The variability effect in high and low guidance instruction: A cognitive load perspective

Vicki Likourezos
University of New South Wales
< v.likourezos@unsw.edu.au >

Slava Kalyuga
University of New South Wales
<s.kalyuga@unsw.edu.au>

The variability effect occurs when learners' exposure to highly variable tasks results in better learning. It was hypothesised that learners who studied high variability worked examples would obtain higher post-test scores compared to learners who studied low variability examples, and learners who self-generated problem solutions for the same high or low variability tasks. This hypothesis was not supported. However, subjective ratings of difficulty supported the assumptions based on cognitive load theory. From a practical perspective, these results suggest that novice learners should initially be presented with low variability, high-guidance learning tasks to reduce a potential cognitive overload.

Effective mathematical learning occurs when students can apply mathematical skills in new contexts by generalising from their prior knowledge. This capacity is known as transfer, and it represents a critical teaching goal in mathematics. For this reason, understanding the factors that impact learner capacity to transfer more efficiently is essential for improving instruction. Prior studies within cognitive load theory suggested that transfer performance may be improved by exposure to worked examples and variability (Cooper & Sweller, 1987; Paas & van Merriënboer, 1994; Sweller & Cooper, 1985). Given the importance of transfer, the present study investigated whether presenting students with high variability mathematical tasks, with or without worked examples, would boost transfer performance, compared to low variability tasks.

The study involved two levels of guidance (full guidance or no guidance) and two levels of variability (high or low). Full guidance tasks required students to study problems with worked examples, and no guidance tasks required students to attempt to generate problem solutions without any guidance. Variability was achieved by changing the range of tasks under which worked examples were studied or problems had to be solved. In the case of high variability, this was done via applying the same process in a wider variety of contexts by changing the surface features (i.e. numbers) and varying the structure of the problems (i.e. question format), as opposed to the low variability problems whereby the surface features changed but not the structure. The objective of the study was to examine which guidance-variability combination would result in superior transfer performance outcomes.

Theoretical Framework

Transfer of learning occurs when students understand mathematical procedures and develop generalised mathematical techniques to solve new problems beyond a single context. Such understanding takes place when students construct deep, organised, domain-specific knowledge structures known as schemata. Despite the importance of promoting activities that enable students to develop schemata, schema acquisition could be hindered by the limited capacity of working memory. In view of this, Sweller (1988) developed cognitive load theory by arguing that appropriate instructional design procedures could be used to reduce cognitive load during learning (see Sweller, Ayres, and Kalyuga (2011), for a comprehensive review of the theory). According to cognitive load theory, conventional problem-solving techniques (e.g., means-ends analysis) impose a heavy working memory load on inexperienced problem solvers.

2019. In G. Hine, S. Blackley, & A. Cooke (Eds.). Mathematics Education Research: Impacting Practice *(Proceedings of the 42nd annual conference of the Mathematics Education Research Group of Australasia)* pp. 444-451. Perth: MERGA.

Multiple studies within a cognitive load framework have provided evidence for the worked example effect, whereby learning is enhanced for novice learners if they are exposed to instruction that relies on worked examples, rather than instruction that directs them to attempt to solve problems with minimal explicit guidance (Atkinson, Derry, Renkl, & Wortham, 2000; Carrol, 1994; Cooper & Sweller, 1987; Paas & van Merriënboer, 1994; Schwonke et al., 2009; Sweller & Cooper, 1985). Despite the overwhelming evidence for the effectiveness of worked examples for novice learners, previous research has also established the expertise-reversal effect, whereby solving problems becomes superior to studying worked examples for more experienced learners who are able to rely on their available solution schemas. For these learners, worked examples become redundant, and processing and interpreting them may impose unnecessary cognitive load and inhibit learning (Bokosmaty, Sweller, & Kalyuga, 2015; Kalyuga, 2007; Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Kalyuga, Rikers, & Paas, 2012).

Introducing variability of learning tasks was proposed as a means of making learning more effective for learners, especially in enhancing their transfer capabilities. Research has shown that exposure to highly variable example-based instruction (compared to less variable, homogeneous examples), from a cognitive load perspective, gives learners the opportunity to engage in deeper processing, enabling new knowledge to be adapted to novel situations, resulting in enhanced transfer performance (Clark, Nguyen, & Sweller, 2006; Paas & van Merriënboer, 1994; Quilici & Mayer, 1996; van Merriënboer & Sweller, 2005). Indeed, Paas and van Merriënboer's (1994) study of variability was the first to show that worked examples led to better transfer compared to problem solving tasks. Their study compared high and low variability versions of worked examples and conventional problem-solving formats, and showed that the worked examples-high variability combination resulted in superior transfer outcomes. In addition, the investigation also showed less cognitive load in the worked examples groups compared to the problem-solving groups, with an interaction that indicated that variability was effective in the worked examples groups but not in the problem-solving groups.

The present investigation is built on Paas and van Merriënboer's (1994) study by exploring variability further, with a particular focus of the effects of variability on problem solving. Given the frequency of problem-solving formats used in post-secondary and tertiary education, it is critical to understand the impact of variability at this higher academic level. The specific hypotheses were:

H_1: Learners who study worked examples that provide explicit solution steps will yield better post-test performance, compared to learners who generate problem solutions without the provision of any solution steps.

H_2: Providing learners with high variability tasks under worked example conditions will generate a variability effect with better post-test performance, compared to using low variability tasks, while problem-solving conditions will not generate this difference.

H_3: Subjective ratings of difficulty for attempting to solve problems (without guidance) will be higher compared to studying fully-guided worked examples, irrespective of the level of variability of the task.

H_4: Subjective ratings of difficulty for completing high variability tasks will be higher compared to completing low variability tasks, irrespective of the level of guidance provided.

Method

Participants

The participants were 68 mathematics students, aged between 18 and 65 (M_{age} = 26.68, SD_{age} = 8.40), enrolled in a university preparation program at the University of New South Wales, Sydney. This post-secondary education program gives the students the qualification to apply for a university program. The sample comprised of 24 females (35%) and 44 males.

Materials

The materials used in the experiment focused on the topic area involving the definition of a quadratic function; the roots of a quadratic function; the axis of symmetry and the vertex of a parabola; and how to draw the graph of a quadratic function. All participants were regarded as novice learners in relation to quadratic functions, as this topic had not yet been taught to them at the time when the experiment was conducted, and it was the next scheduled topic in the mathematics preparation program.

During the first part of the Learning Phase, the experiment convenor provided explicit instruction for the topic whereby solutions of the relevant tasks were demonstrated on the board. During the second half of the Learning Phase, each participant received a handout in accordance with the experimental group they were randomly allocated to: 'worked examples-high variability'; 'worked examples-low variability'; 'problem solving-high variability'; or 'problem solving-low variability'. Worked examples were developed according to the principles of cognitive load theory by removing any redundant information and inserting arrows to assist learners with physically integrating disparate sources of information (redundancy and split-attention effects in cognitive load theory; see Sweller et al., 2011, for an overview) to remove extraneous cognitive activities that could interfere with learning, such as processing redundant and split-source information. Four different versions of Question 1, according to the four experimental conditions, are presented in Figures 1, 2, 3 and 4. The worked example-based tasks (in Figures 1 and 2) contain step-by-step solutions on how to solve the problem. The problem-solving tasks (in Figures 3 and 4) contain only problem-solving statements with no written instructions or diagrams.

Given $f(x) = 4x + 8$, for all real x,

(i)
Find $f(0)$:

$$f(0) = 4(0) + 8$$
$$= 0 + 8$$
$$= 8$$

(ii)
Find $f(-2)$:

$$f(-2) = 4(-2) + 8$$
$$= -8 + 8$$
$$= 0$$

(iii)
Find $f(a)$:

$$f(a) = 4(a) + 8$$
$$= 4a + 8$$

Figure 1. Worked example – low variability.

Given $g(x) = 5 - 2x$, for all real x,

(i) Find $g(-3x)$:

$$g(-3x) = 5 - 2(-3x)$$
$$= 5 + 6x$$

(ii) Find $g\left(\frac{1}{4}\right)$:

$$g\left(\frac{1}{4}\right) = 5 - 2\left(\frac{1}{4}\right)$$
$$= 5 - \left(\frac{1}{2}\right)$$
$$= 4\frac{1}{2}$$

(iii) Find $g(a + 5)$:

$$g(a + 5) = 5 - 2(a + 5)$$
$$= 5 - 2a - 10$$
$$= -5 - 2a$$

Figure 2. Worked example – high variability.

Given $f(x) = 4x + 8$, for all real x,

(i) Find $f(0)$

(ii) Find $f(-2)$

(iii) Find $f(a)$

Figure 3. Problem solving – low variability.

Given $g(x) = 5 - 2x$, for all real x,

(i) Find $g(-3x)$

(ii) Find $g(\frac{1}{4})$

(iii) Find $g(a + 5)$

Figure 4. Problem solving – high variability.

All participants were given a single-item, 9-point Likert-type rating scale to complete which was used to measure cognitive load imposed during their completion of the Learning Phase handout. The Post-Test contained non-transfer questions (questions that were structurally similar to the questions used during direct instruction on the board), as well as transfer questions (questions that were structurally different from those used during direct instruction). The transfer questions required a higher level of understanding applied in a relatively new task situation. On the other hand, the non-transfer questions tested for procedural skills.

Procedure

The participants were randomly assigned to one of the four experimental conditions at the start of the experiment, namely: worked examples-high variability group (17 students); worked examples-low variability group (17 students); problem solving-high variability group (17 students); and problem solving-low variability group (17 students).

Correspondingly, the duration of the experiment, which was one and a half hours, was conducted during the participants' normal mathematics lecture and tutorial time, and consisted of a Learning Phase (60 minutes) and a Post-Test Phase (30 minutes).

Results

There were two independent variables: level of variability and level of guidance, and four dependent variables: Post-Test (total) scores; Post-Test (similar questions) scores; Post-Test (transfer questions) scores; and subjective ratings of difficulty. Table 1 shows the descriptive statistics for the participants' performance.

Prior Knowledge

Prior mathematical knowledge of each participant was measured by averaging scores for four previous class tests that were completed before the commencement of the experiment. A one-way between-groups analysis of variance was conducted for the average class test scores to compare the level of prior mathematical knowledge for the four groups. It involved one independent variable (the condition group) across four levels (worked examples-high variability group, worked examples-low variability group, problem solving-high variability group, and problem solving-low variability group) and one dependent variable (average class test score). The results were not statistically significant, $F(3,64) = .45$, $MSE = 204.73$, $p = .72$, partial $\eta^2 = .02$. Therefore, the average class test scores were not used to control for any differences between the experimental groups for any Post-Test performance results.

Post-Test Scores

Three 2-by-2 between-groups analyses of variance were conducted on the Post-Test (total) scores, Post-Test (similar questions) scores, and Post-Test (transfer questions) scores. The results for guidance were not statistically significant for the Post-Test (total) scores, $F(1,64) = .07$, $MSE = 44.86$, $p = .80$, partial $\eta^2 = .001$, the Post-Test (similar questions) scores, $F(1,64) = .61$, $MSE = 542.12$, $p = .44$, partial $\eta^2 = .01$, and the Post-Test (transfer questions) scores, $F(1,64) = 1.10$, $MSE = 634.22$, $p = .30$, partial $\eta^2 = .02$. Likewise, the results for variability were not statistically significant for the Post-Test (total) scores, $F(1,64) = 1.92$, $MSE = 1300.43$, $p = .17$, partial $\eta^2 = .03$, the Post-Test (similar questions) scores, $F(1,64) = 1.78$, $MSE = 1582.12$, $p = .19$, partial $\eta^2 = .03$, and the Post-Test (transfer questions) scores, $F(1,64) = 1.45$, $MSE = 836.29$, $p = .23$, partial $\eta^2 = .02$. The results of these analyses show no evidence of a relationship between levels of guidance (worked examples or problem solving) or levels of task variability (high or low) for the completion of post-test tasks in similar and novel situations. Hence the expected worked example effect and variability effect were not obtained.

Table 1
Means (Standard Deviations) of the Average Class Test Scores, Post-Test (Total) Scores, Post-Test (Similar Questions) Scores, Post-Test (Transfer Questions) Scores and Subjective Ratings of Difficulty

	Experimental Conditions							
	Worked Examples-High Variability Group		Worked Examples-Low Variability Group		Problem Solving-High Variability Group		Problem Solving-Low Variability Group	
Dependent variables	*M (SD)*	*N*	*M (SD)*	*N*	*M (SD)*	*N*	*M (SD)*	*N*
Average Class Test Scores (%)	72.42 (22.09)	17	71.35 (18.05)	17	64.87 (21.22)	17	71.47 (23.69)	17
Post-Test (Total) Scores (%)	39.63 (28.53)	17	40.71 (28.18)	17	30.34 (21.92)	17	46.75 (24.94)	17
Post-Test (Similar Questions) Scores (%)	42.59 (29.39)	17	45.41 (33.91)	17	30.12 (27.35)	17	46.59 (28.18)	17
Post-Test (Transfer Questions) Scores (%)	33.94 (29.68)	17	31.67 (25.35)	17	30.77 (16.54)	17	47.06 (22.73)	17
Subjective Ratings of Difficulty (1-9)	3.65 (1.37)	17	1.76 (1.09)	17	6.82 (1.55)	17	6.06 (2.30)	17

The variability by guidance interactions were not statistically significant for the Post-Test (total) scores, $F(1,64) = 1.47$, $MSE = 998.48$, $p = .23$, partial $\eta^2 = .02$, the Post-Test (similar questions) scores, $F(1,64) = .89$, $MSE = 791.53$, $p = .35$, partial $\eta^2 = .01$, and the Post-Test (transfer questions) scores, $F(1,64) = 2.53$, $MSE = 1462.62$, $p = .12$, partial $\eta^2 = .04$. Hence there was no simultaneous effect of the two independent variables (guidance and variability) on any of the dependent variables (total, similar and transfer post-test scores) in which one of the independent variables differed depending on the level of the other independent variable.

Subjective Ratings of Difficulty

A 2-by-2 between-groups analysis of variance was conducted to assess the effect of the two independent variables on subjective ratings of difficulty (cognitive load). The results showed a statistically significant main effect for guidance, $F(1,64) = 88.08$, $MSE = 237.19$, $p < .001$, partial $\eta^2 = .58$, and for variability, $F(1,64) = 11.06$, $MSE = 29.78$, $p = .001$, partial $\eta^2 = .15$. As anticipated, the subjective ratings of difficulty were less for the worked examples groups compared to the problem-solving groups, and less for the low variability groups compared to the high variability groups. This indicates that lower cognitive load was imposed on learners who studied fully-guided worked examples compared to learners who generated solutions to problem-solving tasks (without any guidance). Equivalently, learners who worked on low variability tasks experienced lower cognitive load compared to learners who worked on high variability tasks. The variability by guidance interaction for cognitive load was not statistically significant, $F(1,64) = 1.97$, $MSE = 5.31$, $p = .17$, partial $\eta^2 = .03$.

Discussion and Conclusion

This experiment tested the hypotheses that when learners study worked examples (compared to learners that generate problem solutions without any guidance) and when they are provided with high variability tasks (compared to low variability tasks) under worked example conditions, they will attain superior transfer skills. Neither of these two hypotheses, H_1 and H_2, were supported, Unlike the results obtained by Paas and van Merriënboer (1994), who had found that students who studied high-variability worked examples achieved better transfer performance (with an interaction indicating that variability was effective in the worked examples groups but not in the problem-solving groups), the present study did not demonstrate either a worked example effect or a variability effect on both similar or transfer post-tests. It is possible that the lack of significance may have been due to the relatively small sample size. Also, it should be noted that the average class test score of 70.03% was sufficiently high. It is known from the expertise reversal effect that as the learner level of expertise increases, the possibility of obtaining a worked example effect decreases.

Despite the absence of a worked example effect, it seems that the worked examples in the present experiment successfully reduced extraneous cognitive load as evidenced by subjective ratings of cognitive load. Learners could understand the high and low variability tasks more easily by studying worked examples compared to solving problems. This is in line with cognitive load theory which argues for the superiority of worked examples to problem-solving due to reducing unnecessary cognitive load. Additionally, the results for cognitive load demonstrated a significant advantage for participants in the low variability groups, compared to the high variability groups. Completion of low variability tasks was associated with a significant reduction in cognitive load as less mental effort was required to identify a surface match between the similarly structured questions without the need to go any further. In contrast, processing high variability tasks required more mental effort to process the deeper features until the underlying common features were found. Thus, both hypotheses, H_3 and H_4, in relation to cognitive load were supported. An instructional implication of these results is that novice learners who possess inadequate knowledge schemas should avoid processing excessive amounts of interactive elements of information contained in high variability tasks, otherwise they are more likely to experience cognitive overload. To reduce a potential cognitive overload, learners should initially be presented low variability, high-guidance (worked example-based) learning tasks.

References

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181-214.

Bokosmaty, S., Sweller, J., & Kalyuga, S. (2015). Learning geometry problem solving by studying worked examples: Effects of learner guidance and expertise. *American Educational Research Journal*, *52*(2), 307-333.

Carrol, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, *86*(3), 360-367.
Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.
Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*(4), 347-362.
Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*(4), 509-539.
Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23-31.
Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, *93*(3), 579-588.
Kalyuga, S., Rikers, R., & Paas, F. (2012). Educational implications of expertise reversal effects in learning and performance of complex cognitive and sensorimotor skills. *Educational Psychology Review*, *24*(2), 313-337.
Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*(1), 122-133.
Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, *88*(1), 144-161.
Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., & Salden, R. (2009). The worked-example effect: Not an artefact of lousy control conditions. *Computers in Human Behavior*, *25*(2), 258-266.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.
Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59-89.
van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147-177.