it was expected to be a significant predictor of performance. For Step 2 (Model 2), the strategy variable was entered next as it was highly correlated with test data. Finally, the Prior Knowledge × Strategy variable was entered last in Step 3 (Model 3) as it was poorly correlated with test data. In the following analyses, $R^2$ represents the explained variance by Step 1, $\Delta R^2$ represents the change in variance on subsequent steps, and $\beta$ represents the standardized coefficients. Only significant models that include the most predictors are reported.

Performance test scores. Mean group scores are shown for all tests in Table 5. For acquisition scores, Model 2 was significant, $F(2, 37) = 9.80, p < .001$, with prior knowledge a close to significant predictor ($R^2 = 0.063, \beta = .26, p = .055$). Strategy was a significant predictor ($\Delta R^2 = 0.28, \beta = .53, p < .001$) where the worked example strategy (estimated $M = 14.0, SE = 0.83$) led to higher scores than the general problem-solving strategy (estimated $M = 9.3, SE = 0.87$). On similar test scores, Model 1 was significant, $F(1, 38) = 4.50, p < .05$ with prior knowledge the only significant predictor ($R^2 = 0.11, \beta = .33, p < .05$). Similarly, on
the transfer test scores, Model 1 was significant, $F(1, 38) = 17.1, p < .001$ with prior knowledge the only significant predictor ($R^2 = 0.31, \beta = .56, p < .001$). In contrast, for far-transfer test scores, Model 2 was significant, $F(2, 37) = 4.19, p < .05$ with strategy the only significant predictor ($\Delta R^2 = 0.11, \beta = -.33, p < .05$) where the general problem-solving strategy (estimated $M = 11.7, SE = 0.81$) led to higher scores than the worked example strategy (estimated $M = 9.2, SE = 0.77$).

**Angles calculated.** The mean number of angles calculated in the geometry tests are shown in Table 6. For acquisition angles, Model 2 was significant, $F(2, 37) = 3.43, p < .05$ with prior knowledge just reaching significance ($R^2 = 0.06, \beta = .31, p = .05$), but not strategy ($\Delta R^2 = 0.06, \beta = -.24, p = .12$). For similar and transfer angles, no model was significant.

Two hypotheses were tested in this experiment. The first hypotheses (Hypothesis 3) predicted that on tasks most closely related to those practiced during acquisition, the worked example strategy would be superior to a general problem-solving strategy. Although, on the similar and transfer tasks, the worked example group had higher mean scores than the general problem-solving group, these differences were not significant. The only significant predictor for these two tasks was prior knowledge. However, on the acquisition problem the worked example group scored significantly higher than the general problem-solving group, providing some support for the hypothesis. It is notable that the overall mean (percent correct) score for the prior knowledge test was 66.9%, suggesting that most students had a considerable prior knowledge in the geometry domain. The advantage gained during acquisition for the worked example group dissipated after this short practice period. As no differences were found between groups, this also suggests that learners may have continued to use their domain-specific knowledge rather than rely on the general problem-solving strategy. However, on the far-transfer problems where the geometry domain-specific knowledge was no longer relevant, the general problem-solving strategy was more effective than a worked-example approach, providing support for the second hypothesis (Hypothesis 4).

### Experiment 4

The result of the previous three experiments indicated that a general problem-solving strategy was effective in promoting mathematical transfer performance compared with a conventional problem-solving strategy (Experiments 1 and 2) and a worked examples strategy (Experiment 3). The main aim of Experiment 4 was to extend the investigation of a general problem-solving strategy to a different learning domain. The domain chosen was the economics topic *Supply and Demand*, which required learners to construct graphical representations from several central economic concepts. Again, as in the previous experiments, it was hypothesized that the general problem-solving strategy would increase performance on transfer tasks.

For the current topic, the general problem-solving strategy was specified as “draw as many graphs as you can.” Instructions to draw as many graphs as possible could potentially assist problem solvers in recognizing that generating moves in this way may be a useful, general strategy when trying to solve problems where the solution is not immediately available from long-term memory. Students in the general problem-solving group were also asked to select which of their graphs they thought best answered the question. This was necessary so that a final answer could be identified. To determine the effectiveness of the general problem-solving strategy, a comparison was made with a conventional problem-solving group asked to solve the problem without mention of the general problem solving strategy. The intention was to verify whether the general problem-solving strategy would facilitate problem solving on transfer problems. For this experiment, novices on the topic were chosen to test the hypothesis that a general problem-solving strategy would lead to higher learning than a conventional problem-solving strategy (Hypothesis 1).

### Method

**Participants.** Thirty Year 11 (mean age of 16.6 years) economics students from one Sydney high school participated in the study initially (13 females and 17 males). Each student was randomly assigned to a group representing one of two experimental conditions: the conventional problem-solving group ($N = 15$) or the general problem-solving group ($N = 15$). All students had some limited familiarity with the basic economics content (supply and demand) having previously studied it in their regular school classes, but nevertheless were considered novices.

**Materials and Procedure**

**Problem tasks.** There were four phases (prior knowledge test, acquisition phase, similar test, and transfer test) to this experiment requiring participants to complete a set of problems in each phase.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Group Means (SDs) of Scores on Each Test During Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prior knowledge (max = 16)</td>
</tr>
<tr>
<td>Worked example</td>
<td>10.6 (3.3)</td>
</tr>
<tr>
<td>General problem-solving</td>
<td>10.8 (2.7)</td>
</tr>
<tr>
<td>Total</td>
<td>10.7 (3.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Group Means (SDs) for Number of Angles Calculated in Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategy</td>
<td>Acquisition</td>
</tr>
<tr>
<td>Worked example</td>
<td>17.8 (6.3)</td>
</tr>
<tr>
<td>General problem-solving</td>
<td>22.3 (11.0)</td>
</tr>
<tr>
<td>Total</td>
<td>19.9 (9.0)</td>
</tr>
</tbody>
</table>
The prior knowledge test consisted of 10 multiple-choice questions that tested basic knowledge of supply and demand in an everyday sense. For example, one problem asked—"if there is an increase in the price of a good or service does this lead to an increase or decrease in demand?" A follow-up question asked— "does this lead to an increase or decrease in the supply of the good or service?" Students received one mark for each correct response in the prior knowledge test.

The acquisition phase consisted of two main stages (explanatory and problem-solving). In the explanatory stage, learning materials were provided through an explanatory 11-page (A4) summary of the topic (supply and demand) to both groups, which they were instructed to read and understand. This summary consisted of explanations of the following concepts: the law of demand, the demand curve and demand schedule, a change in demand, the law of supply, the supply curve and supply schedule, a change in supply, and market equilibrium. For example, it was explained that when a purchasable item is priced high there is less demand and when an item is priced low there is more demand, which is the law of demand. The explanatory materials also provided five example graphs illustrating the demand curve, an increase in demand, the supply curve, an increase in supply, and a supply and demand curve.

The second stage of acquisition consisted of problem-solving tasks based on the information provided in the explanatory stage. Students were required to solve problems using the information they had learned previously. For both groups, a set of six problems was constructed. All problems incorporated both supply and demand, only asked for one movement in either supply or demand and never both. Students were told to draw a graph or graphs (see below) but not explicitly told to draw a supply and demand graph. Based on how much they understood from the explanatory learning material they had to deduce that a supply and demand graph was required. Figure 2 provides an example of a supply and demand graph where there is also an increase in demand but no shift in the supply curve. In this example, students were required to draw a shift in the demand curve and no change to the supply curve.

To generate two different problem-solving strategies, the instructions given for each problem were different for the two groups. For the general problem-solving group they were instructed: Using the information given to you in the learning material, draw one graph illustrating this change in demand and/or supply. Answer this question by drawing as many graphs as you can and picking the best one. Make sure you show all of the graphs drawn. Graphs do not have to be drawn to scale. For the conventional problem-solving group they were instructed: Using the information given to you in the learning material, draw one graph illustrating this change in demand and/or supply. The graph does not have to be drawn to scale.

Students were familiar with drawing graphs as they had learned to draw graphs in mathematics and also had basic instruction in drawing graphs in economics. Furthermore, they were also provided with three different types of graphs during the previous explanatory stage.

Each problem in this acquisition set was awarded a maximum of two marks. If students drew more than one graph the last graph drawn on the page was taken as their solution to the problem, unless otherwise indicated. Full marks were awarded for drawing a correct supply and demand graph that illustrated an increase or decrease in demand or supply correctly as required. One mark was awarded for drawing a correct supply and demand graph but an incorrect increase or decrease in supply or demand. Half a mark was awarded for only drawing a correct supply or demand graph. No marks were awarded for drawing an incorrect supply or demand graph. All tests in the study including this test were double marked by two experienced scorers. If a discrepancy occurred on any question, then consensus was reached following further discussion.

For both groups, the problems were printed back to back on A4 sheets of paper with each page containing one problem and the written instruction printed at the top of each page. Both groups were instructed that they were required to answer Questions 1 to 6 in the space provided. Students were able to refer back to their explanatory learning materials when completing these problems. No feedback was provided.

For the similar test phase, a set of six problems, similar to those presented during acquisition, was provided to both groups. Unlike during acquisition, the general problem-solving group was not given any instructions to draw multiple graphs. Both groups were instructed only to draw one graph, consistent with the instructions given previously to the conventional problem-solving group. Each question was worth a maximum of two marks. The same marking criteria used in acquisition was adopted in the similar test phase. Students were instructed at the top of the first page to answer Questions 1 to 6 in the spaces provided and that graphs did not have to be drawn to scale. Sufficient space was provided for both groups to draw more than one diagram.

A transfer test was designed to determine whether students could transfer the strategy they learnt in acquisition to different domains within economics. This test was made up of four questions and was printed on A4 paper back-to-back. Students were instructed at the top of the first page of the transfer test to answer Questions 1 to 4 in the spaces provided. The set of transfer
problems were printed on two separate pages, one for Questions 1 and 2 and another for Questions 3 and 4.

This test introduced new material based on various topics in economics. The questions were based on the topics the circular flow of income and consumer behavior. All four questions in the transfer test were more open-ended style questions, as there were several possible answers to each of the questions. For example, students were asked “why” questions, where more than one response can be given. Even though it was possible to have more than one correct answer, correct answers were limited according to predetermined correct responses that were aligned to the economics topic area being questioned.

Question 1 was initially awarded a maximum of four marks. In this question, students were provided with a diagram of the circular flow of income and were asked to suggest reasons why this model is called the “circular flow of income.” Four key aspects need to be considered and therefore one mark each was awarded for (a) correctly explaining what a leakage is, (b) correctly explaining what an injection is, (c) correctly providing an explanation, and (d) correctly providing an example. Question 2 was initially awarded a maximum of 10 marks (five sections worth a maximum two marks each). In Question 2 students were required to complete a table. Students were provided with five different changes in a leakage or injection and had to explain how the five changes affect the level of economic activity and provide reasons why. One mark was awarded for correctly stating that the change led to a rise or fall in economic activity and one mark was awarded for providing a correct reason for such a rise or fall. Questions 3 and 4 were worth a maximum of two marks each and were economic comprehension questions not based on the circular flow of income diagram. Question 3 was based on business management, whereas Question 4 focused on market price adjustments. Two marks for each question were awarded for giving two correct answers in response to the questions. After the initial marking procedure all four questions were scaled to a maximum of two marks during analysis of the data to provide problem equality.

The experiment was conducted in the students’ regular economic 80 min class under strict supervision. A brief introduction regarding the format of the study was provided. Students were instructed that all answers should be written on the question papers provided. The entire experiment consisted of four phases: prior knowledge (15 min), acquisition (30 min), similar test (20 min), and transfer test (15 min). For each phase, the materials were distributed and then collected after the allotted time had expired.

Results and Discussion

Mean test scores are shown in Table 7. Consistent with Experiment 3 with a continuous measure of prior-knowledge, stepwise linear regression was completed on the data in the order: Step 1 (prior knowledge), Step 2 (strategy), and Step 3 (prior knowledge × strategy interaction).

Performance scores. For acquisition scores, Model 2 was significant, \( F(2, 77) = 8.64, p = .001 \), with prior knowledge a significant predictor \((R^2 = 0.32, \beta = .61, p < .001)\) but not strategy \((R^2 = 0.07, \beta = -.26, p = .10)\). On test scores, Model 3 was significant. \( F(3, 26) = 6.18, p < .01 \), but only strategy was found to be a significant predictor \((R^2 = 0.26, \beta = -.51, p < .01)\). The general problem-solving strategy (estimated \( M = 10.2, SE = 0.75 \)) led to higher test scores than the conventional problem-solving strategy (estimated \( M = 6.6, SE = 0.75 \)). For transfer scores, Model 2 was significant, \( F(2, 27) = 5.79, p < .01 \). Prior knowledge just failed to reach significance \((R^2 = 0.06, \beta = .32, p = .06)\), but strategy was a significant predictor \((R^2 = 0.24, \beta = -.50, p < .01)\). The general problem-solving strategy (estimated \( M = 4.5, SE = 0.31 \)) led to higher transfer test scores than the conventional problem-solving strategy (estimated \( M = 2.9, SE = 0.33 \)).

Graphs sketched. Group scores for the mean number of graphs drawn are shown in Table 8. For acquisition graphs, Model 2 was significant. \( F(2, 27) = 6.10, p < .01 \). Prior knowledge was not a significant predictor \((R^2 = 0.05, \beta = .40, p = .10)\), but strategy was significant \((\Delta R^2 = 0.26, \beta = -.51, p < .01)\). The general problem-solving strategy (estimated \( M = 7.9, SE = 0.69 \)) led to more graphs completed than the conventional problem-solving strategy (estimated \( M = 4.9, SE = 0.69 \)). On test graphs, Model 2 was significant, \( F(2, 27) = 6.24, p < .01 \). Prior knowledge was a significant predictor \((R^2 = 0.12, \beta = .42, p < .05)\), as was strategy \((\Delta R^2 = 0.20, \beta = -.45, p < .01)\). The general problem-solving strategy (estimated \( M = 6.0, SE = 0.41 \)) led to more graphs completed than the conventional problem-solving strategy (estimated \( M = 4.4, SE = 0.41 \)).

Ideas generated. To determine the extent to which the general problem-solving strategy was applied to the transfer phase the number of ideas written down per question were counted. Each contained statement provided was considered to be an idea. For example, the first question in the transfer phase asked students to suggest reasons why the circular flow of income model has this specific name. Responses included “because there are injections and outflows of money” and “money circulates around the economy.” Such responses were considered to be an individual idea.

For these data, regression analysis revealed that Model 2 was significant, \( F(2, 27) = 4.58, p < .05 \). Prior knowledge was a significant predictor \((R^2 = 0.13, \beta = .42, p < .05)\), as was strategy \((\Delta R^2 = 0.12, \beta = -.35, p < .05)\). The general problem-solving strategy (estimated \( M = 10.5, SE = 0.91 \)) led to significantly more

Table 7

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Prior knowledge (max = 10)</th>
<th>Acquisition (max = 12)</th>
<th>Similar (max = 12)</th>
<th>Transfer (max = 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional problem-solving</td>
<td>4.7 (2.2)</td>
<td>7.1 (3.9)</td>
<td>7.0 (4.1)</td>
<td>2.9 (1.3)</td>
</tr>
<tr>
<td>General problem-solving</td>
<td>4.1 (2.1)</td>
<td>8.3 (3.8)</td>
<td>10.1 (2.0)</td>
<td>4.4 (1.3)</td>
</tr>
<tr>
<td>Total</td>
<td>4.4 (2.1)</td>
<td>7.7 (3.9)</td>
<td>8.6 (3.6)</td>
<td>3.7 (1.5)</td>
</tr>
</tbody>
</table>
ideas being generated than the conventional problem-solving strategy (estimated $M = 7.8, SE = 0.91$).

This experiment used a different learning domain than the geometry tasks of the previous three experiments—the economics topic, supply and demand—to test whether a general problem-solving strategy could facilitate problem solving on transfer tasks (Hypothesis 1). Using novices, a significant advantage was found in favor of the general problem-solving strategy compared with a conventional problem-solving strategy on both similar and transfer tests. When considering the number of graphs drawn, the general problem-solving strategy not only generated more graphs during the acquisition problems, but also during the test phase, indicating that the strategy continued to be used after it was no longer required. Furthermore, more ideas were also generated during testing. Hence, consistent support was found in favor of Hypothesis 1.

The results differed from the geometry experiments in that the general problem solving strategy facilitated performance on both similar and transfer tasks. In the geometry experiments, although learners varied in mathematical ability, most had a significant amount of prior knowledge and so that domain-specific knowledge could be used to solve the similar problems, rendering the general problem solving strategy redundant. In the current experiment, the learners may have had less domain-specific knowledge and so on the similar test, less knowledge was available to generate solutions than on the similar tests in the first three experiments. In the absence of domain-specific knowledge, the use of the general problem solving strategy on the similar problem test should be beneficial, resulting in improved performance by the group alerted to the advisability of using that strategy.

**General Discussion**

The hypotheses of this article were generated directly from the evolutionary-based cognitive architecture used by cognitive load theory. The suggestion that human cognition functions in a manner similar to evolution by natural selection is very old and can be traced back to Darwin (1871/2003) with both philosophers (Popper, 1979) and psychologists (Campbell, 1960) adding more recent considerations to this viewpoint. Cognitive load theory has provided details to the analogy including the roles of working memory, long-term memory, and problem solving. It was suggested via the randomness as genesis principle that random generation and test, known to provide a default process to overcome problem-solving impass, was analogous to random mutation in evolutionary biology. It was suggested further that while random generate and test was a biologically primary skill to which problem solvers already had access, they may not be aware of which biologically secondary materials a random generate and test strategy could be applied. Accordingly, it was hypothesized that alerting problem solvers to the possibility of using this general problem-solving strategy should be beneficial on transfer problems with little or no overlap in the domain-specific knowledge required for solution. In contrast, where relevant knowledge was available, it was hypothesized via the information store and environmental organizing and linking principles that domain-specific knowledge would have primacy and overcome any advantage of the general strategy. These hypotheses were tested in the current experiments.

Evidence in support of these hypotheses was found in all four experiments. In the first three experiments conducted in the mathematical learning domain of geometry, evidence was found on transfer problems. In direct comparison to a conventional problem-solving strategy, the general problem-solving strategy was superior on both transfer (Experiment 1) and far transfer (Experiment 2) tasks. In comparison with a worked examples approach (Experiment 3), the general problem-solving strategy was superior on far-transfer problems only. It was also predicted that if the problems were fairly familiar to learners then the general problem-solving strategy would lose its effectiveness. Evidence in support of this argument was also found in the first three experiments on similar test problems, where no significant differences between strategies were found. In addition on the transfer problems of Experiment 1 and far-transfer problems in Experiment 2, strategy—ability interactions were found demonstrating that lower ability students benefitted the most from the general problem-solving strategy.

In Experiment 4, which featured the learning domain of economics that required graph sketching, a general problem-solving strategy was found to be superior to a conventional problem-solving strategy on problems both similar to those used during acquisition as well as on transfer problems. As this experiment, unlike the previous experiments, chose students with little experience of the given tasks, an advantage was found on similar problems.

The influence of prior-knowledge on the effectiveness of a general problem solving strategy was predicted by cognitive load theory, and supported by the empirical evidence. General problem-solving strategies such as the heuristics used in these experiments may be relatively effective only in the absence of relevant domain-specific knowledge. Learning to solve problems means acquiring the domain-specific knowledge required for particular classes of problems (Tricot & Sweller, 2014). When available, that knowledge is likely to be more effective and used in preference to a general strategy. However, when no or very little prior knowledge is available, the use of domain-specific solutions is not possible. Under such circumstances, general problem-solving strategies based on biologically primary knowledge may be effective.

General problem-solving strategies are central to the randomness as genesis principle (Sweller & Sweller, 2006). When faced with a novel problem to solve for which only limited, domain-specific knowledge is available, at various points humans may have little choice but to randomly generate problem-solving moves and test them for effectiveness. The ability to randomly generate moves is likely to be acquired as biologically primary knowledge and so used automatically without generating unhelpful extraneous cognitive load. While in the absence of sufficient knowledge, the use of a random generation strategy appears to be unavoidable if further problem-solving progress is to be achieved. However, it

---

**Table 8**

*Group Test Means (SDs) for Number of Graphs Calculated in Experiment 4*

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Acquisition</th>
<th>Similar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional problem-solving</td>
<td>5.1 (1.4)</td>
<td>4.5 (2.3)</td>
</tr>
<tr>
<td>General problem-solving</td>
<td>7.8 (3.6)</td>
<td>5.9 (1.0)</td>
</tr>
<tr>
<td>Total</td>
<td>6.43 (3.0)</td>
<td>5.2 (1.9)</td>
</tr>
</tbody>
</table>
must be emphasized that few, if any, moves on any problem are likely to be fully random. Some knowledge always is available if only to exclude the possibly very large number of moves that could be made on many problems.

Nevertheless, it is legitimate to ask how we know that moves are being generated randomly apart from the fact that under some circumstances there seems to be no alternative to random generation? The fact that learners shown the general problem-solving strategy in this study generated more moves than learners simply attempting to reach the problem goal, may indicate the influence of random generation. In all experiments, both the conventional and random generation groups were presented the problem goal in the same manner. The additional moves made by the general problem-solving groups presumably were not needed to reach the goal. The existence of those additional moves makes sense if learners are generating moves randomly and so provides possible empirical evidence for random generation.

Nevertheless, no direct evidence was collected showing students used a random generation process. Our theoretical argument predicted that the general problem-solving strategy would facilitate such a process. We know that the general problem-solving strategy led to better learning outcomes for some students, but we do not know about the actual problem-solving processes that took place. It was also argued that students with high levels of prior knowledge would not need to use a random generation problem-solving strategy, as they would be able to use domain-specific knowledge. Consistent with cognitive load theory the results also supported this prior knowledge argument but it was not directly tested. It was also predicted that using a general problem-solving strategy would reduce extraneous cognitive load compared with a conventional problem-solving strategy. Again no direct evidence was collected to support this assumption. Further research is required to test these hypotheses directly. Individual testing and the collection of verbal protocols, as well as more forensic analyses of problem-solving strategies and cognitive load, are needed in future studies to identify the actual cognitive processes used and the cognitive load generated by the various groups.

While the current experiments were generated by the randomness as genesis principle that is part of the cognitive architecture used by cognitive load theory, the results could be explained or perhaps related, post factum, to other areas. First, generative learning theory (Mayer, 2003, 2012; Mayer, Steinhoff, Bower, & Mars, 1995; Wittrock, 1989) and the generation effect (Bertsch, Pesta, Wiscott, & McDaniel, 2007) both suggest that learning is enhanced when learners generate actions rather than when they are shown the same actions. The current work may be related to generative learning theory and the generation effect although it needs to be noted that with the exception of Experiment 3 that used worked examples, the current experiments used generation for both the control and experimental groups. What was generated was changed and compared rather than the presence or absence of generation.

The second area to which the current results might be related is brainstorming (Meadow, Parnes, & Reese, 1959). Brainstorming research has indicated that when people are asked to generate as many ideas as they can, they are likely to generate more ideas, and possibly more good ideas than people not asked to generate many ideas. Although the bulk of this research has been carried out using materials unrelated to those used in an instructional context (e.g., experimental participants may be asked to find as many uses as they can for a brick), the general procedure is analogous to the procedures used in the current experiments.

The study had a number of limitations that could have influenced the results. First, the experiments lacked power in that most cell sizes across the four experiments were small. It was predicted that when learners had sufficient prior knowledge, the general problem-solving strategy would lose its effectiveness. Results generally supported this prediction; however, it is possible that a number of nonsignificant group differences may have been because of lack of power, leading to Type 2 errors. Nevertheless, significant differences with large effect sizes were found on transfer problems, as predicted, suggesting that lack of power may not have been an issue.

Second, in the first two experiments school test records were used to classify students as low or high mathematical ability. Furthermore, in Experiment 1 students were chosen from different schools, potentially exposing them to different school-based strategies, as well as making it difficult to categorize mathematical ability consistently across schools. Consequently, more accurate prior-knowledge testing closely aligned with the materials presented could have been conducted to eliminate these factors.

Third, consistent with much research that looks at transfer as well as retention effects, the sequence order was fixed in the common pattern of retention followed by transfer tests, with no counterbalancing. It is feasible that there was an order effect in favor of transfer problems (except in Experiment 4 where both tests produced significant differences) because of additional learning accrued during the retention testing. Nevertheless, if such learning took place it seemed to only advantage the general problem solving groups and not the conventional problem solving groups. As the different treatments were given only during acquisition, any additional advantage gained during retention tested should have benefited both groups.

Fourth, the acquisition times were quite short and may have reduced the impact of the various strategies used. We do not know how the different groups used this time, or even if the whole time was used. Future research could investigate the impact of learning time, and also collect some verbal protocol information to identify the type of processes applied by the various groups. Such research could also address the other limitations, documented above, by increasing power, using more consistent prior knowledge tests, and allowing for potential test order effects. It would also be informative to conduct testing over a longer period of time to determine if the positive effects are lasting, and also use different domains to see if the general problem-solving strategy is more suited to some domains and types of questions than others.

In summary, from an instructional perspective, the current results suggest that students should be alerted to using the random generate and test, problem-solving strategy. When faced with problems for which they do not have knowledge of a solution, these strategies can enhance problem-solving performance. Arguably, general problem-solving strategies do not need to be taught because they are biologically primary (Tricot & Sweller, 2014) but when faced with a biologically secondary task for which random generation and test may be beneficial, indicating to learners that they should consider using the strategy on that task may facilitate
problem solving performance. However, it should not be seen as a substitute for domain-specific knowledge indicating to problem solvers that moves are appropriate when faced with particular problem states.

Although the students in this study were able to use the general problem-solving strategies without explicit instruction, based on the current results, knowing when to apply those strategies in a novel, biologically secondary task does require explicit instruction. It was necessary to draw a link for the students between the strategy and the targeted domain.

It can be concluded that tapping into primary knowledge through general problem-solving strategies can increase problem-solving success rates. Using such strategies seems most applicable when learners are novices in the domain.

References


**Appendix**

**Far-Transfer Test Problems Used in Experiment 2**

**Question 1.** Find the value of x

![Diagram 1](image1)

**Question 2.** Find the value of x

![Diagram 2](image2)

**Question 3.** AB = 3, AC = 4, BC \parallel ED, CD \parallel BE
If DF = 2DE, find the value of DF

![Diagram 3](image3)

**Question 4.** 0 is the center of the circle. A, B, C are on the circle, find the value of angle x

![Diagram 4](image4)

**Question 5.** a, b, c and d are related by the following formulae:

\[a + b = c + d\]

\[4a = b\]

\[c = 2a - b\]

If \(a = 3\), what is the value of d?

**Question 6.** In April, Janice ate 80 sweets. Each day she ate at least 2 sweets. Every 5th day she ate exactly 4 sweets. What could have been the maximum she could have eaten on 1 day?

**Question 7.** Consider the sequence 1, 3, 6, 10, 15, 21, 28, etc. If the 99th term is 4950, what is the 100th term?