MODESTUM
OPEN ACCESS

EURASIA Journal of Mathematics, Science and Technology Education, 2019, 15(10), em1760
ISSN:1305-8223 (online)
Research Paper https://doi.org/10.29333/ejmste/105482

Tracing Effect in the Worked Examples-based Learning: An Exploration of Individual Differences in Working Memory Capacity

Lian-Ming Yeo [1*], Yuh-Tsuen Tzeng [1]

[1] National Chung Cheng University, TAIWAN

Received 3 December 2018 ▪ Revised 8 March 2019 ▪ Accepted 9 March 2019

ABSTRACT

Previous study has shown that tracing gesture may enhance the worked example-based learning by reducing cognitive load. The present study attempted to replicate the previous results and further explored the individual differences in tracing effect in relation to the learners' working-memory capacity. Specifically, 11- to 13-year-old students varies in their working-memory capacity were asked to study worked examples on angles relationships involving parallel lines either without tracing instruction or with tracing instruction. The results showed that the tracing group outperformed the non-tracing group on a subsequent test and reported lower levels of test difficulty. In addition, the learning outcomes between the low- and high-capacity individuals in both conditions were comparable, suggesting that tracing gesture offers a simple yet effective embodied technique that may further enhances the worked example-based learning by reducing cognitive load, however, its facilitation effect was not affected by the individuals' working-memory capacity.

Keywords: cognitive load theory, embodied cognition, tracing gesture, worked examples, working-memory capacity

INTRODUCTION

Since 1980s, cognitive load theory has developed rapidly as an influential theory in the fields of educational psychology. Based on the current knowledge about human cognitive architecture and functions, the theory has provided guidelines to generate effective instructional designs to support learning process (Paas, Van Gog & Sweller, 2010; Sweller, Van Merriënboer & Paas, 1998). It was assumed that human cognitive architecture consists of an effectively unlimited long-term memory which interacts with a working memory that is very limited in both capacity and duration (Baddeley, Eysenck, & Anderson, 2009). Long-term memory stores huge amounts of schemas that vary in their degree of automaticity allowing us to categorize different problem states and solve the problem based on the most appropriate solution moves. Working memory deals with all conscious activities including the construction and automation of schemas either prior to it being stored in long-term memory or after it has been stored. Accordingly, the theory suggested that the learners' working memory load should be reduced when engaging in novel activities that are directed at high-quality schemas construction to enhance the learning processes and outcomes (Paas et al., 2010; Sweller et al., 1998; Sweller, 1994).

The emerging embodied cognition perspectives put forward the role of body in shaping human mind (Barsalou, 2008; Glenberg, Witt & Metcalfe, 2013; Wilson, 2002). Recently, the use of bodily resources such as hands gestures has been shown positively affected learning process (Alibali & Nathan, 2012; Edwards, Ferrara & Moore-Russo, 2014; Novack & Goldin-Meadow, 2015). The present study focuses on the effect of a specific type of hand gesture, namely tracing gesture (or finger tracing), an action of an individual using his or her fingers (e.g. index finger) to make a trace of an object's outline (e.g. the outline of a triangle) whether in the air or on the surface by simulating of the act of writing or drawing without using other instruments such as pen or pencil. In the context of teaching and learning, this kind of finger movement has been over century practiced in the Montessori educational pedagogy (Montessori, 1912) to help children recognizing geometrical shapes, letters of alphabet and numbers better than just only visualizing them. Recently, there were experimental studies found that tracing gesture also

© 2019 by the authors; licensee Modestum Ltd., UK. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).
✉ yeolm17@hotmail.com **(*Correspondence)** ✉ ttcytt@ccu.edu.tw

Contribution of this paper to the literature

- The present study corroborated the previous results on the facilitation effect of tracing gesture in the mathematics worked examples-based learning by using visuo-spatial-based instructional materials.
- The present study provided the first evidence on the role of individual differences in working-memory capacity to investigate the effect of tracing gesture on cognitive load and geometrical problem solving.
- The present study made an initial attempt to measure the working-memory capacity of Malaysian elementary school students with operation span task. The findings suggested that multiple working memory span tasks should be administered to obtain more accurate measurement of the individuals' working-memory capacity.

may facilitate relatively more complex learning tasks such as anatomy and physiology learning (Macken & Ginns, 2014); geometrical problem solving (Hu, Ginns & Bobis, 2015); and route learning (So, Ching, Lim, Cheng, & Ip, 2014).

To the best of our knowledge, Hu and her colleagues (2015) has made an initial attempt to examine the effect of tracing gesture on cognitive load and problem-solving performance, by using worked-example-based instructional materials which has been demonstrated as a more effective and efficient way to teach novice learners (Carroll, 1994; Cooper & Sweller, 1987; Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Rourke & Sweller, 2009; Sweller & Cooper, 1985). Based on the theoretical frameworks of cognitive load theory and embodied cognition perspectives, Hu et al. (2015) argued that tracing gesture may serve as an effective means in directing the learners' attention to the location of relevant information in the worked examples for further processing, which could reduce the waste of cognitive resources in visual search to generate extra cognitive resources available for the use of effective problem-solving schema construction. In their study, the students in tracing condition were instructed to trace out the elements of worked examples with their dominant hand's index finger. For instance, the students use their index fingers to trace out the parallel lines based on the written instruction in worked examples. Meanwhile, the students in non-tracing condition were instructed to put their hands on the laps while studying the content of worked examples. The results showed that the students in tracing condition outperformed the non-tracing group as measured by the number of correct answers, error rate, and time to solution with large statistical reliable effects for the advanced test questions and marginal effects for basic questions in solution time. Moreover, the tracing group also reported lower levels of test difficulty for both the basic and advanced test questions, reflecting lower levels of the experienced cognitive load as a result of enhanced schema construction. Thus, the findings suggested that tracing gestures used in the worked examples facilitate the construction of better problem-solving schemas over and above the typical benefits of learning from the materials that already been generated from cognitive load-based instructional design.

Given the initial evidence that the explicit instruction to trace out the essential elements in the worked examples might facilitate the learning for novices in terms of test performance and the experienced cognitive load, the present study was designed to replicate the previous results and further explore whether there are individual differences in the tracing effect in relation to the learners' working-memory capacity. Working-memory capacity which varies among individuals is the ability to store and manipulate limited relevant information in memory over a short duration (Baddeley, Eysenck, & Anderson, 2009; Daneman & Carpenter, 1980). Recently, Marstaller and Burianová (2013) has examined the differences in working-memory accuracy (i.e. recalling list of letters) between individuals who were instructed to produce deictic gesture (i.e. pointing) and those who were instructed not to produce the gesture in relation to their working-memory capacity. The researchers found that only low-capacity individuals who did not gesture showed a reduced recalling performance but not high-capacity individuals who did not gesture. On the other hand, low-capacity and high-capacity individuals who gestured did not showed a decrease in the recalling performance. These findings indicated that individuals who have lower working-memory capacity were more likely to benefit from producing the deictic gesture in recalling list of letters, and thus suggesting that these gesture effects are affected by individuals' working-memory capacity.

The existing study has demonstrated that the effects of deictic gesture (i.e. pointing gesture) depends on the individual's working-memory capacity. However, it remains to be established whether the facilitation effects of other deictic gesture (i.e. tracing gesture) are affected by the individual differences in working-memory capacity (cf. Van Wermeskerken et al., 2016). Thus, the present study was designed to address this issue by incorporating tracing gesture in the worked-example-based instructional materials. Based on the evidence from Hu et al.'s (2015) works that tracing gesture could help to reduce working memory load so that more cognitive resources are available for schema construction, it is possible that producing this gesture is more effective for those who have lower working-memory capacity. In contrast, individuals who have higher working memory-capacity are supposed to benefit to a lesser extent as they have larger capacity to process the information irrespective of whether they use such external aids or not. Thus, it was hypothesized that low-capacity individuals in the tracing condition

would demonstrate better test performance and report lower levels in ratings of test item difficulty, compared to their counterparts in the non-tracing condition. On the other hand, high-capacity individuals in the tracing condition would demonstrate comparable results with their counterparts in the non-tracing condition. Note that the tracing gesture effect in relation to the learners' working memory capacity is still understudied, the present hypotheses should be regarded as exploratory in nature.

METHOD

Participants

Participants were 72 students, including 29 boys and 43 girls, aged between 11 and 13 years old (M = 12.35, SD =0.28), randomly selected from two Chinese elementary schools located in Johor Bahru city, Johor, Malaysia. The participants were novices to the mathematics topic of angle relationships involving parallel lines. To ensure the participants in all conditions were equally novice with the tested topic, they were asked whether this topic had been learned in advance before recruiting them in the experiment. Children's parents submitted written consent prior to their child's participation in the present study.

Materials and Procedure

Participants were presented with two independent tasks: first, an operation span task to assess working-memory capacity and then followed by the worked examples-based learning task which consisted of three phases: initial instruction phase, acquisition phase, and test phase. For the operation span task, in groups of four to ten students, depending on the availability of computer resources, each participant completed the task individually for approximately 15 minutes. By using median split, the participants were divided into two groups (i.e. high-capacity and low capacity) based on their scores on the operation span task. These two groups were then randomly assigned to the tracing or the non-tracing condition. This pseudo-randomization was performed to assure the number of high-capacity and low-capacity individuals were divided equally across the two conditions. A week later, the participants were tested individually according to their designated condition for approximately 20 minutes in a quiet classroom.

Operation span task. The automated shortened operation span task running on E-Prime 2.0 was used to measure the participants' working-memory capacity (Foster, Shipstead, Harrison, Hicks, Redick, & Engle, 2015). The shortened version was used to reduce the administration time without decreasing the reliability and validity in measuring the working-memory capacity. Oswald, McAbee, Redick and Hambrick (2014) have demonstrated in their Study 2 that the coefficient alpha for shortened operation span task was .71 with the mean time saved 6.3 minutes in measuring the participants' working-memory capacity. More importantly, it was shown that the verbal working-memory span task such as operation span task load on the same factor in a factor analysis as spatial working-memory span task (Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004) based on the claim that working-memory capacity primarily reflect domain-general executive function (Engle, Tuholski, Laughlin, & Conway, 1999). Thus, the operation span task can account for individual's working-memory capacity same as other span tasks.

For each trial in this operation span task, the participants were presented with a series of simple arithmetic equations and were asked to solve them as fast and accurate as possible. Each equation was followed by the presentation of a possible answer and the participants were asked to judge whether the presented answer was correct or not by clicking one of the two buttons (i.e. true or false). Each answer was followed by a single letter which the participants were asked to retain for later recall. After four, five, or six such sequences of equations, possible answers, and letters (i.e. processing-and-storage presentations), a recall grid is presented, and participants must click on the letters they saw during the trial in correct serial order. After the recall, the computer provided feedback about the number of letters correctly recalled, the number of math errors, and the percentage of math accuracy which presented in the upper right-hand corner of the screen (see **Figure 1**). The participants were encouraged to maintain their math accuracy at or above 85% at all times in order to have the data of those participants who were attempting to solve both the math equations and remember the letters (Unsworth, Heitz, Schrock, & Engle, 2005). In addition to that 85% accuracy criterion, the participants were also encouraged to solve the math equations as fast as possible to prevent them from rehearsing the letters. If the participants took more time to solve the math equations than their mean time (based on participant's performance during the processing-task only condition) plus 2.5 SD, the program would automatically move on and count that trial as an error.

Each participant solved two practice trials for the following three conditions respectively before proceeding to the real trials: (i) storage-task only, (ii) processing-task only, and (iii) processing-and-storage tasks interleaved. In the real trials for this shortened operation span task, the participants were required to solve 6 trials with 2 each of the list lengths 4, 5, and 6 (instead of 15 trials with 3 each of list lengths 3, 4, 5, 6, and 7 in the standard operation

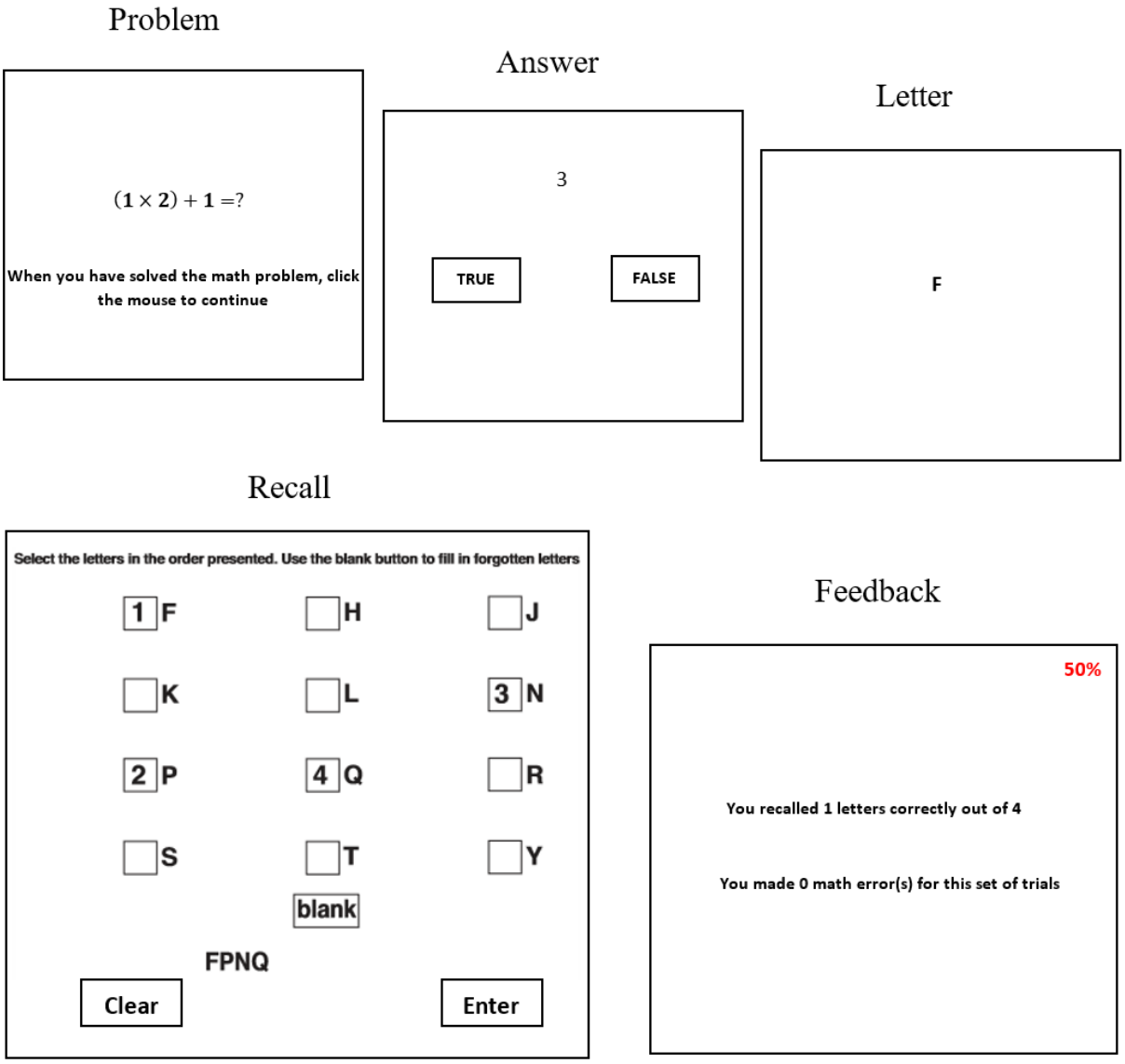

Figure 1. Illustration of the operation span task. First, a math problem is presented. After click the mouse, participants judge whether the presented answer is correct or incorrect. Then followed by a letter to be memorized for 800 msec. For recall, participants select the letters in the correct order. The feedback is presented for 2,000 msec

span task). Thus, the shortened version allows the researchers obtain the measure of working-memory capacity in less time than using the full-length version of the task.

Scoring. The operation span task provided span scores based on an absolute-scoring and partial-scoring method. The partial span score is the total number of letters recalled in the correct serial position on memory trials. For the absolute span score, the participants only received credit if they were 100% accurate within that trial (i.e. set size). For example, in a set trial of M-S-L, if the participant recalled all 3 letters completely correct in the serial position, both the absolute and partial score will be a 3 for the trial. However, if the participant gave a response M-H-L where 2 of the 3 letters are correctly recalled in the serial position, the partial score will be a 2 and the absolute score will be a 0 (because the participant was not 100% accurate on that trial). The partial score will always be equal to or higher than the absolute score.

As a measure of an individual's working-memory capacity, the partial span scoring was used in the present study because it has more variance and thus allows for better discrimination between high and low capacity participants (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005). The empirical evidence has shown that the partial scoring is superior to absolute scoring based on higher test-retest correlations and internal consistencies (Conway et al., 2005; Redick, Broadway, Meier, Kuriakose, Unsworth, Kane, & Engle, 2012). Thus, the partial scoring was strongly recommended as it is more reliable and sensitive to individual differences in working-memory

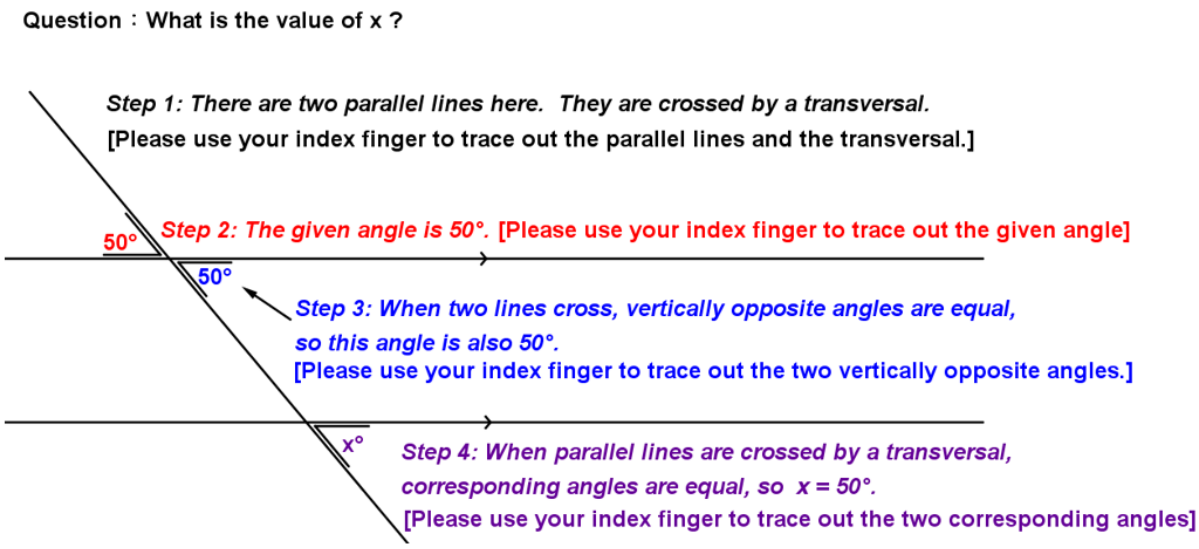

Figure 2. The first worked example for tracing condition adapted from Hu et al.'s (2015) works

capacity. As mentioned previously, this shortened operation span task consisted of 6 trials with 2 each of list lengths 4, 5, and 6, thus the task providing a maximum possible span score of 30.

Initial instruction phase. All the participants in both conditions were given five minutes to study a 5-pages instructional material with each page contains the following information respectively: (1) the definition of parallel lines and transversal; (2) the definition of interior and exterior angles; (3) the vertically opposite angles are equal; (4) the corresponding angles are equal; (5) the sum of co-interior angles is 180 degrees. For each angle relationship as presented from page 3 to page 5, the instruction consisted of its definition, diagrams displaying the locations of the specific angles, and an example demonstrating the application of this angle relationship to solve a problem.

Acquisition phase. After completing the initial instruction phase, all participants in both conditions were given two worked examples containing the solution steps to find a missing angle by using the three angle relationships. For each worked example, they were asked to spend two minutes to read and study carefully. The participants in tracing condition were required to trace out the specified elements of the diagram by using their index finger of their writing hand as stated in the bracketed instruction of every solution step (see **Figure 2**). For the non-tracing condition, the participants were required to read and understand the solution steps in the worked examples with their hands placed on their laps. Each worked example was paired with a similar practice problem, with a maximum of two minutes to obtain the solution. Once the answer was given, the participants received feedback from the experimenter whether their answer was correct or incorrect. The participants who provided an incorrect answer were asked to try again within the 2-minute time limit. If these participants could not provide the correct answer within the time limit, they were required to study the worked example again and then went back to solve the practice problem until the correct answer was attained.

Test phase. Once the participants completed the acquisition phase, they were asked to solve six problems to measure their learning outcomes. The first two problems were similar questions which possessed similar diagrams and solution steps to the worked examples but with different numbers. These similar questions could be solved with two-step solution using the correct angle relationships. The next four problems were considered as transfer questions which had different surface features such as orientation and number of lines (parallel and transversal). These transfer questions required three-step solution of the angle relationships with various permutations in their combinations to obtain the correct solution. For instance, see **Figure 3**, the three-step solution for one of the transfer questions might be as follows: (1) the vertically opposite angles are equal; (2) the corresponding angles are equal; and (3) the sum of co-interior angles is 180°, with these three solution steps, the value of x could be obtained. The answer's accuracy and solution time for each question provided by the participants were recorded for the analysis of their test performance.

Test item difficulty self-reports. After the participants solved each test question, they were required to rate the difficulty of the question immediately. The test item difficulty of similar questions and transfer questions were defined as the average of the difficulty ratings across the two items and four items respectively. The average of these ratings served as an indicator of their experienced cognitive load in the attempts to find the solution for each question during the test phase, that is, the participants' experience of working memory load when solving the problems. In the present study, the experimenter had adapted a 5-point illustrated subjective rating scale of test difficulty which was designed by Hu et al. (2015), ranging from 1 being "very easy" to 5 being "very difficult". In

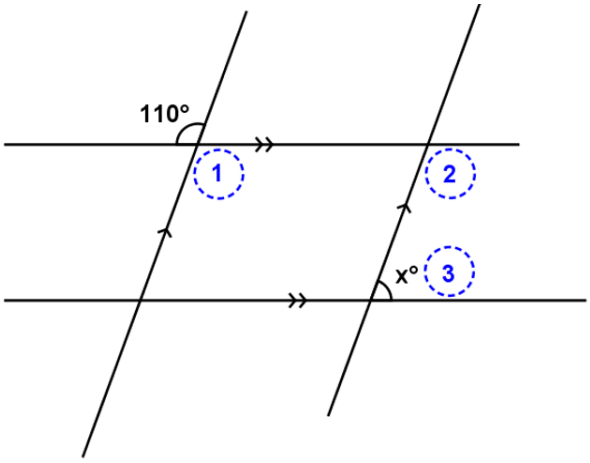

Figure 3. Example of a transfer question in the test phase

Table 1. Descriptive statistics of the dependent variables for the similar questions

	Dependent Variables	Tracing	Non-Tracing	Total
Low WMC	Accuracy	1.83 (.38)	1.78 (.43)	1.81 (.41)
	Total solution time	32.91 (26.40)	34.79 (13.64)	33.85 (20.02)
	Test difficulty	1.94 (.59)	2.47 (.85)	2.21 (.72)
High WMC	Accuracy	1.83 (.38)	1.67 (.59)	1.75 (.49)
	Total solution time	29.64 (15.50)	26.86 (14.19)	28.25 (14.85)
	Test difficulty	1.69 (.69)	2.25 (.86)	1.97 (.78)
Total	Accuracy	1.83 (.38)	1.73 (.51)	1.78 (.45)
	Total solution time	31.28 (20.95)	30.83 (13.92)	31.06 (17.44)
	Test difficulty	1.82 (.64)	2.36 (.86)	2.09 (.75)

Note: Mean (and *SD*); WMC, working-memory capacity

the current rating scale, five cartoon faces were positioned above the numeric anchors (from 1 to 5) to help young-aged participants indicate how they felt about the difficulty of the test questions while solving them. For instance, on the top of number 1 "very easy", a big smiling face was displayed, indicating that a problem has been solved relatively effortless (i.e. on the basis of a well-constructed schema). In contrast, a frowning face was put on top of the number 5 "very difficult", indicating that a question requires considerable conscious effort to be solved based on the problem-solving schema that might be incomplete and/or difficult to be activated in the working memory.

After completing all the experimental procedures, the participants were asked not to share any information regarding the content of the experiment to the other students and finally they were thanked for their voluntary participation in the study.

DATA ANALYSIS

A two-way multivariate analysis of variance (MANOVA) was performed separately for the similar questions and transfer questions with two independent variables (gesturing condition and working-memory capacity) and three dependent variables (number of correct answers, total solution time, and ratings of test item difficulty). The results of data analysis reported the tests of significance at alpha level of .05 together with the estimates of mean difference effect size, η^2.

RESULTS

Table 1 and **Table 2** provide the descriptive statistics of the dependent variables under analysis for the similar questions and transfer questions respectively. To explore the effect of working-memory capacity, the median of this variable was used to divide the participants into two groups and categorized participants who had partial span scores less than or equal to 20 as low-capacity individuals (N = 36) and participants who had the scores greater than or equal to 21 as high-capacity individuals (N = 36). The categorical measure was used because it showed broad differences between groups.

Table 2. Descriptive statistics of the dependent variables for the transfer questions

	Dependent Variables	Tracing	Non-Tracing	Total
Low WMC	Accuracy	2.89 (1.02)	2.06 (1.31)	2.48 (1.17)
	Total solution time	103.36 (56.88)	122.42 (67.88)	112.89 (62.38)
	Test difficulty	2.68 (.63)	3.38 (.68)	3.03 (0.66)
High WMC	Accuracy	3.11 (.68)	2.56 (.71)	2.84 (.70)
	Total solution time	102.83 (32.88)	101.38 (45.67)	102.11 (39.28)
	Test difficulty	2.65 (.63)	3.11 (.69)	2.88 (.66)
Total	Accuracy	3.00 (.85)	2.31 (1.01)	2.66 (.93)
	Total solution time	103.10 (44.88)	111.90 (56.78)	107.50 (50.83)
	Test difficulty	2.67 (.63)	3.25 (.69)	2.96 (.66)

Note: Mean (and *SD*); WMC, working-memory capacity

Table 3. Correlation matrix

	Accuracy	Total solution time	Test difficulty
Similar questions			
Accuracy	1	-.237*	-.374**
Total solution time	-.237*	1	.449**
Test difficulty	-.374**	.449**	1
Transfer questions			
Accuracy	1	.066	-.385**
Total solution time	.066	1	.278*
Test difficulty	-.385**	.278*	1

*. Correlation is significant at the .05 level (2-tailed)
**. Correlation is significant at the .01 level (2-tailed)

Since the Pearson's *r* correlation matrix in **Table 3** showed that the three dependent variables for similar questions and transfer questions were correlated with one another at a low to moderate levels ($|r| < .50$) that can avoid the risk of multicollinearity (Leech, Barrett, & Morgan, 2005), thus the use of two-way MANOVA was decided to assess whether tracing group and non-tracing group with higher and lower working-memory capacity have different learning outcomes, and whether there was an interaction between gesturing condition and working-memory capacity. The MANOVA was performed separately for the similar questions and transfer questions to assess the linear combination of the three dependent variables (i.e. accuracy, total solution time, and test difficulty) which represent the students' learning outcomes, specifically their test performance and the experienced cognitive load in the mathematical problem solving.

In the present study, participants with low and high working-memory capacity were randomly assigned to the tracing and non-tracing conditions to ensure the independence of observation. Moreover, the sample size for all four groups were equal. MANOVA is robust to violations of homogeneity of variance/covariance matrices if the sizes of groups are nearly equal (Leech et al., 2005). Therefore, all the assumptions of MANOVA in terms of independent observation, multivariate normality, and homogeneity of variance/covariance matrices were met.

For the similar questions, the results of MANOVA showed that the interaction between gesturing condition and working-memory capacity was not significant, Wilk's Lambda = .988, F (3, 66) = .26, p = .856, multivariate η^2 =.012. The main effect for working-memory capacity was also not significant, Wilk's Lambda = .948, F (3,66) = 1.21, p = .315, multivariate η^2 =.052. However, the main effect for gesturing condition was significant, Wilk's Lambda = .847, F (3, 66) = 3.97, p = .012, multivariate η^2 =.153.

For the transfer questions, the MANOVA revealed a significant multivariate effect for gesturing condition, Wilk's Lambda = .785, F (3, 66) = 6.01, p = .001, multivariate η^2 =.215, and a non-significant multivariate effect for working-memory capacity, Wilk's Lambda = .947, F (3,66) = 1.22, p = .309, multivariate η^2 =.053. A non-significant multivariate effect for interaction was also found, Wilk's Lambda = .981, F (3, 66) = .44, p = .727, multivariate η^2 =.019. Both the results of MANOVA for similar questions and transfer questions indicated that the linear composite of accuracy, total solution time and test difficulty differs for tracing and non-tracing groups.

Follow up univariate ANOVAs (**Table 4**) indicate that the effect of gesturing condition was significant for accuracy of transfer questions, F (1, 68) = 9.38, p = .003, η^2 = .121; test difficulty of similar questions, F (1, 68) = 9.24, p = .003, η^2 = .120; and test difficulty of transfer questions, F (1, 68) = 13.94, p < .001, η^2 = .170. The tracing group solved more transfer questions correctly (M = 3.00, SD = .85) than non-tracing group (M = 2.31, SD = 1.01). Moreover, the tracing group reported lower level in the ratings of test item difficulty for similar questions (M = 1.82, SD = .64) than non-tracing group (M = 2.36, SD = .86). The tracing group also reported transfer questions as less difficult (M = 2.67, SD = .63) than the non-tracing group (M = 3.25, SD = .69). These results indicated that the

Table 4. Effects of working-memory capacity and gesturing on students' learning outcomes

Source	Dependent Variable	*df*	*F*	*η*	*p*
Working-memory Capacity	*Similar questions*				
	Accuracy	1	.27	.06	.607
	Total solution time	1	1.70	.15	.196
	Test difficulty	1	1.76	.16	.190
	Transfer questions				
	Accuracy	1	2.54	.19	.116
	Total solution time	1	.76	.10	.386
	Test difficulty	1	.89	.11	.348
Gesturing	*Similar questions*				
	Accuracy	1	1.07	.13	.304
	Total solution time	1	.01	0	.918
	Test difficulty	1	9.24	.35	.003
	Transfer questions				
	Accuracy	1	9.38	.35	.003
	Total solution time	1	.51	.08	.479
	Test difficulty	1	13.94	.41	.001
Capacity * Gesturing	*Similar questions*				
	Accuracy	1	.27	.06	.607
	Total solution time	1	.30	.06	.589
	Test difficulty	1	.01	0	.938
	Transfer questions				
	Accuracy	1	.38	.07	.542
	Total solution time	1	.69	.10	.410
	Test difficulty	1	.59	.09	.447
Error	*Similar questions*				
	Accuracy	68			
	Total solution time	68			
	Test difficulty	68			
	Transfer questions				
	Accuracy	68			
	Total solution time	68			
	Test difficulty	68			

tracing group demonstrated better test performance (the more numbers of correct answers) and the reduction of experienced cognitive load (the lower level in the ratings of test item difficulty), as compared with the non-tracing group during subsequent test phase.

DISCUSSION

The present study was designed to replicate the findings of previous research (Hu et al., 2015) on the effect of tracing gesture in the worked example-based instruction and further explore whether there are individual differences in the tracing effect in relation to the learners' working-memory capacity. It was predicted that low-capacity individuals would benefit more than high-capacity individuals because tracing gesture might help to reduce the working memory load so that the total cognitive load does not exceed the individual's available capacity. In contrast, high-capacity individuals are supposed to benefit to a lesser extent as they have larger capacity to process the information irrespective of whether they produce this gesture or not.

The results showed the significant main effect of tracing gesture; however, neither the main effect of working-memory capacity nor an interaction between both factors was significant. This indicated although the exploratory hypothesis that tracing effects might be affected by working-memory capacity was not supported, the hypothesis about the facilitation effects of tracing gesture in the previous study was confirmed. Participants in the tracing condition significantly outperformed the non-tracing condition across most of the variates during the test phase. Specifically, the participants regardless of low or high-capacity in the tracing condition demonstrated better test performance (i.e. accuracy of transfer questions) and reported lower levels in the ratings of test item difficulty (for both similar and transfer questions), compared to their counterparts in the non-tracing condition. These findings indicated that tracing gesture not only assisted low-capacity individuals to construct better schema in the worked example-based learning to solve the test questions with lower level of cognitive load but also confer its facilitation effect to high-capacity individuals. In other words, both the low and high-capacity individuals benefit similarly from producing tracing gesture in the present worked example-based learning.

As the significant main effect of working-memory capacity was not found, the results did not show a significant difference between the low- and high-capacity individuals in the tracing condition on test performance and the ratings of test item difficulty. Similarly, there was no significant difference between low- and high-capacity participants in the non-tracing condition across all those variates. These findings indicated that the low- and high-capacity individuals in the tracing condition demonstrated comparable better learning outcomes as a result of better schema construction; whereas in the non-tracing condition, high-capacity individuals yielded comparable worse results with low-capacity individuals as they did not construct better problem-solving schemas from the worked examples to solve the test questions with lower level of cognitive load.

Taken together, the findings of the present study showed that facilitation effect of tracing gesture was not affected by individual's working-memory capacity. However, these findings were in contrast with previous study that gesture effect depends on individuals' working-memory capacity which using a novel gesture task and an operation span task (Marstaller & Burianová, 2013). The possible explanation for the present findings might be argued that the use of operation span task which is a verbal working-memory span task to measure working-memory capacity might not have lined up perfectly with the kind of working memory needed to complete the present worked-example learning task successfully even though this span task has been shown to be a good measurement of working-memory capacity that primarily reflect domain-general executive attention (Kane et al., 2004). This is because the present learning task not only consisted of verbal component but also spatial component. Thus, it might be more valid and adequate to administer spatial working-memory span task along with verbal working-memory span task and then use the average of scores on both tasks in measuring the individual's working-memory capacity (Conway et al., 2005). Another possible explanation is that the use of working-memory capacity's median to categorize the participants into low- and high-capacity groups might cause misclassification especially to those participants whose capacity scores were close to the median. The lack of significant effect of working-memory capacity in the present study might due to the two groups that were not distinctive in capacity. To overcome this median-split design, Conway et al. (2005) recommended the extreme-groups design which using upper and lower quartiles of capacity scores to categorize high- and low-capacity individuals respectively in order to reduce the problem of misclassification. Future research might shed more light onto the relationship between working-memory measures and the tracing effect in the present worked example-based learning.

Given the paucity of studies investigating the tracing effect in worked example-based learning, there are several other limitations still can be examined in order to extend its applicability in the mathematical problem solving. First, the previous works have demonstrated that tracing gesture can lighten cognitive load and enhance performance in tasks that have a spatial component, nonetheless, the role of tracing gesture in non-spatial tasks can be further explored (cf. Ginns, Hu, Byrne, & Bobis, 2016). Second, previous study measured tracing effect on learning immediately after instruction, future investigations with follow-up tests will provide more robust tests of learning, to examine whether or not tracing gesture has significant effect on acquisition of new knowledge and helps in the retention of that knowledge during instruction (cf. Cook, Duffy & Fenn, 2013; Cook, Mitchell & Goldin-Meadow, 2008). Third, in the attempt to investigate the effect of gesture on cognitive load and learning, previous studies did not consider the role of individual differences in terms of general cognitive abilities such as fluid intelligence. Whether there are individual differences in the effect of tracing gesture on general cognitive abilities is certainly worth further investigation to develop more effective instructional designs that match with the learners' cognitive dispositions.

The findings of present research not only provide empirical evidence to support the cognitive load theory and embodied cognition perspective, but also provide useful educational implications. At theoretical level, the findings of the present study particularly in tracing gesture effect on learners' working memory load would provide support for cognitive load theory to expand its scope to "embodied processing" in mathematical learning. At educational practice level, mathematics teachers might gain insights of using tracing gesture as an alternative and potential embodied instructional strategy that can be taken to promote learning in mathematics as well as raise the students' interests and motivation. Moreover, the students also benefit from receiving better instruction which emphasizes the strategies that exploit the advantage of using bodily resources to facilitate their mathematical understanding and problem-solving skills. In terms of other educational endeavor, the writers of text books or reference books can adopt the idea of adding the instruction to trace out elements in the worked examples. This simple written instruction can be taken as a cost-effective strategy to improve and enrich the mathematical experience of the students. The tracing activity not only suitable for individualize learning, but also can use as collaborative and investigative approaches to learning, increasing concept development and understanding.

While independent study on the issue of using embodied resources in schools appears to be increasing, studies of the tracing effect in mathematical learning are yet scarce. It is hoped that the findings of the present study will lead to an increased understanding among education professionals to anticipate the predictable strengths and challenges to the use of tracing gesture, and to consider its positive interventions in the progressive development of "embodied learning" as an established practice in schools.

ACKNOWLEDGEMENT

Ministry of Science and Technology Taiwan Scholarship Program for Malaysian students, serial no. 2015001.

REFERENCES

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences, 21*, 247–286. https://doi.org/10.1080/10508406.2011.611446

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory*. Hove & New York: Psychology Press.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology, 86*(3), 360–376. https://doi.org/10.1037/0022-0663.86.3.360

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development, 84*, 1863–1871. https://doi.org/10.1111/cdev.12097

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. https://doi.org/10.1016/j.cognition.2007.04.010

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347–362. http://dx.doi.org/10.1037/0022-0663.79.4.347

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Edwards, L. D., Ferrara, F., & Moore-Russo, D. (2014 Eds.). *Emerging perspectives on gesture and embodiment in mathematics*. USA: Information Age Publishing, Inc.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*(3), 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S. & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory and Cognition, 43*(2), 226–236. https://doi.org/10.3758/s13421-014-0461-7

Ginns, P., Hu, F. T., Byrne, E., & Bobis, J. (2016). Learning by tracing worked examples. *Applied Cognitive Psychology, 30*(2), 160–169. https://doi.org/10.1002/acp.3171

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science*, *8*(5), 573–585. https://doi.org/10.1177/1745691613498098

Hu, F. T., Ginns, P., & Bobis, J. (2015). Getting the point: tracing worked examples enhances learning. *Learning and Instruction, 35*, 85–93. https://doi.org/10.1016/j.learninstruc.2014.10.002

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. https://doi.org/10.1037/0022-0663.93.3.579

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Use and interpretation*. USA: Lawrence Erlbaum Associates, Inc.

Macken, L., & Ginns, P. (2014). Pointing and tracing gestures may enhance anatomy and physiology learning. *Medical Teacher, 36*(7), 569–601. https://doi.org/10.3109/0142159X.2014.899684

Marstaller, L. & Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin & Review*, *20*(3), 496–500. https://doi.org/10.3758/s13423-012-0365-0

Montessori, M. (1912). *The Montessori method*. Cambridge, MA: Robert Bentley, Inc.

Novack, M., & Goldin-Meadow, S. (2015). Learning from gesture: How our hands change our minds. *Educational Psychology Review, 27*(3), 405–412. https://doi.org/10.1007/s10648-015-9325-3

Oswald, F. O., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods,47*(4), 1343–1355. https://doi.org/10.3758/s13428-014-0543-2

Paas, F., Van Gog, T., & Sweller, J. (2010). Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review, 22*(2), 115–121. https://doi.org/10.1007/s10648-010-9133-8

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*(3), 164–171. https://doi.org/10.1027/1015-5759/a000123

Rourke, A., & Sweller, J. (2009). The worked-example effect using ill-defined problems: Learning to recognize designers' styles. *Learning and Instruction, 19*(2), 185–199. https://doi.org/10.1016/j.learninstruc.2008.03.006

So, W. C., Ching, T. H., Lim, P. E., Cheng, X., & Ip, K. Y. (2014). Producing gestures facilitates route learning. *PLoS ONE, 9*(11), 1–21. https://doi.org/10.1371/journal.pone.0112543

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

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. http://www.jstor.org/stable/3233555

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498–505. https://doi.org/10.3758/BF03192720

Van Wermeskerken, M., Fijan, N., Eielts, C., & Pouw, W. T. J. L. (2016). Observation of depictive versus tracing gestures selectively aids verbal versus visual-spatial learning in primary school children. *Applied Cognitive Psychology, 30*(5), 806–814. https://doi.org/10.1002/acp.3256

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322

http://www.ejmste.com