2015

Giving learning a helping hand: finger tracing of temperature graphs on an iPad

Shirley Agostinho
University of Wollongong, shirleya@uow.edu.au

Sharon K. Tindall-Ford
University of Wollongong, sharontf@uow.edu.au

Paul Ginns
University of Sydney

Steven J. Howard
University of Wollongong, stevenh@uow.edu.au

Wayne Leahy
Macquarie University

See next page for additional authors

Publication Details

Agostinho, S., Tindall-Ford, S., ginns, p., Howard, S. J., Leahy, W. & Paas, F. (2015). Giving learning a helping hand: finger tracing of temperature graphs on an iPad. Educational Psychology Review, 27 (3), 427-443.

Giving learning a helping hand: finger tracing of temperature graphs on an iPad

Abstract

Gesturally controlled information and communication technologies, such as tablet devices, are becoming increasingly popular tools for teaching and learning. Based on the theoretical frameworks of cognitive load and embodied cognition, this study investigated the impact of explicit instructions to trace out elements of tablet-based worked examples on mathematical problem-solving. Participants were 61 primary school children (8-11 years), who studied worked examples on an iPad either by tracing temperature graphs with their index finger or without such tracing. Results confirmed the main hypothesis that finger tracing as a form of biologically primary knowledge would support the construction of biologically secondary knowledge needed to understand temperature graphs. Children in the tracing condition achieved higher performance on transfer test questions. The theoretical and practical implications of the results are discussed.

Keywords

learning, giving, hand, finger, tracing, helping, temperature, ipad, graphs

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Agostinho, S., Tindall-Ford, S., ginns, p., Howard, S. J., Leahy, W. & Paas, F. (2015). Giving learning a helping hand: finger tracing of temperature graphs on an iPad. Educational Psychology Review, 27 (3), 427-443.

Authors

Shirley Agostinho, Sharon K. Tindall-Ford, Paul Ginns, Steven J. Howard, Wayne Leahy, and Fred Paas

GIVING LEARNING A HELPING HAND: FINGER TRACING OF TEMPERATURE GRAPHS ON AN IPAD

Abstract

Gesturally-controlled information and communication technologies, such as tablet devices, are becoming increasingly popular tools for teaching and learning. Based on the theoretical frameworks of cognitive load and embodied cognition, this study investigated the impact of explicit instructions to trace out elements of tablet-based worked examples on mathematical problem-solving. Participants were 61 primary school children (8-11 years), who studied worked examples on an iPad either by tracing temperature graphs with their index finger or without such tracing. Results confirmed the main hypothesis that finger tracing as a form of biologically primary knowledge would support the construction of biologically secondary knowledge needed to understand temperature graphs. Children in the tracing condition achieved higher performance on transfer test questions. The theoretical and practical implications of the results are discussed.

Keywords

Cognitive load theory, embodied cognition, tracing effect, tablets, iPads,

Giving Learning a Helping Hand: Finger Tracing of Temperature Graphs on an iPad

Introduction

There is an increasing uptake of gesturally-controlled information and communication technologies (ICTs) in the classroom (Johnson, Adams Becker, Estrada, & Freeman, 2014). These technologies enable the user to interact with a computer device in ways that mimic natural communication forms such as speaking, using body movements and hand gestures (Sheu & Chen, 2014) and as such are also referred to as 'intuitive technology' (Johnson et al, 2014). Some examples of use in the classroom are ICTs based on gestural recognition, such as Microsoft Kinect, and input, such as tablets, e.g., the iPad (for a review of gesture-based computing research in education, see Sheu & Chen, 2014). While ICTs based on gestural recognition interfaces are likely to remain niche products in classrooms, the prevalence of tablet use for learning in both schools and workplaces is increasing as part of a growing trend towards a "bring your own device" model (Docebo, 2014). Perhaps the defining distinction between the interface of a typical personal computer and a tablet is the latter's use of gestural input through finger touching and swiping. It may seem obvious that the ease-of-use afforded by gestural input should support learning. Nonetheless, Spitzer (2013) has cautioned against the uncritical adoption of tablets for learning, particularly by the very young, since "swiping" may be a relatively impoverished means of learning through interaction with the physical world. More generally, there are relatively few experimental evaluations of tablets' roles in teaching and learning, including the design of gestural interactions for learning from tablet-based instruction. In this study the use of personal tablets as the mode of instruction was both for theoretical and practical purposes. Theoretically, this study builds on the work of Hu, Ginns, and Bobis (2014, 2015) to examine whether the benefits of tracing on learning can generalize to other interfaces such as a digital tablet. The practical significance of this study is that tablets are increasingly being adopted in classrooms (Johnson, Adams Becker, Estrada, & Freeman, 2014), yet there is little research on how they may optimally be used for learning. The present study was designed to investigate the impact of explicit instructions to trace out elements of tablet-based worked examples on mathematical problem-solving. Our hypotheses are based on cognitive load theory and embodied cognition perspectives, which are reviewed below.

An evolutionarily informed cognitive load theory

Cognitive load theory (Sweller, Ayres, & Kalyuga, 2011) offers empirically supported instructional solutions to manage cognitive load in learning environments. Cognitive load is the amount of working memory resources allocated when studying learning materials. Cognitive load theory has drawn on a number of theoretical and empirical perspectives during its development, including schema theory (Bartlett, 1932; Marshall, 1995), working memory (Baddeley, 2012), expert-novice differences (Kalyuga, 2007; Van Gog, Ericsson, Rikers, & Paas, 2005), and evolutionary theory (Geary, 2008; Sweller, 2003). This last development holds the promise of novel instructional designs drawing on Geary's (2008) distinction between biologically primary and biologically secondary knowledge. The first of these categories refers to knowledge, which we are genetically predisposed to develop quite easily, usually through cultural immersion. Examples include learning to recognise faces, and speaking a mother tongue. The second of these categories refers to knowledge that requires extended, conscious and effortful learning, often within the context of purpose-built educational institutions. Examples include learning to write, and understanding conceptual and procedural knowledge of the sciences, and mathematics.

Drawing on this theorizing, Paas and Sweller (2012) posited the substantial limitations of working memory, in terms of both capacity (Cowan, 2001; Miller, 1956) and duration (Peterson & Peterson, 1959) when dealing with novel information, may apply much more clearly to biologically secondary knowledge than biologically primary knowledge. The authors also considered whether some forms of biologically primary knowledge might support the development of biologically secondary knowledge. A number of existing cognitive load theory-based effects could be interpreted in the light of this possibility. For example, the collective working memory effect (Kirschner, Paas, & Kirschner, 2009, 2011) occurs when "collaborating learners can gain from each other's working memory capacity during learning" (Paas & Sweller, 2012, p.30). Under an evolutionary account of this effect, face-to-face communication is biologically primary: humans have evolved to communicate and coordinate, because doing so enhances the chances of survival. Likewise, the human movement effect (Van Gog, Paas, Marcus, Ayres, & Sweller, 2009; Wong, Marcus, et al., 2009) occurs when learners studying materials incorporating hand movements benefit more from animations than transient animations than static graphics; Paas and Sweller interpret these findings in terms of the survival value of observational learning from conspecifics' movements, a mode of learning that may have its neurological basis in the mirror neuron system (for a review see Rizzolatti & Craighero, 2004).

Notions of *embodied cognition* (for reviews see Foglia & Wilson, 2013; Glenberg, Witt, & Metcalfe, 2013) formed the final body of scholarship identified by Paas and Sweller (2012) as productive for an evolutionarily informed cognitive load theory. Research has shown that gesturing has a fundamental impact on a range of educationally relevant cognitive functions (for reviews see Goldin-Meadow, 2011; Hostetter, 2012). While a learner might employ a range of different gestures while learning, either alone or in company (see Kendon, 2004, for a discussion of gesture taxonomies), the discussion below focuses on research demonstrating substantial linkages between hand positioning and visual processing. Such findings form the basis for hypothesizing gestures involving pointing that might constitute biologically primary knowledge capable of supporting construction of biologically secondary knowledge (Paas & Sweller, 2012).

Hand position, pointing and touching gestures, and attentional processes

Over the past decade, basic research on the neural basis of multi-modal perception and attention (e.g., Abrams, Davoli, Du, Knapp, & Paull, 2008; Cosman & Vecera, 2010; Spence, 2010; Shams & Seitz, 2008) has demonstrated the close connections between hand positioning and visually based cognitive processing. For example, examining participants' performance across three visual attention tasks (visual search, inhibition of return, and attentional blink), Abrams et al. systematically varied the proximity of participants' hands to the stimulus display. They found that whether or not the hands were visible, participants' vision was enhanced for objects near the hands. They drew on evolutionary theorizing to explain these results, arguing hand proximity "could facilitate the detailed evaluation of objects for potential manipulation, or the assessment of potentially dangerous objects for a defensive response" (p.1035).

The emergence across all cultures of *protodeclarative pointing* gestures (Liszkowski, Brown, Callaghan, Takada, & de Vos, 2012; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004) by most children by the age of 12 months supports the argument for pointing as a form of biologically primary knowledge. Such gestures occur when an infant points to an object or event as part of an attempt to share attention and interest (Liszkowski et al., 2004), and act as an effective means of guiding joint attention (Liszkowski et al., 2012), forming the basis for an infant's identification and understanding of the world. Evidence for this latter claim comes from a meta-analysis by Colonnesi, Stams, Koster, and Noom (2010) of both concurrent and longitudinal investigations of infants' pointing and language development. Colonnesi et al. concluded the existence of statistically reliable concurrent and longitudinal meta-analytic correlations means "the pointing gesture not only predates language, but is also prospectively associated with the development of language" (p. 361).

There is now substantial research demonstrating interactions between visual attention processes and hand position while pointing. For example, Chum, Bekkering, Dodd, and Pratt (2007) found that encoding spatial arrays with pointing movements towards the visual display led to better memory performance. They suggested that pointing makes the feature of spatial arrangement more salient and hence improves the encoding of the arrays. Such benefits extend to situations involving interaction with the environment through touch, which might include simultaneously looking at or listening to elements in the environment. Talsma, Senkowski, Soto-Faraco, and Woldorff (2010) review research demonstrating synergistic effects on attentional processes when visual, auditory, and/or tactile inputs are synchronised. Of particular relevance to the present study is their general conclusion that "temporally and spatially aligned sensory inputs in different modalities have a higher likelihood to be favored for further processing, and thus to capture an individual's attention, than do stimuli that are not aligned" (Talsma et al., 2010, p. 400). For example, Van der Burg, Olivers, Bronkhorst, and Theeuwes (2009) found participants' search times for targets in highly distracting visual environments were substantially improved when the target's color change was accompanied by a tactile signal. They concluded synchrony between visual and tactile inputs acts to guide attention in visually complex environments by making the visual event more salient.

Pointing and tracing to support learning of biologically secondary knowledge

Several lines of research support the contention that pointing gestures – including those that incorporate touch – may support the learning of biologically secondary knowledge. The practice of *print referencing* involves an adult tracking the printed text with a finger while reading the text to a pre-literate child, and pointing to the print in text or in illustrations while asking or answering questions (Evans, Williamson, & Pursoo, 2008; Justice, Kaderavek, Fan, Sofka, & Hunt, 2009). In a large field experiment, Piasta, Justice, McGinty, and Kaderavek (2012) found the use of preschool teachers' print references enhanced children's early literacy skills, including reading, spelling, and comprehension, 2 years after the intervention's conclusion. Other recent multimedia-based studies with adult participants have found improved learning when attention-capturing signals such as animated cursors were replaced with images of the human hand pointing (e.g., Atkinson, Lin, & Harrison, 2009; Castro-Alonso, Ayres, & Paas 2014; de Koning & Tabbers, 2013), or an animated avatar pointing to specific aspects of on-screen information (Mayer & DaPra, 2012; Moreno, Reislein, & Ozogul, 2010).

Learning gains related to observing either an actual or virtual pointing hand also extend to pointing actions performed by the students themself. "Sandpaper Letters" have been used in Montessori schools for over a century, requiring students to trace letters cut out of sandpaper with their fingers in the same sequence as writing the letter. At the same time as tracing, students listen to the letter's sound as pronounced by their teacher (Montessori, 1912). Subsequent experimental investigations have supported the efficacy of this practice for letter learning and phoneme identification (e.g., Bara, Gentaz, & Colé, 2007; Bara, Gentaz, Colé, & Sprenger-Charolles, 2004; Hulme, Monk, & Ives, 1987) as well as recognition of geometrical shapes in kindergarten children (e.g., Kalenine, Pinet & Gentaz, 2011). Alibali and deRusso (1999) used a within-subjects design to test preschoolers' accuracy in counting presented objects across a range of conditions (no gesture, puppet pointing, child pointing, puppet touching, and child touching). Across these conditions, they found a strong positive gradient in counting accuracy. They speculated this pattern of results was the result of greater proximity of the finger to the object when touching rather than pointing, consistent with the basic research discussed above, and/or reduced working memory load by providing an external place-holder (the finger) in the set of counted objects.

Recently, Alibali (2005) and Ping and Goldin-Meadow (2010) have drawn on Kita's (2000) information-packaging hypothesis to explain results in gesturing studies. This hypothesis holds that "the production of a representational gesture helps speakers organize rich spatio-motoric information into packages suitable for speaking" (Kita, 2000, p.163). Drawing on this hypothesis, Ping and Goldin-Meadow (2010) suggested gestures "can provide an overarching framework that serves to organize ideas conveyed in speech, in effect chunking mental representations to reduce the load on working memory" (p.616). While the information packaging hypothesis has typically been applied in considerations of co-speech gestures' impacts on information processing and learning, it might also apply to situations where pointing gestures are used to trace out elements of spatially rich instructional materials, either in the air or on a surface. Cognitive load theory has historically discussed the prior knowledge base as the primary determinant of element interactivity and hence intrinsic cognitive load, through schemas' capacity to chunk multiple elements of information (Sweller et al., 2011). Consistent with the information-packaging hypothesis, tracing elements of instruction might likewise reduce intrinsic cognitive load.

Several studies have explored this possibility while drawing on cognitive load theory-based theorizing. Macken and Ginns (2014) found adult learners who traced elements of expository text and associated diagrams on the structure and function of the human heart outperformed learners who simply studied the materials visually on both terminology and comprehension tests. Importantly, the materials were designed to reduce sources of split attention (Ginns, 2006) wherever possible, meaning the above results were obtained from a relatively high baseline in instructional design quality. In a pilot study, Hu, Ginns, and Bobis (2014) investigated whether tracing the graphical elements of paper-based worked examples on parallel lines geometry would enhance problem solving. Again, the materials were explicitly designed to minimize sources of split attention through physical integration and color-coding of related referents. The tracing condition outperformed the non-tracing condition on acquisition phase problems and as well as errors during the test phase; however, ceiling effects precluded analysis of test performance. In a follow-up experiment, using more difficult test questions with a younger sample of participants, Hu, Ginns, and Bobis (2015) found the tracing group solved more test questions than the non-tracing group, solved them more quickly, made fewer errors, and reported lower levels of test difficulty. Test difficulty ratings were interpreted as an index of schema quality constructed during the acquisition phase. A subsequent study compared conditions where students studied the paper-based materials, studied the materials while tracing graphical elements in the air above the materials, or studied the material while tracing graphical elements on the surface of the paper. Hypothesized gradients of performance (no tracing < tracing in the air < tracing on the paper) were found for acquisition phase problem-solving and test phase basic and advanced question problem-solving. Hypothesized gradients of performance (no tracing > tracing in the air > tracing on the paper) were also found for acquisition phase errors, and time to solution, error rate, and test question difficulty ratings for advanced questions in the test phase. Hu, Ginns, and Bobis (2015) interpreted these results as indicating that the more working memory modalities (visual, tactile and kinesthetic) activated during learning, the better the problem-solving schemas constructed, especially since the strongest results were found on advanced test questions requiring transfer from the worked examples presented (cf. Cooper & Sweller, 1987).

Rationale of current study

The tracing effect found across the above studies provides the foundation for the current experiment. While the above findings provide initial support for a tracing effect, they are limited in several ways. First, all of the above findings were generated through interactions with paper-based instructional materials, meaning it is as yet unclear whether the results generalize to interactions with other interfaces used for learning, such as computers, tablets or interactive whiteboards. In particular, the ubiquity of touch-based interactions while students use tablets to learn poses a fundamental question: do these interactions enhance learning over and above that which would be achieved simply through reading from the screen? Secondly, evidence for a cognitive load-based explanation of the learning gains found in the above studies is mixed. Macken and Ginns (2014) did not find statistically reliable differences between conditions on single-item ratings of intrinsic, extraneous and germane cognitive load. Hu, Ginns, and Bobis' (2015) use of test phase difficulty ratings rather than acquisition phase difficulty or mental effort ratings limits the claims that can be made about tracing's effects on cognitive load while learning. The present experiment was thus designed to test the following hypotheses:

Hypothesis 1: participants who trace elements of worked examples while studying will solve more problems than those who study by viewing only, reflecting more effective schemas.

Hypothesis 2: participants who trace elements of worked examples while studying will solve problems more quickly than those who study by viewing only, reflecting more efficient schemas.

Hypothesis 3: participants who trace elements of worked examples while studying will rate test problems as lower in difficulty than students who study by viewing only, reflecting lower intrinsic cognitive load during problem-solving.

Hypothesis 4: participants who trace elements of worked examples while studying will rate worked examples as lower in difficulty than students who study by viewing only, reflecting lower intrinsic cognitive load during acquisition.

Method

Participants and Design

Sixty-one participants participated in the experiment. The research study was conducted in accordance with the ethical standards of the University of Wollongong. Participants were from Years 3 and 4 (an age range from 8 to 11 years) in a primary school located south of Sydney, NSW, Australia. One Year 3 class (12 girls and 14 boys) and two Year 4 classes (17 girls and 18 boys) participated in the experiment. Parental/carer consent to participate was obtained. Upon receipt of the consent forms, participants were randomly assigned to one of two conditions:

Tracing condition: Participants were instructed to trace information that was highlighted and circled on an iPad app using their index finger (an example is provided in Figure 1).

Non-tracing condition: Participants were instructed to look at the information that was highlighted and circled on an iPad app (an example is provided in Figure 2).

Materials

The topic covered in the experiment was learning how to read single and double lined temperature line graphs. Materials were based on the Australian National Curriculum Stage 3 (Years 5-6) Mathematics Syllabus Data Units 1 and 2 (Board of Studies NSW Syllabuses for the Australian curriculum, 2012). The specific outcome directly addressed by this topic is: "uses appropriate methods to collect data and constructs, interprets and evaluates data displays, including dot plots, line graphs and two-way tables" (Outcome MA3-18SP).

The topic of temperature line graphs was selected because it was a topic from Stage 3 curriculum that the participants would be taught, therefore the intrinsic cognitive load of this content was considered high for Years 3 and 4 (8-11 years) students and it was determined from the participants' teachers that the participants had limited knowledge about reading temperature line graphs before the experiment.

The temperature line graph instructional materials were designed based on previous research conducted by Leahy, Chandler and Sweller (2003); Leahy and Sweller (2004, 2005, 2011); and Wong, Leahy, Marcus, and Sweller (2012). The information was considered to be high in element interactivity as there was a high degree of interrelatedness between the different elements the participants were required to comprehend to read the temperature line graphs. All instructions and materials were delivered in the form of an iPad app. Two iPad apps were developed (one for each condition). Each app consisted of 50 slides comprising an introduction to temperature line graphs and the *x* and *y* axes followed by three worked examples explaining how to find the temperature at a particular time of day, followed by a series of ten test questions. A cognitive load Likert rating scale, based on the work by Paas (1992), was also included for participants to rate their perceived difficulty in understanding the information and answering the questions.

The first worked example explained how to find what the temperature was at a certain time of day on a single-lined temperature graph (see Figure 1). The x-axis represented the time of day and the y-axis represented the range in temperature. The problem: “Let’s find the temperature at 10am” was posed. For novices, this is considered high in element interactivity because each word and its associated reference (elements) must be comprehended in relation to the other words and the graph. Knowledge elements such as "10am" are represented by a line, with a dot matching the temperature and the intersection between an x- and y-axis (which are also related to other elements of time and temperature) are elements that interact and must be fully grasped before a correct answer is possible. In addition there is a set procedure to correctly find or comprehend the answer.

The second worked example explained how to find what was the time of day for a certain temperature on the same single-lined temperature graph as illustrated in Figure 1. This is considered high in element interactivity for novices, as it is similar to the first worked example. The third worked example introduced a double temperature line graph and explained how to find what was the temperature on a particular day at a certain time (see Figure 2). This worked example was more complex than the previous two worked examples, demonstrating increased element interactivity.

Procedure

Participants were individually tested, and the experiment took approximately 25 minutes. The experiment began with an introductory phase, followed by an acquisition phase and concluded with a test phase (detailed below). A perceived difficulty rating followed each worked example in the acquisition phase and each question in the test phase. Due to the participants’ age, the researcher read all the text to ensure the participants understood the information and instructions.

Introductory Phase: The researcher started the iPad app and entered a participant code. The first slide was untimed and explained the focus of the "lesson" and what the participant had to do based on the condition assigned. Participants touched (using their index finger) the icon labeled "Next" to proceed to slide two. The second slide (untimed) introduced participants to a cognitive load Likert rating scale. The rating scale was used to measure participants' perceived ease or difficulty in understanding the worked examples and answering the problems in the test phase. The scale consisted of 1, *very easy*, to 5, *very difficult.* This five-scale rating is a modified version of the nine-scale rating developed by Paas (1992) and is similar to that used in Hu, Ginns, and Bobis (2015). To ensure the participants understood the cognitive load rating scale, the researcher asked the participant to try and solve two problems; one perceived easy and one perceived difficult (e.g. 2+7; 3x4x5) and identify the rating scale score that best reflected their perceived ease or difficulty in solving each problem. The researcher made sure the participant understood the cognitive load rating scale before proceeding with the experiment.

Acquisition phase*:* Slides 3 to 5 introduced the basic elements of single-lined temperature line graphs such as points on a graph indicating temperature at different times of the day, x-axis displaying the time of day and the y-axis displaying the temperature. Participants in the trace condition were asked to trace the x- and y-axes whereas participants in the non-trace condition were asked to look at the highlighted x- and y-axes.

Slides 6 to 12 comprised the first worked example and explained the steps required to calculate what was the temperature at 10am on a single-lined temperature line graph. Figure 1 represents the last slide in this worked example. The first six slides were timed for 20 seconds. Participants in the trace condition traced the highlighted circles and lines whereas participants in the non-trace condition looked at the highlighted information. Each step was shown on a single slide. The instructions presented in the Trace condition for the first worked example is as follows:

Slide 6: “Let’s find the temperature at 10am.”

Slide 7: “Firstly *look* at 10am circled on the x-axis and then *trace* the circle around 10am”.

Slide 8: “Now *look* at the orange line drawn up from 10am. *Trace* this line up from 10am to see how it meets the temperature line.”

Slide 9: “Now *look* how the orange line is drawn from the temperature line to y axis. *Trace* this line to see how it meets the y axis.”

Slide 10: “*Look* how 30º is circled. *Trace* the circle around 30º to see that at 10am the temperature is 30º.”

Slide 11: “Starting at 10am, trace the circles and lines again to understand that at 10am the temperature is 30º.”

Participants were then asked to circle the number on the 5-point Likert cognitive load rating scale that indicated their perceived ease or difficulty to understand how to find the temperature. They then touched the icon labeled “Next” to proceed to the next worked example.

Slides 13 to 19 comprised the second worked example showing each step required to calculate the time of day when the temperature was 34° on a single-lined temperature line graph, followed by the cognitive load rating scale (similar to the first worked example). Slides 20 to 29 presented the third worked example that introduced a double lined temperature graph displaying temperatures for two days (see Figure 2). Participants were showed the steps for calculating the temperature for a specific day - Tuesday at 1pm. As per the previous worked examples, participants completed a rating scale indicating how easy or difficult it was to understand how to find the temperature on a specific day.

Test Phase: Immediately after completing the acquisition phase, participants proceeded to the test phase where they were required to answer ten questions based on their understanding of temperature line graphs (slides 30 to 50). Participants were given a maximum of 2 minutes to answer each question. Each question was presented on one slide and when participants answered each question they then touched the icon labeled "Next" to proceed to the next slide. If a question was not answered within two minutes, the iPad app automatically proceeded to the next slide. Following each test question, participants were asked to circle the number on the 5-point Likert cognitive load rating scale that indicated their perceived ease or difficulty to answer the question.

The first three questions were based on the information provided in the acquisition phase. The questions were considered similar questions as the problems had exactly the same structure and required the same procedural steps to solution as presented in the acquisition phase. The first question required participants to draw on a temperature line graph to show the x-axis. The second question required participants to draw on a temperature line graph to show the y-axis. The third question asked participants to calculate the temperature when the time of day was 1pm. Questions 4 to 10 required participants to transfer their understanding of temperature line graphs to questions that were different in structure or steps to solution to the worked example problems provided in the acquisition phase. Questions 4 and 5 required participants to find the highest and lowest temperature on a single-lined temperature line graph. Question 6 asked participants to find the time of day for a specified temperature on a double lined (two-day) temperature line graph, this procedure was presented on a single but not a double line graph. Question 7 required participants to calculate what day the temperature did not go below a certain degree for a two-day temperature line graph. Question 8 provided a two-day temperature line graph and required participants to calculate between a certain time frame, what day had the highest temperature and what it was. Question 9 presented a two-day temperature line graph and required participants to state what days and the times the temperature was 30. The last question represented a far transfer task where a two-line graph displayed the number of soccer goals two children kicked over a number of months, participants had to state who kicked the most goals in May. Questions 1 to 8 were each marked out of 1. Question 9 was marked out of 3 as there were six responses required with each answer scoring 0.5. Question 10 was marked out of 1. In summary, the maximum total test score was 12: total test score for similar problems (Questions 1 to 3) was 3 and the total test score for transfer questions (Questions 4 to 10) was 9. The answer to each question, the time to complete each question, and the cognitive load rating scale response for each worked example and test question was recorded by the iPad app. The iPad app recorded the participant's finger movement for each slide as a screen shot. The screen shots were examined during analysis to check that the participants (in the trace condition) had complied with the trace instructions.

Results

The data was first screened to evaluate trace instruction compliance, normality, and homogeneity of variance. Examination of compliance in the trace condition with 'trace' requirements indicated at least 80% of participants complied with all trace instructions. There were no participants in the trace condition who did not comply with any aspect of a worked example. Shapiro-Wilk statistics identified skewness in the distribution of test scores, although skewness was not extreme (as evidenced by z skewness statistics less than 4). Levene's test showed that the variances were equal for proportional accuracy scores for both conditions: Total, $F = .021$, $p = .886$; Similar, $F = 1.593$, $p = .212$; Transfer, $F = .057$, $p = .812$. Nevertheless, a non-parametric Mann-Whitney U test was run to evaluate potential differences in patterns of significance. Given that patterns of significance did not differ across these analyses, the non-extreme nature of the skewness, and equal variances for proportional accuracy scores across both conditions, the parametric analyses are presented. Table 1 presents descriptive statistics for total raw test scores, proportional accuracy scores and response times for similar and transfer questions. Table 2 presents the cognitive load ratings for the worked examples and test questions.

	Accuracy			Response Time		
	Total	Similar	Transfer	Total	Similar	Transfer
Condition	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Non-Trace (N=31)	7.50 (2.40)	73.12 (31.53)	58.97 (22.21)	31.10 (7.08)	17.70 (6.52)	36.84 (8.78)
Trace (N=30)	9.07 (2.42)	82.22 (24.34)	73.33 (22.57)	31.38 (6.85)	16.00 (4.86)	37.97 (8.56)

Table 1: Total raw test scores, proportional accuracy scores and response times for similar and transfer questions. *Note*. Given that similar questions (out of three possible marks) and transfer questions (out of nine possible marks) are differently weighted, proportional accuracy scores are presented for these questions to facilitate comparison. Total scores, in contrast, reflect total raw scores for similar and transfer questions combined (maximum score of 12). Ceiling effects were

not found; only 8% of participants (5 of 61) scored 100%.

	Worked Examples	Test Questions		
	Total	Total	Similar	Transfer
Condition	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Non-Trace (N=31)	1.62 (0.58)	1.83 (0.61)	1.45 (0.55)	1.99 (0.71)
Trace (N=30)	1.53 (0.60)	1.81 (0.62)	1.44 (0.56)	1.97 (0.70)

Table 2: Cognitive load ratings for the worked examples and test questions. *Note*. Two participants did not complete the cognitive load rating scale for one test question.

A 2 (Instructional Condition: trace vs. non-trace) x 2 (Type of Test: similar vs. transfer) mixed analysis of variance (ANOVA) with type of test as the repeated measure was performed on the following dependent variables: proportional accuracy scores, and response times for similar and transfer questions, and cognitive load ratings for the worked examples and test questions. An ANOVA on proportional accuracy scores identified main effects for condition, $F(1, 59) = 4.99$, $p = .029$, partial $\eta^2 = .08$, and question type, $F(1, 59) = 8.88$, $p = .004$, partial $\eta^2 = .13$, but no significant interaction, $F(1, 59) = 0.46$, $p = .498$, partial $\eta^2 = .01$. The main effect of instructional condition indicated that the trace condition outperformed the non-trace condition. The main effect of type of test indicated that performance was higher on similar than on transfer questions.

Despite the non-significant interaction, examination of the data suggested the possibility of differential benefits of tracing on similar and transfer questions. As such, independent-samples t-tests were run on similar and transfer scores. Results of these analyses indicated a significant effect of condition for transfer questions, $t(59) = -2.51$, $p = .015$, $\eta^2 = .10$. Performance did not differ on similar questions, $t(59) = -1.26$, $p = .213$, $\eta^2 = .03$. These results suggest that tracing only benefited performance on transfer questions..

The 2 x 2 ANOVA on the average time to complete similar and transfer questions (response time) revealed a main effect for type of test, $F(1, 59) = 385.27$, $p < .001$, partial $\eta^2 = .87$, but not instructional condition, $F(1, 59) = 0.03$, $p = .857$, partial $\eta^2 < .01$. The main effect of type of test indicated that similar questions were responded to more quickly than transfer questions (see Table 1). There was no significant interaction conditioning these effects, $F(1, 59) = 1.83$, $p = .182$, partial $\eta^2 = .03$.

The 2 x 2 ANOVA conducted on the average cognitive load rating for similar and transfer questions revealed a main effect for type of test, $F(1, 59) = 52.32$, $p < .001$, partial $\eta^2 = .48$, but not condition, $F(1, 59) = 0.04$, $p = .850$, partial $\eta^2 < .01$. The main effect for type of test indicated that the cognitive load ratings for similar questions were lower than for transfer questions (see Table 2). There was no significant interaction conditioning these effects, $F(1, 59) = 0.02$, $p = .890$, partial $\eta^2 < .01$.

To investigate the effects of condition on cognitive load ratings of the worked examples provided in the instruction phase, an independent-samples t-test was run on these cognitive load ratings. Analyses indicated no significant difference in the average cognitive load rating on worked example, $t(59) = 0.59$, $p = .556$, $\eta^2 < .01$. This was also the case for cognitive load ratings for each worked example separately (all ps < .05). Inspection of the distribution of responses suggested floor effects on this variate.

Discussion

The purpose of this study was to investigate whether the act of tracing on an iPad when learning about temperature line graphs would be more effective for learning than looking at the same content on the screen. The results showed that children who were instructed to trace on the temperature line graphs while studying them achieved a higher performance on transfer test problems than children who studied the same materials without tracing. This confirmed our first hypothesis. In terms of response time to answer the test questions, it was found that overall the similar questions were answered more quickly than the transfer questions but there was no significant difference between the conditions, thus our second hypothesis was not confirmed. It was expected that participants in the trace condition would rate the test problems as lower in difficulty than participants in the non-trace condition. This hypothesis (H3) was not confirmed but the cognitive load ratings results showed that overall the cognitive load ratings for similar questions were lower than for the transfer questions. It was also expected that participants in the trace condition would rate the worked examples as less difficult than those in the non-trace condition. The findings showed no significant difference in the average cognitive load rating for the worked examples, so this hypothesis (H4) was also not confirmed.

This study provides further evidence that builds on current research (Macken & Ginns, 2014; Hu, Ginns & Bobis, 2015) about the effectiveness of biologically primary knowledge, e.g., tracing, to support the understanding of biologically secondary knowledge, e.g., mathematical knowledge, as the findings showed that tracing increased test performance on transfer questions. There was, however, no comparable pattern of results for similar questions or response time. This may have been due to the explicitness of the worked examples, whereby each step to solution was presented on one slide. This may have supported all participants to develop basic temperature graph schematic knowledge, as demonstrated by non-significant differences on similar questions. The explicit step-by-step instructions may have diminished the effectiveness of tracing. It is conjectured that if participants were presented with an initial instruction phase, where all the steps to solution in a worked example is explained first, followed by an acquisition phase, where participants are then to repeat the steps by either tracing on, or looking at, the temperature line graph, (similar to the research design reported in Hu, Ginns and Bobis, 2015), the effect of tracing may have been more evident on similar questions. The relationship between the amount of explicit instruction and the effectiveness of tracing is an important aspect that requires further exploration.

This study did not show evidence for a subjective cognitive load-based explanation of the learning gains found in the trace condition. The results showed that irrespective of condition, the participants rated the transfer test questions slightly higher in difficulty than the similar questions. The lack of significance on these subjective cognitive load ratings raises questions about how to effectively collect subjective cognitive load rating data from young participants. In this experiment the cognitive load rating scale was explained to each participant by presenting two examples; each representing the opposite ends of the rating scale. The researcher ensured each participant understood the cognitive load rating scale before proceeding with the experiment. A possible reason for the non-significant result may be attributed to the participants' age and the lack of a visual cue. For example, in the study by Hu, Ginns, and Bobis (2015) where the tracing group reported lower levels of test difficulty than the non-tracing group, the participants were Year 5 primary students whereas in this study the participants were Years 3 and 4 primary students. This may have influenced the participants' ability to reflect on perceived task difficulty.

Futhermore, the 5-point rating scale in Hu et al. (2015) included two cartoon-face icons positioned above the 1 and 5. A smiling face and the text label 'very easy' was positioned above the 1, a frowning face and the text label 'very difficult' was positioned above the 5, and the anchors 2,3,4 appeared without additional text labels or icons. For this experiment the 5-point rating scale was provided with text labels for each anchor. The visual cue may have helped participants to rate their perceived level of ease/difficulty. Cognitive load rating scales for use by young children thus needs further investigation. Possible alternative measurements of young children's cognitive load need to be investigated. One possible alternative could be exploring other mental load rating mechanisms, such as Leppink, Paas, Van der Vleuten, Van Gog & Van Merrienboer (2014).

Another limitation in this study was whilst the data collected from the screen shots were examined to check that the participants in the trace condition had performed the required trace instructions, a similar check of compliance was not possible for the non-trace condition as evidence as to whether the participants actually viewed and looked at each screen as per the instructions, was not possible to collect. Testing was conducted one-on-one with a researcher and this enabled a check on overall participant engagement and focus on the instructional materials. One suggestion of how this could be considered in future studies is by utilizing eye-tracking technology to check for compliance of participants looking at the highlighted lines and circles of the worked examples. Due to this limitation, it is unclear, if participants in the non-trace condition attended to the highlighted parts of the worked examples.

To extend this current research, future studies could include a delayed test phase in addition to the immediate test phase as conducted in this study, to examine the robustness of a tracing effect. As iPads are increasingly being used in classrooms and thus it is important to investigate the efficacy of tracing as an instructional technique within naturalistic settings. In this current study, participants were tested individually and thus were requested to leave the classroom for approximately 25 minutes to partake in this experiment. The topic selected for this research was directly related to the school curriculum thus future studies could focus on curriculum specific topics and be implemented in a naturalistic classroom-based environment such that the acquisition and testing phases could take place in a whole class environment with participants each using an iPad. Previous cognitive load effects have shown that as expertise increases, the effectiveness of an instructional technique can diminish (Kalyuga, 2007). This research study showed the benefits of tracing (biologically primary knowledge) to support novice problem solving. Of interest is the possible interaction between biologically primary knowledge (pointing, tracing) and expertise. Future research may explore the proposition that as expertise increases the benefits of utilizing biologically primary knowledge (pointing, tracing) decreases.

In order to further examine whether the positive effect of tracing was due conclusively to the act of gesturing, a future study could include an additional condition where participants watch an animated hand with a pointed finger conduct the tracing. In this study, tracing may have been beneficial because it forced attention to the worked examples in a way that viewing/looking cannot, thus, including an additional condition that serves as an attention cue, could help to isolate these variables.

Thus, while this experiment provided some evidence of finger tracing positively effecting learning, more conceptual work that can elaborate and extend our previous understanding on the impact of hand gesture or gesture-controlled devices on learning is still needed.

Conclusion

The results reported in this paper adds to a growing evidence base that finger tracing (a form of biologically primary knowledge) can support problem solving (a form of biologically secondary knowledge) as learners who traced tablet-based worked examples about temperature line graphs outperformed learners who only looked at these worked examples on transfer test questions. Where previous research has investigated the effectiveness of tracing worked examples using paper-based materials, this current work has explored the effectiveness of tracing on a tablet device to support mathematical problem solving. With the increasing use of tablet technology in classrooms there is a need to provide evidence-based research to inform teaching practices. As such, the results from this experiment provide some evidence for the use of tracing. A practical implication that can be inferred from this study is, give learners a helping hand by encouraging finger tracing when learning mathematical concepts using a tablet device.

References

Abrams, R., Davoli, C., Du, F., Knapp III, W., & Paull, D. (2008). Altered vision near the hands. *Cognition*, *107*,1035-1047.

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, *5*, 307-331.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*, 37-56.

Atkinson, R., Lin, L., & Harrison, C. (2009). Comparing the efficacy of different signaling techniques. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 954-962). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/31608

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Bara, F., Gentaz, E., & Colé, P. (2007). Haptics in learning to read with children from low socio-economic status families. *British Journal of Developmental Psychology*, *25*, 643-663.

Bara, F., Gentaz, E., Colé, P., & Sprenger-Charolles, L. (2004). The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetic principle. *Cognitive Development*, *19*, 433-449.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Board of Studies NSW Syllabuses for the Australian curriculum (2012). *Mathematics K-10, Stage 3, Statistics and Probability*. Retrieved March 20, from http://syllabus.bos.nsw.edu.au/mathematics/mathematics-k10/content/1147/

Castro-Alonso, J. C., Ayres, P., & Paas, F. (2014). Learning from observing hands in static and animated versions of non-manipulative tasks. *Learning and Instruction*, *34*, 11-22.

Chum, M., Bekkering, H., Dodd, M. D., & Pratt, J. (2007). Motor and visual codes interact to facilitate visuospatial memory performance. *Psychonomic Bulletin & Review*, *14*, 1189-1193.

Colonnesi, C., Stams, G.J.J.M., Koster, I., & Noom, M.J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, 352–366.

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

Cosman, J., & Vecera, S. (2010). Attention affects visual perceptual processing near the hand. *Psychological Science*, *21*, 1254-1258.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87-185.

de Koning, B., & Tabbers, H. (2013). Gestures in instructional animations: A helping hand to understanding non-human movements? *Applied Cognitive Psychology*, *27*, 683-689.

Docebo (2014). *E-Learning Market Trends & Forecast 2014 - 2016 Report*. Retrieved from http://www.docebo.com/elearning-lms-resources/papers-researches/

Evans, M.A., Williamson, K., & Pursoo, T. (2008). Preschoolers' attention to print during shared book reading. *Scientific Studies of Reading*, *12*, 106-129.

Fischer, M. H., & Hoellen, N. (2004). Space- and object-based attention depend on motor intention. *Journal of General Psychology*, *131*, 365-377.

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*, 319-325.

Geary, D. C. (2008). An evolutionarily informed education science. *Educational Psychologist*, *43*, 179-195.

Ginns, P. (2006). Integrating information: Meta-analyses of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, *16*, 511-525.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the Revolution to Embodiment: 25 Years of cognitive psychology. *Perspectives on Psychological Science*, *8*, 573-585.

Goldin-Meadow, S. (2011). Learning through gesture. *WIREs Cognitive Science*, *2*, 595-607.

Hostetter, A. B., (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*, 297-315.

Hu, F.-T., Ginns, P., & Bobis, J. (2014). Does tracing worked examples enhance geometry learning? *Australian Journal of Educational & Developmental Psychology*, *14*, 45-49.

Hu, F.-T., Ginns, P., & Bobis, J. (2015). Getting the point: Tracing worked examples enhances learning. *Learning and Instruction*, *35*, 85-93.

Hulme, C., Monk, A., & Ives, S. (1987). Some experimental studies of multi-sensory teaching: The effects of manual tracing on children's paired-associate learning. *British Journal of Developmental Psychology*, *5*, 299-307.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). *NMC Horizon Report: 2014 K-12 Edition*. Austin, Texas: The New Media Consortium

Justice, L.M., Kaderavek, J.N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers'early literacy development through classroom-based teacher-child storybook reading and explicitprint referencing. *Language, Speech, and Hearing in Schools*, *40*, 67-85.

Kalenine, S., Pinet, L., & Gentaz, E. (2011). The visual and visuo-haptic exploration of geometrical shapes increases their recognition in preschoolers. *International Journal of Behavioral Development*, *35*, 18-26.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*, 509-539.

Keselman, H. J., Othman, A.R., & Wilcox, R.R. (2013). Preliminary testing for normality: Is this a good practice? *Journal of Modern Applied Statistical Methods*, 12, Article 2. Retrieved from http://digitalcommons.wayne.edu/jmasm/vol12/iss2/2

Kendon, A. (2004). *Gesture: Visible action as utterance.* Cambridge: Cambridge University Press.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, *21*, 31–42.

Kirschner, F., Paas, F., & Kirschner, P. A. (2011). Task complexity as a driver for collaborative learning efficiency: The collective working memory effect. *Applied Cognitive Psychology*, *25*, 615–624.

Kita, S. (2001). How representational gestures help speaking. In D. McNeill (Ed.), *Language and Gesture*. Cambridge, UK: Cambridge University Press.

Leahy, W., Chandler, P., & Sweller, J. (2003). When auditory presentations should and should not to be a component of multimedia instruction. *Applied Cognitive Psychology, 17*(4), 401-418.

Leahy, W., & Sweller, J. (2004). Cognitive load and the imagination effect. *Applied Cognitive Psychology, 18*(7), 857-875.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*(6), 943-951.

Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology: Applied, 11*(4), 266-276.

Leppink, J., Paas, F., Van Der Vleuten, C.P., Van Gog, T., & Van Merrienboer, J.J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30,* 32-42.

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & de Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cognitive Science*, *36*, 698-713.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*, 297-307.

Macken, L., & Ginns, P. (2014). Pointing and tracing gestures may enhance anatomy and physiology learning. *Medical Teacher, 36*, 569-601.

Marshall, S.P. (1995). *Schemas in problem solving.* Cambridge: Cambridge University Press.

Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied*, *18*, 239-252.

Mehta, C. R., & Patel, N. R. (2012). *IBM SPSS exact tests*. Retrieved from ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/21.0/en/client/Manuals/IBM_SPSS_Exact_Tests.pdf

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

Montessori, M. (1912/1964). *The Montessori method.* Cambridge, MA: Robert Bentley, Inc.

Moreno, R., Reislein, M., & Ozogul, G. (2010). Using virtual peers to guide visual attention during learning. *Journal of Media Psychology, 22*, 52-60.

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

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, *24*, 27-45.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193-198.

Piasta, S. B., Justice, L. M., McGinty, A. S., & Kaderavek, J. N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development*, *83,* 810-820.

Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science*, *34*, 602-619.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169-192.

Sheu, F-R., & Chen, N-S. (2014). Taking a signal: A review of gesture-based computing research in education. *Computers in Education*, *78*, 268 – 277.

Shams, L., & Seitz, A. (2008). Benefits of multisensory learning. *Trends in Cognitive Science*, *12*, 411-417.

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Science*, *1191*, 182-200.

Spitzer, M. (2013). To swipe or not to swipe? – The question in present day education. *Trends in Neuroscience and Education*, *2*, 95-99.

Sweller, J. (2003). Evolution of human cognitive architecture. *Psychology of Learning and Motivation*, *43*, 216-266.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*, 400-410.

Wong, A., Marcus, N., Smith, L., Cooper, G. A., Ayres, P., Paas, F., et al. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior*, *25*, 339–347.

Van der Burg, E., Olivers, C., Bronkhorst, A., & Theeuwes, J. (2009). Poke and pop: Tactile-visual synchrony increases visual saliency. *Neuroscience Letters*, *450*, 60-64.

Van Gog, T., Ericsson, K. A., Rikers, R. M., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development*, *53*, 73-81.

Van Gog, T., Paas, F., Marcus, N., Ayres, P., & Sweller, J. (2009). The mirror-neuron system and observational learning: Implications for the effectiveness of dynamic visualizations. *Educational Psychology Review*, *21*, 21–30.

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449-457.

Acknowledgement

This research was funded by the University of Wollongong's Research Council Small Grant Scheme, 2013.

Table 1

	Accuracy			Response Time		
	Total	Similar	Transfer	Total	Similar	Transfer
Condition	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Non-Trace (N=31)	7.50	73.12	58.97	31.10	17.70	36.84
	(2.40)	(31.53)	(22.21)	(7.08)	(6.52)	(8.78)
Trace (N=30)	9.07	82.22	73.33	31.38	16.00	37.97
	(2.42)	(24.34)	(22.57)	(6.85)	(4.86)	(8.56)

Table 1: Total raw test scores, proportional accuracy scores and response times for similar and transfer questions. *Note*. Given that similar questions (out of three possible marks) and transfer questions (out of nine possible marks) are differently weighted, proportional accuracy scores are presented for these questions to facilitate comparison. Total scores, in contrast, reflect total raw scores for similar and transfer questions combined (maximum score of 12). Ceiling effects were not found; only 8% of participants (5 of 61) scored 100%.

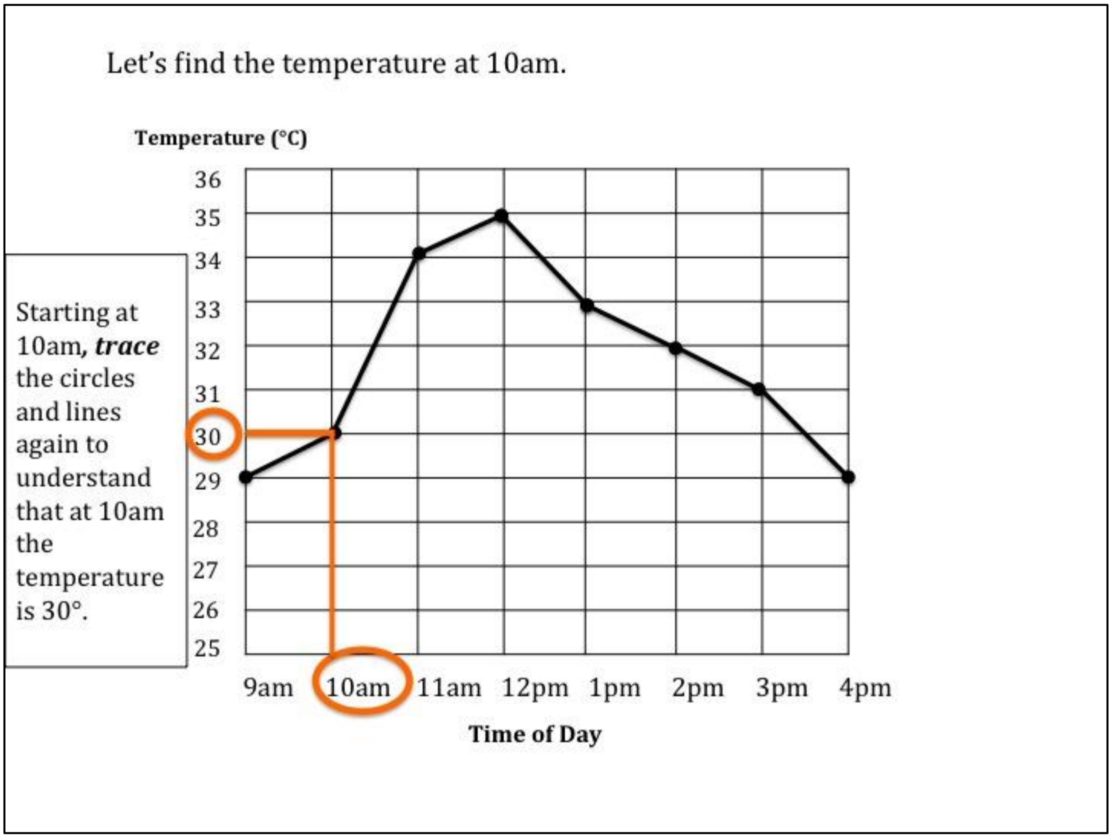

Figure 1: Trace condition – Last screen of the first worked example

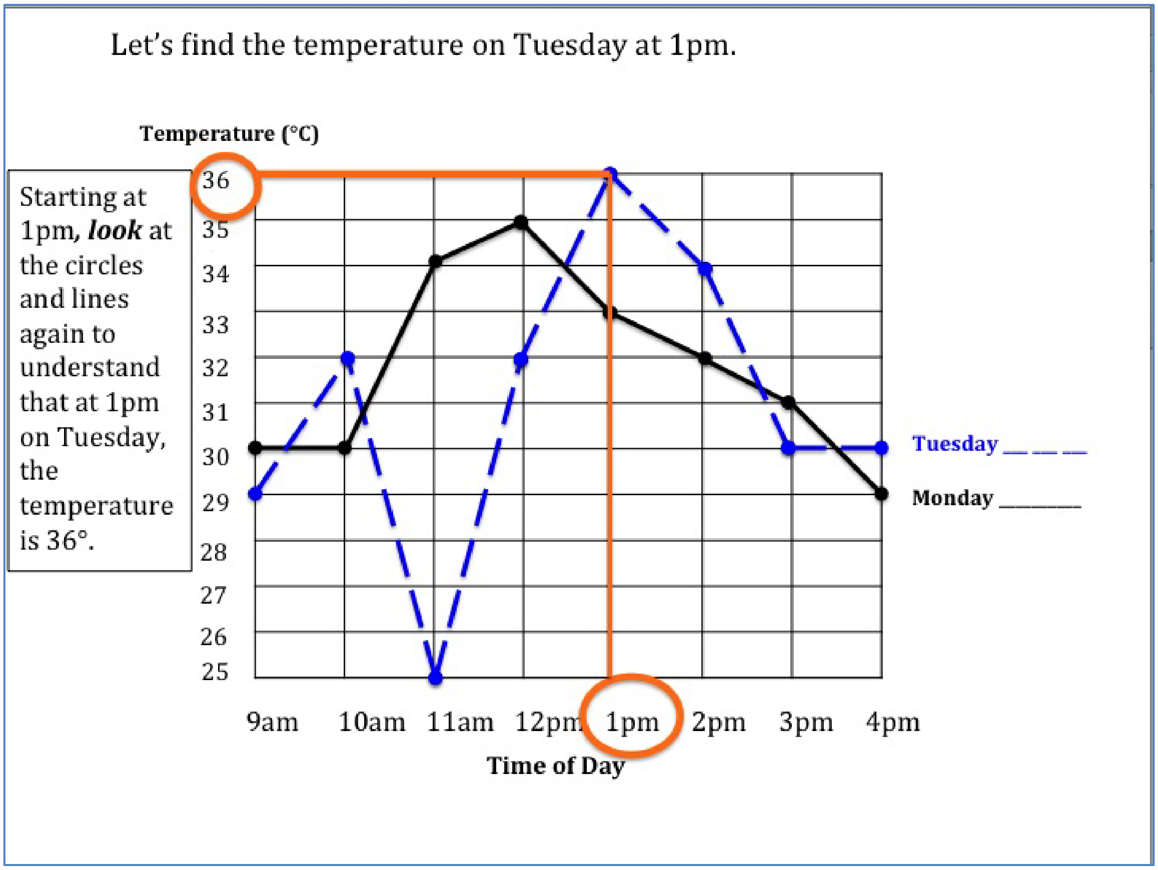

Figure 2: Non-Trace condition – Last screen of the third worked example