INTERVENTION STUDY

CrossMark

Tracing Enhances Recall and Transfer of Knowledge of the Water Cycle

Michael Tang[1] · Paul Ginns[1] (iD) · Michael J. Jacobson[1]

Abstract
Cognitive load theory has incorporated evolutionary perspectives to consider how *biologically primary knowledge* (such as physical movement and pointing), acquired through evolutionary processes, might support the acquisition of *biologically secondary knowledge* (such as reading or writing), requiring explicit teaching. Tracing (a physical movement) against a surface with the index finger may be one form of biologically primary knowledge that can enhance learning biologically secondary knowledge. We investigated whether tracing lesson materials (about the water cycle) presented on A4 pieces of paper in an initial phase, then on an A1 poster in a subsequent phase, would affect primary school students' reports of intrinsic versus extraneous cognitive load, as well as recall and transfer test performance. Students who traced while studying reported lower extraneous cognitive load than those who simply studied and scored higher on subsequent recall and transfer tests. Considerations for instructional designers, educators and researchers are discussed.

Keywords Cognitive load theory · Biologically primary knowledge · Tracing · Recall · Transfer

Cognitive load theory (CLT; Sweller 2011; Sweller et al. 2011) is concerned with the implications of human cognitive architecture for instructional design. This architecture includes a working memory system whose limits in capacity and duration (Miller 1956; Peterson and Peterson 1959) are often substantially reduced when task-specific problem-solving schemas (Marshall 1995) are cued by environmental signals and are retrieved from long-term memory. A range of variables affecting optimal use of working memory resources have been investigated by CLT researchers, and by targeting instructional sources of intrinsic, extraneous and germane load (Kalyuga 2011; Sweller 2010; Sweller et al. 1998), a variety

✉ Paul Ginns
paul.ginns@sydney.edu.au

[1] Sydney School of Education and Social Work, The University of Sydney, Sydney, NSW 2006, Australia

🍎 Springer

of cognitive load effects enhancing the effectiveness and efficiency of instructional materials have been generated.

Inspired by Geary's (2008) evolutionary educational psychology, CLT has incorporated an evolutionary view of human cognitive architecture (Sweller 2011). Geary (2008) proposed that human evolution has produced a range of cognitive modules with associated information-processing domains; such domains are "biased to process specific forms of information, and … predicted to be plastic to the extent that sensitivity to variation in the corresponding information patterns provides an advantage" (p. 180). Based on this evolutionary perspective, concepts regarding the most fundamental categorization of knowledge have been introduced, namely biologically primary knowledge and biologically secondary knowledge.

Biologically primary knowledge refers to abilities that humans have evolved to acquire without much mental effort. A commonality shared between these abilities is that no explicit instruction is required for most people to learn this category of knowledge, nor could such abilities be taught. For example, according to Geary's (2008) hypothesized 'Individual' module, people acquire the capacity to produce and interpret non-verbal behaviours such as gestures automatically through interacting with other people, rather than being taught explicitly. Humans may have presumably evolved to acquire these primary skills over many generations in order to survive and reproduce in an ever-changing natural environment (Sweller 2011).

In contrast, biologically secondary knowledge refers to abilities that humans have not evolved to acquire based on effortless natural processes. Skills such as reading or writing, or scientific knowledge, could not be learned via simple exposure to one's environment, nor without explicit teaching. Such knowledge is held to be culturally important but also recent in its development when compared to human evolutionary history. Because of the complexity of such knowledge, effortful and deliberate practice is required to acquire such knowledge, and educational institutions have been invented for the purpose of acquiring such knowledge by providing explicit curricula (Geary 2008).

Since humans are able to easily acquire biologically primary knowledge without specific instructions, it seems that the acquisition of this category of knowledge is not bounded by the limitations of working memory capacity (Sweller 2008). In other words, when learning is associated with biologically primary knowledge, the conscious cognitive resources required are negligible. Given such an advantage in working memory management, using the primary system to facilitate the learning of secondary knowledge could be a way to reduce the load imposed on working memory while learning (Paas and Sweller 2012). The immediate question raised following such an assumption is whether those primary systems can be used to facilitate the acquisition of secondary knowledge. In this study, we investigate how tracing gestures, as a type of biologically primary knowledge, might enhance learning about a physical process, the water cycle.

Pointing and Tracing as Biologically Primary Knowledge

The pointing gesture, defined as static deictic pointing in space (cf. McNeill 1992), has long been discussed in association with attentional processes; for example, Vygotsky (2004) argued that an infant's pointing gesture marks the beginning of voluntary attention development, and such attention allows the formation of concepts by associating the attended item with meaning. Early eye-tracking research by Steinbach and Held (1968) found more precise visual tracking

of a target mounted on a fingertip when the observer voluntarily moved his/her arm, compared to when the observer relaxed the arm and the arm was moved by the experimenter. Steinbach and Held speculated that "the ability to keep an image of a manipulated object on the area of sharpest vision on the retina is highly adaptive" (p. 188), an early example of evolutionary theorizing linked to pointing gestures.

Pointing has also been recognized as a means to communicate in early stages of language development and appears to be the most primitive mode of communication prior to the development of language. In infants raised in normal social milieus, pointing gestures generally emerge prior to the first birthday. This observed phenomenon, known as *protodeclarative pointing* (see Bates et al. 1975; Camaioni 1993), occurs when infants show their understanding of intentions, attention and knowledge of their communicative partners (Tomasello et al. 2007). Pointing gestures are recognized as one of the key components in infants' joint attention during interactions with others, in particular parents and caregivers (de Villiers Rader and Zukow-Goldring 2010), and their frequency is strongly reflected in early stages of subsequent language acquisition (Colonnesi et al. 2010). Liszkowski et al. (2012) have argued that this prelinguistic gestural communicative behaviour provides a language-independent universal foundation for diversified linguistic communication. Interestingly, this prelinguistic behaviour seems to be unique to the human race. A cross-species study (Povinelli et al. 1997) further signifies humanity's evolutionary uniqueness in using pointing gestures by asserting the fundamental differences between humans and chimpanzees in the use of pointing for understanding their immediate environment. Evidence from these studies has not only demonstrated the universality of pointing gestures in infants before cultural influences, but also implies that communicative pointing gestures may have been an evolutionary result of survival advantage for our ancestors over other similar species. In other words, pointing gestures are genetically programmed behaviours in our species and, most importantly, provide an example of biologically primary knowledge.

Tracing, which we define as dynamic movements of one or more fingers (typically the index finger) against a surface (e.g. paper-based instructional materials), is distinguished from static deictic pointing in space (cf. McNeill 1992). Circa 3500–3000 BCE, preceding the invention of writing, tracing movements have appeared to support human expression and story-telling (cf. Houston 2004); for example, analysis of finger flutings—lines drawn with fingers in soft surfaces—in Franco-Cantabrian Upper Paleolithic caves provides evidence that children as young as 2 years of age participated in artistic activities alongside adults as long as 40,000 years ago (Sharpe and Van Gelder 2006; Van Gelder 2015). Furthermore, finger fluting has been found to be the basis of artistic expression in Paleolithic groups which were wildly dispersed geographically, suggesting the cultural universality of tracing, like pointing (cf. Liszkowski et al. 2012). Given the antiquity of these activities, these archaeological studies suggest that tracing should be categorized as biologically primary knowledge.

Taken together, these lines of research suggest pointing and tracing gestures can be categorized as biologically primary, providing the basis for communication prior to language development.

Effects of Hand Position and Movement on Cognitive Processing

As noted above, links between visual processing and observer-generated movements were established in early research by Steinbach and Held (1968). However, the emerging

frameworks of embodied and grounded cognition (Barsalou 1999; Beilock and Holt 2007; Foglia and Wilson 2013; Glenberg 1997) have provided the basis for substantially increased understanding of these linkages over the past two decades. For example, targets near the hand are detected faster and scrutinized longer (Reed et al. 2006) and are more likely to be perceived as figures in figure-ground processing (Cosman and Vecera 2010). In other words, the targets near the hand are prioritized by the brain, and more visual working memory may be assigned towards processing targets nearby the hands. In addition, hand proximity improves change detection and visual working memory performance (Tseng and Bridgeman 2011), which may be a result of the hands shielding attentional processes from visual interference (Davoli and Brockmole 2012). The effects of hand position affecting cognitive resource management are further demonstrated in Dodd and Shumborski's (2009) study on memorizing a spatial array while pointing with the index finger. Learners who pointed throughout the session outperformed those who merely read the material. However, when learners only pointed at specific objects in the array as opposed to pointing to all of the objects, the performance was enhanced at the pointed objects while performance on non-pointed objects was inhibited. Taken together, targets near the hand appear to receive prioritized access to visual working memory, while the attentional focus created with the hands is able to effectively reduce the interference from non-targets.

These studies demonstrate the connection between hand movements and positions and altered cognitive processes, including working memory management—a fundamental concern of cognitive load theory.

Learning by Pointing and Tracing

Tracing with the index finger can be distinguished from simply pointing with the index finger by its greater spatial dynamics, as well as the greater sensory input resulting from tracing out a substantial curve's surface as opposed to pointing towards an object. While pointing is an extremely common and effective gesture that assists communication (for a meta-analysis, see Hostetter 2011), learning through tracing will necessarily involve manipulation of the object of a surface, usually with the index finger.

The incorporation of such movements into instruction to promote learning has a long history in education. *Sandpaper letters* (Montessori 1912, 1914), a pedagogy from Montessori education, involves children tracing letters of the alphabet made of sandpaper with their fingertips while they learn the phonetic sound associated with each letter. This simultaneous multi-sensory involvement of sound, vision, tactile and kinesthetic senses from the motion and sandpaper is argued to "fix the visual image and auditory sound quickly through the cooperation of senses" (Montessori 1912, p. 172).

Subsequent experimental research beginning in the 1970s investigated effects of tracing while learning. Hulme (1979) compared students' memory for abstract shapes when the traced the object versus pointed at it, concluding "…tracing produces a distinct motor memory trace and that this additional source of information aids visual recognition" (p. 249). Under laboratory settings, the efficacy of tracing has been verified for letters (Hulme et al. 1987) and geometric shapes (Kalenine et al. 2011). A number of experiments (Bara et al. 2007; Bara et al. 2004) have demonstrated that tracing enhances recognition of various stimuli. For instance, kindergarten children who explored letters via training incorporating haptic, visual and auditory elements performed better in pseudo-word reading, letter recognition and

phonological awareness than those who explored using only visual and auditory training (Bara et al. 2007). Another pedagogy incorporating finger pointing and tracing is 'print referencing' (Justice and Ezell 2004; Justice et al. 2009). Print referencing methods are designed to help students learn to read by requiring them to point to specific words on a page or to track under words as the text is read.

Tracing Effect and Cognitive Architecture

Research framed by CLT on the tracing effect has considered a range of possible mechanisms for tracing effectiveness. In this section, we consider two of the main mechanisms discussed in the literature.

Information Packaging Hypothesis

Self-generated gestures, including tracing, might affect cognition with a substantial visuo-spatial component in several ways, including supporting retrieval of both lexical and spatial representations from long-term memory, enhancing focus on spatial information and 'packaging' spatial information with speech (Alibali 2005). At the same time, drawing on the *information packaging hypothesis* (cf. Alibali et al. 2000), Ping and Goldin-Meadow (2010) argued that 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). According to CLT, chunking multiple elements of information into a single element acts to reduce intrinsic cognitive load of the learning task, enhancing schema construction and/or automation. In terms of CLT, schemas have typically played the role of chunking, but under this analysis, suitable gestures such as tracing could also reduce intrinsic cognitive load by visuo-spatially chunking elements of to-be-learned information. This account of the tracing effect would thus predict students who point and trace while learning should report lower intrinsic cognitive load, but not extraneous cognitive load, than those who only study the material.

Signalling Effect

When learning in a multi-sensory environment, instructional elements that direct students' attention to the most relevant information across multiple sensory inputs should aid learning. One of the possible effects of tracing could be its capacity to act as an attention-guiding cue. Attention cueing/signalling (de Koning et al. 2009) has been applied in instructions to reduce split attention effects and improve instructional effectiveness. Attention cueing refers to increasing the visual salience of task-relevant information in an instructional animation and can be achieved in several ways, including the spotlight cue (i.e. luminance contrast; e.g. see de Koning et al. 2007, 2010), arrows (e.g. Imhof et al. 2013) and spreading-colour cues (e.g. Boucheix and Lowe 2010). Empirical studies have suggested that attention cueing, with the intention of redirecting attention to aspects of the display that requires the learners' attention, would effectively reduce visual search and hence decrease extraneous load imposed by split attention (e.g. Britton et al. 1982; de Koning et al. 2007; Loman and Mayer 1983). Since self-generated pointing and tracing gestures can be viewed as a form of 'self-signalling', under this hypothesis, students who point and trace while learning should report lower extraneous cognitive load, but not intrinsic cognitive load, than those who only study the material.

Tracing Effect Studies

The effects of tracing have been investigated in a range of experiments informed by CLT. Macken and Ginns (2014) investigated whether learning about heart physiology presented in the form of expository text and diagrams could enhance university students' learning. Participants who were instructed to point and trace outperformed those who sat on their hands as a control for non-tracing on terminology and comprehension tests; no effects were found on single-item subjective ratings of intrinsic, extraneous and germane cognitive load. Although the expected effects on cognitive load ratings were not detected, possibly due to the prolonged instruction time prior to measurement in cognitive load, positive effects on test performance provided initial evidence of biologically primary knowledge facilitating learning of biologically secondary knowledge.

Hu et al. (2014) investigated the tracing effect through an alternative instructional design, worked examples, using geometry topics with primary school–level students (years 5 and 6). They found that students who traced out elements of paper-based geometry-worked examples solved more practice problems and made fewer errors on the subsequent test. Another study (Hu et al. 2015), also using paper-based geometry-worked examples with more difficult test items, found that students who traced had a lower error rate, solved more test problems and took less time to solve test problems compared to the non-tracing condition. As predicted by the information packaging hypothesis discussed above, test difficulty ratings made by the tracing condition were lower than the non-tracing condition and were interpreted as a reduced intrinsic load for students who traced. In a second experiment, Hu et al. (2015) tested a hypothesized performance gradient across control (no tracing), tracing in the air and tracing on the paper conditions, based on an expected ordinal correlation between the number of working memory modalities activated (visual only; visual + kinesthetic; visual + kinesthetic + tactile) and performance. Performance gradient hypotheses were supported for most variates under analysis, including acquisition phase problem solving and error rate, test phase basic question test performance and advanced question test performance, time to solution and ratings of test question difficulty. These latter results for difficulty ratings were interpreted as evidence for tracing lowering intrinsic cognitive load through information packaging. Together, these results support the argument that tracing gestures will enhance learning most when they incorporate multiple sensory modalities.

Agostinho et al. (2015) extended investigations of the tracing effect into a computer-based medium using worked examples on interpreting temperature line graphs presented on iPads. They found that year 3 and 4 participants who traced out the temperature graphs with their index finger achieved higher transfer test scores, but not similar test scores. Self-reported test difficulty as a measure of intrinsic cognitive load did not differ across conditions; however, the authors noted that mean scores on this variate suggested a floor effect.

Ginns et al. (2016) found that tracing enhanced adolescents' learning from worked examples on triangle geometry, but also investigated the tracing effect using mathematical worked examples for a less obviously 'spatial' mathematics topic, order of operation operators (i.e. the symbols for addition, subtraction, multiplication, division and brackets). Controlling for prior ability in arithmetic, students who traced out the order of operations solved more transfer problems on a posttest. However, hypotheses regarding lowered intrinsic cognitive load during the test phase were not supported in either experiment.

The present experiment was designed to test whether instructions to trace out elements of a science lesson reduce cognitive load and enhance learning when incorporated into a novel

lesson sequence—a lesson with isolated elements (cf. Pollock et al. 2002) initially presented on A4-sized paper, followed by an integrated presentation of the water cycle on an A1-sized poster. Research on the tracing effect has largely used lessons presented on A4-sized paper, with one study (Agostinho et al. 2015) using iPads. While A4-sized paper and tablet computers are straightforward media for educators to access and design for, larger formats for lessons such as A1 posters or interactive whiteboards may allow clearer displays of relatively complex integrated graphics and text displays that the same information displayed on a smaller medium, e.g. A4 paper. The present study thus aimed to test whether the tracing effect would generalize to a different sequence of media (isolated presentation of elements followed by integrated elements) than previously investigated.

Taken together, the above studies informed by CLT support the inclusion of explicit instructions to point and/or trace while learning. Novices who were instructed to trace typically outperformed students in non-tracing conditions on a range of variates including terminology recall and text comprehension (Macken and Ginns 2014), error rates (Hu et al. 2014, 2015) and performance on transfer test problem-solving (Ginns et al. 2016; Hu et al. 2015). These studies provided some evidence for effects on intrinsic cognitive load, but results are mixed across the reported studies, and it remains possible that tracing affects both intrinsic and extraneous cognitive load. Importantly, no tracing studies to date have measured students' self-reported cognitive load immediately following the intervention phase using multi-item scales, which can be expected to yield more sensitive tests of hypotheses.

The Present Study

The present study tests hypotheses related to intrinsic versus extraneous cognitive load using multi-item rather than single-item self-report measures, as the use of the former in prior research might account at least in part for the mixed findings noted above through insufficiently reliable measurement (e.g. the use of single-item ratings, e.g. in Macken and Ginns 2014). Our hypotheses were as follows. Participants in the tracing conditions will

1. Report lower intrinsic cognitive load than participants in the non-tracing condition (H_1).
2. Report lower extraneous cognitive load than participants in the non-tracing condition (H_2).
3. Outperform participants in the non-tracing condition on a subsequent test of recall of terms (H_3).
4. Outperform participants in the non-tracing condition on a subsequent transfer test (H_4).

The experiments presented in this study were approved by the Human Research Ethics Committee of the University of Sydney under Protocol No. 2015/351. Informed consent was obtained from all individual participants included in the study.

Participants Forty-six grade 5 and grade 6 students aged between 9.7 and 12.0 years old ($M = 10.73$, SD = 0.62) from a public coeducational school in Sydney, Australia, voluntarily participated in this experiment. Participants were randomly assigned to conditions. Prior to analysis, two participants were removed due to unwillingness to attempt the posttests and transfer tests, leaving 44 participants in tracing ($n = 22$) and non-tracing ($n = 22$) conditions. Participants may have had previous knowledge of the water cycle from their earlier school curriculum, but the contents in the instructional materials included concepts and definitions

suggested by stage 4 (year 7 and year 8) of the New South Wales curriculum (BOS NSW 2006) to ensure the novelty of the contents for the participants.

Materials and Procedure The lesson materials consisted of a short lesson on the water cycle. The lesson materials and tests of understanding were aligned with the Australian science curriculum for stage 4 (years 7 and 8) and developed based on the contents and learning goals found in various educational resources, such as textbooks, books and websites (e.g. Bunn 1996; Nowicki 2012; Pefani 2005). In addition, instructional materials developed by Gordon (2012) were also used as a major reference. The material was reviewed by junior high school and primary school science teachers, as well as trialled with several students, ranging from year 5 to year 7, to calibrate the length and complexity of the materials and tests to fit the intended population and experimental needs.

The lesson materials were presented as a single A1-sized coloured poster. Sky, land, sea, trees and underground were displayed as background on the poster. Eight textboxes explaining the water cycle processes (e.g. evaporation, transpiration) were presented on the poster, consisting of a total of 450 words, ranging from 39 to 77 words long in each textbox. Each textbox only described one process. Eight arrows were presented on the poster, clearly indicating the directions of water moving through the environment. The textboxes were placed right next to the arrows best representing the processes described to minimize a possible split attention effect (Ginns 2006). For example, the textbox describing transpiration was as follows: "The living things on Earth are represented in the diagram by the trees. For plants, water is absorbed by the roots and used in photosynthesis, but it is also released from their leaves through the process of transpiration", and placed on top of the trees, with the arrow starting from the trees, going up into the atmosphere to represent this process.

Drawing on theorizing by Pollock et al. (2002), a two-phase, isolated interacting element learning approach was taken in designing the lesson materials. Rather than studying a resource with all elements of the water cycle presented on a single poster, participants first either studied or traced while studying individual elements (e.g. evaporation) of the water cycle for 2 min each, with each element presented on a separate A4 page. It was predicted that this would allow participants to learn individual processes more effectively in isolation from each other, reducing element interactivity and hence the intrinsic cognitive load during this initial phase. Students studied eight separate pages (2 min per page) about the individual water cycle system processes; Fig. 1 provides an example. Students were then asked to study, or trace while studying, the 'whole system' on an A1 poster for an additional 4 min (see Fig. 2); thus, the overall period of time spent studying the water cycle was 20 min.

In a pilot test of the experimental materials, students were verbally prompted to trace out the 'corresponding arrow(s)' with their index finger after reading each textbox. There was substantial variability in students' responses to these prompts, with some students responding with many more tracing actions (up to ten) than others (one or two). In an attempt to generate a more consistent tracing response from students, three specific instructions were added to the lesson materials for the tracing condition to ensure that the participants would complete the required intervention. On the first page of instructions, two simple practice opportunities were introduced, and the participants were required to complete the practice, "trace using your index finger of your writing hand in the following examples 5 times", before the start the learning phase. On the second page of instruction, it was stated, "As the instructions in the brackets (TRACE) tell you, use your index finger to trace out the arrow in the diagram on the page at least 5 times". Finally, for each process, the most important or the most suitable phrase to trace

Fig. 1 Page 6 of lesson material for tracing group

the arrow was labelled with bold font '(TRACE)'. For example (see Fig. 1), on the sixth page of lesson material describing evaporation, the description says, "Most water returns to the atmosphere through evaporation from the soil, rivers and oceans (TRACE). Evaporation occurs when water absorbs solar energy and changes from a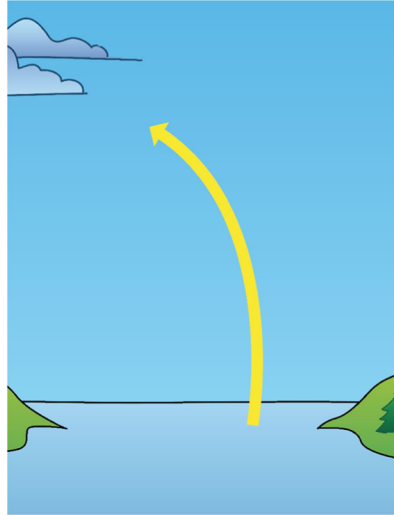 liquid form to a gaseous form, also known as water vapour". The expectation was that these additional instructions would increase the frequency of students' tracing behaviours, and hence the magnitude of the tracing effect.
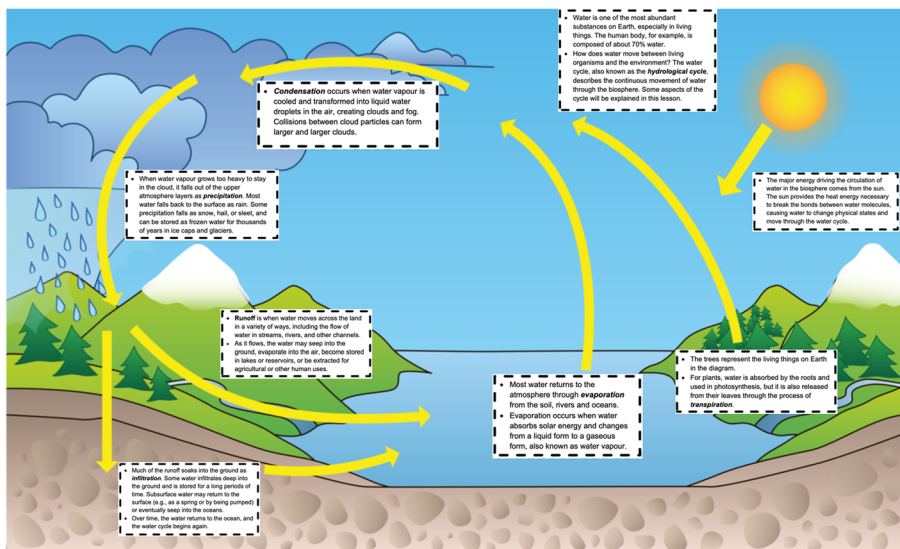

Fig. 2 The 'whole system' on an A1 poster

Pretest For the pretest, participants were given 10 min of working time and were instructed to check their work if they finished before time. The testing material used in this phase consisted of 20 multiple-choice questions with single correct answers on the main terms, concepts and implications of the water cycle. Participants were required to choose one choice from four alternatives. Test questions were modified from textbooks and the questions used in Gordon's (2012) study to match the contents of the materials. The concepts included the processes that move water molecules around the biosphere and the implications of the water cycle. Removing five items that more than 80% of students scored correctly (ceiling effects) or incorrectly (floor effects) gave a test with Cronbach's alpha = 0.51.

Learning Task Cognitive Load Self-Reports After the acquisition phase, participants completed the multi-item intrinsic and extraneous cognitive load questionnaire adapted from Leppink et al. (2013), with wording modified to fit the language capacities of the primary school-aged participants. Three items were used to assess students' experience of intrinsic cognitive load (e.g. "The topic covered in the activity was very difficult"), and three items were used to assess students' experience of extraneous cognitive load (e.g. "The instructions and/or explanations were full of unclear language"). Students responded on an 11-point Likert scale ranging from 0 (not at all the case) to 10 (completely the case). Cronbach's alpha for the intrinsic cognitive load scale was 0.71, and Cronbach's alpha for the extraneous cognitive load scale was 0.46.

Test Phase The paper-based posttest consisted of two sections: a recall test and a transfer test. Participants received a maximum of 10 min to complete each section.

In the recall test, the same diagram of the water cycle from the lesson material was provided with the arrows numbered from 1 to 8. Eight terminology questions required the participants to recall the names of water cycle processes labelled by ? numbered arrows with a single word or phrase. Each terminology question was followed by a subsequent explanatory question, "What happens in this process?", requiring them to recall the process in their own words. The terminology and explanatory questions were intended to measure the student's recall of the presented material, corresponding to factual knowledge in Bloom's taxonomy (Anderson et al. 2013). The first recall question regarding solar energy was used to provide students with a model for how to answer questions in this section of the test, and hence not counted towards the final score.

Since the purpose of the terminology questions was to test participants' recall of elements of the water cycle, the exact word or phrase was required, ignoring identifiable misspellings. For example, when the question asked for 'evaporation', the word 'evaporate' or 'evaporation' or indefinable misspellings such as 'evaprate' (missing 'o') was required in order to score one point (out of seven possible points), and 'water moving up' or 'water breaks into vapour' were not accepted. A different approach was used to score the subsequent explanatory questions because these questions were intended to test participants' recall of concepts in water cycle processes, not specific words. Hence, 20 idea units (2–4 units per question/process) were identified across seven processes by the researcher. The participants were assigned with 1 point (out of 20 points) for each idea unit that the participants wrote down, regardless of wording or spelling, such as 'change of physical state from liquid to gas' and 'moving into air'. Combining terminology and explanatory test questions, the maximum possible score across the recall test was 27.

In the transfer test, participants were required to answer two long answer questions. The first question asked the participants to explain how a solar still works. A short prompt with a labelled solar still diagram was given, "The solar still is a device uses solar energy to purify water. Different versions of a still are used to desalinate seawater, in desert survival kits and for home water purification", and the participants were asked "Please write down what you think will happen, and what will you get in the small cup in the middle of the solar still?" About half a page of blank lines were provided for the participants to answer the question. Five minutes was given to complete this question. The second transfer question asked about how air pollution would harm the environment and humans via the water cycle. The prompt asked, "Sulfur dioxide (SO_2) and nitrogen oxides (NO, NO_2, etc.) are gases commonly emitted by cars and industrial factories, and acids are formed when they are dissolved in water. Starting from how these harmful gases will enter the water cycle, describe how these gases may harm the environment and the humans. Draw on the diagram if necessary". About one third of the page was left blank as lines for answering, and a water cycle diagram with numbered arrows (the same as the comprehension test) was also provided. Scoring for both of the transfer tests were similar to the explanatory questions; 10 and 15 idea units were identified for each transfer question, respectively, while in transfer question 2, participants would also score a point if they drew the same idea on the diagram. For example, the idea unit "When it rains (precipitates), the pollution will go into… (1 point)" can also be shown as raining with dots falling down to the ground in order to score this point instead of writing it out. The maximum possible score across the two transfer questions was 25.

Data Analysis Since the posttests were open-ended short answer questions and are subject to evaluators' judgments, an inter-rater reliability analysis using the weighted kappa statistic (κ; Cohen 1968) was performed on individual questions to determine consistency among raters. Two raters were provided with identical rubrics and training and evaluated a random sample of 20% of test responses. The average inter-rater reliability for individual questions was 0.98. The raters came to complete agreement after discussion, and the rest of the posttests were scored on the basis of this agreement.

Initial checks on distributional properties of data under analysis used the Anderson-Darling normality test to evaluate distributional assumptions, using Keselman et al.'s (2013) recommendation of a type 1 error rate for this test of 0.20. Where data for a variate was non-normally distributed, the Mann-Whitney U using an exact p value was used to compare experimental conditions. Where normality assumptions were not violated, the independent group t test was used for comparison of conditions. Analyses of experimental data reported across the present article combined tests of significance, controlling the type 1 error rate at 0.05, with estimates of the standardized mean difference effect size (d) and its 95% confidence interval (Cumming 2012) using an Excel spreadsheet developed by Newcombe (2006a, 2006b) following Mann-Whitney tests. Based on a major review of over 800 meta-analyses of educational research, Hattie (2009) suggested the following benchmarks for effect size magnitude: small $d = 0.20$, medium $d = 0.40$, large $d = 0.60$ and above.

Results and Discussion

The variables under analysis were prior knowledge level before the acquisition phase, acquisition phase self-reported intrinsic and extraneous cognitive load level and test phase retention

and transfer scores. Means, standard deviations and Anderson-Darling test p values are provided in Table 1.

Prior Knowledge Level The mean score of the tracing condition ($M = 11.00$, SD = 2.11) was not significantly different to that of the non-tracing condition ($M = 10.00$, SD = 2.31), $U = 170.5$, $p = 0.091$, $d = 0.54$ [95% CI − 0.09, 1.11], indicating random assignment to conditions created by groups that were equivalent in terms of prior knowledge.

Cognitive Load Self-Reports The difference in self-reported intrinsic cognitive load rating made by the tracing condition ($M = 3.39$, SD = 1.91) and the non-tracing condition ($M = 3.79$, SD = 1.70) was not statistically significant, $t(42) = 0.72$, $p = 0.473$, $d = 0.22$ [95% CI − 0.41, 0.84]; thus, H_1 was not supported. A significant effect for self-reported extraneous cognitive load rating was found using Welch's t test (based on heterogenous variances across conditions): the rating reported by the tracing group ($M = 1.36$, SD = 1.12) was lower than the non-tracing group ($M = 2.62$, SD = 1.79), $t(35) = 2.79$, $p = 0.008$, $d = 0.84$ [95% CI 0.19, 1.49], supporting H_2.

Recall Questions A significant difference on retention scores was found favouring the tracing group ($M = 10.09$, SD = 1.83) over the non-tracing group ($M = 6.91$, SD = 1.99), $t(42) = 2.45$, $p = 0.019$, $d = 0.74$ [95% CI 0.13, 1.35], supporting H_3.

Transfer Questions Students in the tracing condition scored higher on transfer questions ($M = 5.23$, SD = 3.25) than students in the non-tracing condition ($M = 2.32$, SD = 2.19), $U = 105$, $p = 0.001$, $d = 1.11$ [95% CI 0.42, 1.70], supporting H_4.

General Discussion

The results of the present study thus supported three of the four hypotheses: tracing out elements of the water cycle poster resulted in lower extraneous (H_2) (but not intrinsic—H_1) cognitive load, as well as enhanced recall test performance (H_3) as well as transfer test performance (H_4).

Table 1 Means (M), standard deviations (SD) and Anderson-Darling (A-D) test p value for pretest, postlesson intrinsic and extraneous cognitive load self-report ratings and posttest recall and transfer test scores

Variable	Non-tracing condition			Tracing condition		
	M	SD	A-D	M	SD	A-D
Prelesson						
Water cycle pretest scores (/15)	10.00	2.31	0.178	11.00	2.11	0.506
Postlesson						
Intrinsic cognitive load (/10)	3.79	1.70	0.506	3.39	1.91	0.430
Extraneous cognitive load (/10)	2.62	1.79	0.324	1.36	1.12	0.220
Test phase						
Recall test (/27)	6.91	1.99	0.498	10.09	1.83	0.778
Transfer test (/25)	2.32	2.19	0.002	5.23	3.25	0.241

These results add to the existing literature on the tracing effect in several ways. Focusing first on theorizing, previous studies have suggested that tracing might affect overall cognitive load by reducing intrinsic cognitive load through information packaging (Alibali et al. 2000; Hu et al. 2015). Alternatively, tracing may act to cue attention (cf. de Koning et al. 2009), reducing extraneous cognitive load associated with search and match processes. Evidence for the former hypothesis, based on difficulty ratings of test items, has been mixed (e.g. Agostinho et al. 2015; Ginns et al. 2016; Hu et al. 2015), and evidence for the latter hypothesis is absent. The experiment discussed above tested both hypotheses using a multi-item, multi-scale cognitive load instrument positioned after the lesson. The above experiment provided evidence for a tracing effect on reports of extraneous cognitive load (H_2), but not reports of intrinsic cognitive load (H_1). These results should, however, be considered with caution for several reasons. First, the reliability of the extraneous cognitive load self-report scale was lower than is typically desired, likely a function of substantial skewness in student responses combined with consisting of only three items. Despite these shortcomings, the difference between conditions was substantial enough to emerge even with a somewhat unreliable measure. Second, these results might be idiosyncratic to the particular topic and lesson format investigated here. Other tracing studies are ongoing using Leppink et al.'s (2013) multi-item scale, which will provide a better indication of the generalizability of this result.

Taken together, the results of the present study act to refine the design principles of tracing instructions. Specifically, educators wishing to incorporate tracing instructions into lessons should be mindful of the element interactivity of lesson materials given the prior knowledge levels of students. Students with low prior knowledge who are asked to learn about complex processes presented through text and diagrams will learn more effectively if they trace out separate elements of the lesson before encountering all elements of the lesson presented simultaneously.

Although the present study incorporated a novel larger lesson format, the A1-sized format poster, our study's design does not investigate the relative contribution to learning of tracing on A4-isolated elements (first phase), followed by tracing out the whole water cycle system (second phase). We hypothesize that larger formats will support larger tracing gestures, which in turn should generate larger tracing effects through greater working memory capacity across multiple working memory channels (Hu et al. 2015; Hulme 1979). 'Dosage' questions have begun to be investigated by Hu et al. (2015; experiment 2) and present an opportunity for future research. Beyond printed posters, other large format presentation media such as interactive whiteboards and touchscreen tables present tracing affordances; the multi-touch potential of these media means that future research on tracing may be enhanced by consideration of cognitive load from cooperative learning perspectives (for a review, see Kirschner et al. 2018).

A number of measurement limitations of the study also suggest directions for further research. The reliability of some measures, including the pretest and the self-reported measures of cognitive load, were lower than is typically desired. The pretest was not used as covariates in subsequent analysis due to lower than desired reliability; however, the interpretation of the differential results for intrinsic versus extraneous cognitive load might be questioned due to the low reliability of the extraneous cognitive load measure. As noted above, both cognitive load self-reports were brief (three items), and responses to some items, particularly extraneous load items, were substantially skewed by a large proportion of zero ('not at all the case') responses. Sheng and Sheng's (2012) Monte Carlo study found that Cronbach's alpha underestimates reliability when data is highly skewed, and this underestimation increases when skewed data

are drawn from small sample sizes. Future research should continue to refine these measures; for example, Martin and Evans (2018) adapted Leppink et al.'s (2013) scale to the mathematics classroom experience by adding two additional items to the original four, yielding substantially more reliable measures than those in the present study. Another limitation of self-reports used in the present study was the collection of data only after both learning phases (A4 pages then A1 page) had been completed. Future studies should consider collecting these measures at multiple points across both learning and test phases (cf. Hu et al. 2015) in order to understand the dynamics of cognitive load across a lesson. The results of the present study can also inform power analyses of future studies, preferably with larger samples in order to support more contemporary measures of reliability such as McDonald's omega (Crutzen and Peters 2017; McDonald 1999) as well as moderation and/or mediation hypotheses (Hayes 2017) (e.g. the expertise reversal effect; Kalyuga 2007). Beyond self-report measures, objective measures of cognitive processing such as eye-tracking (e.g. Johnson and Mayer 2012; Pouw et al. 2016) may yield real-time rather than postlesson insights into the effects of tracing, particularly regarding the hypothesized increased visual processing around the hand.

In conclusion, the evolutionary underpinnings of CLT (Paas and Sweller 2012), in particular the distinction between biologically primary and biologically secondary knowledge, have provided the basis for investigations of the tracing effect. In the present study, we sought to expand on previous research by testing whether tracing might support learning a novel science topic, the water cycle, through a novel format, a lesson combining A4 paper–based elements followed by integrated elements on an A1-sized poster. The results of this study suggest a clearer role of tracing in reducing extraneous rather than intrinsic cognitive load.

Compliance with Ethical Standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

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. https://doi.org/10.1007/s10648-015-9315-5.

Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*(4), 307–331. https://doi.org/10.1207/s15427633 scc0504_2.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes, 15*(6), 593–613. https://doi.org/10.1080 /016909600750040571.

Anderson, L. W., Krathwohl, D. R., et al. (2013). *Taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives.* Pearson.

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*(3), 433–449. https://doi.org/10.1016/j.cogdev.2004.05.003.

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*(4), 643–663. https://doi.org/10.1348/026151007 X186643.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(04), 577–660. https://doi. org/10.1017/S0140525X99002149.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development, 21*, 205–226.

Beilock, S. L., & Holt, L. E. (2007). Embodied preference judgments: can likeability be driven by the motor system? *Psychological Science, 18*(1), 51–57. https://doi.org/10.1111/j.1467-9280.2007.01848.x.

Board of Studies New South Wales (BOS NSW). (2006). *Science K-10 syllabus*. Sydney: Board of Studies NSW.

Boucheix, J. M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*(2), 123–135. https://doi. org/10.1016/j.learninstruc.2009.02.015.

Britton, B. K., Glynn, S. M., Meyer, B. J., & Penland, M. J. (1982). Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology, 74*(1), 51–61.

Bunn, T. (1996). *Form 6 biology revision*. Auckland: ESA Publications.

Camaioni, L. (1993). The development of intentional communication: a re-analysis. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicative development* (pp. 82–96). New York: Routledge.

Cohen, J. (1968). Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220.

Colonnesi, C., Stams, G. J. J., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: a meta-analysis. *Developmental Review, 30*(4), 352–366. https://doi.org/10.1016/j. dr.2010.10.001.

Cosman, J. D., & Vecera, S. P. (2010). Attention affects visual perceptual processing near the hand. *Psychological Science, 21*(9), 1254–1258. https://doi.org/10.1177/0956797610380697.

Crutzen, R., & Peters, G. J. Y. (2017). Scale quality: alpha is an inadequate estimate and factor-analytic evidence is needed first of all. *Health Psychology Review, 11*(3), 242–247.

Cumming, G. (2012). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

Davoli, C. C., & Brockmole, J. R. (2012). The hands shield attention from visual interference. *Attention, Perception, & Psychophysics, 74*(7), 1386–1390. https://doi.org/10.3758/s13414-012-0351-7.

de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology, 21*(6), 731–746. https://doi.org/10.1002 /acp.1346.

de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: guidelines for research and design. *Educational Psychology Review, 21*(2), 113–140. https://doi.org/10.1007/s10648-009-9098-7.

de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Attention guidance in learning from a complex animation: seeing is understanding? *Learning and Instruction, 20*(2), 111–122. https://doi. org/10.1016/j.learninstruc.2009.02.010.

de Villiers Rader, N., & Zukow-Goldring, P. (2010). How the hands control attention during early word learning. *Gesture, 10*(2-3), 202–221. https://doi.org/10.1075/gest.10.2-3.05rad.

Dodd, M. D., & Shumborski, S. (2009). Examining the influence of action on spatial working memory: the importance of selection. *Quarterly Journal of Experimental Psychology, 62*(6), 1236–1247. https://doi. org/10.1080/17470210802439869.

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(3), 319–325. https://doi.org/10.1002/wcs.1226.

Geary, D. C. (2008). An evolutionarily informed education science. *Educational Psychologist, 43*(4), 179–195. https://doi.org/10.1080/00461520802392133.

Ginns, P. (2006) Integrating information: a meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction 16*(6):511–525. https://doi.org/10.1016/j.learninstruc.2006.10.001.

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. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1–55.

Gordon, C. (2012). *Self-management of students' cognitive load within a primary school context* (unpublished Honours thesis). University of Wollongong, New South Wales, Australia.

Hattie, J. A. (2009). *Visible learning: a synthesis of 800+ meta-analyses on achievement. Abingdon*. Routledge.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach* (2nd ed.). New York, NY: Guilford Press.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297–315. https://doi.org/10.1037/a0022128.

Houston, S. D. (2004). *The first writing: script invention as history and process*. Cambridge: Cambridge University Press.

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. https://doi.org/10.1016/j.learninstruc.2014.10.002.

Hulme, C. (1979). The interaction of visual and motor memory for graphic forms following tracing. *The Quarterly Journal of Experimental Psychology, 31*(2), 249–261. https://doi.org/10.1080/14640747908400724.

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*(4), 299–307.

Imhof, B., Scheiter, K., Edelmann, J., & Gerjets, P. (2013). Learning about locomotion patterns: effective use of multiple pictures and motion-indicating arrows. *Computers & Education, 65*, 45–55. https://doi.org/10.1016/j.compedu.2013.01.017.

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology. Applied, 18*(2), 178–191. https://doi.org/10.1037/a0026923.

Justice, L. M., & Ezell, H. K. (2004). Print referencing: an emergent literacy enhancement strategy and its clinical applications. *Language, Speech, and Hearing Services in Schools, 35*(2), 185–193. https://doi.org/10.1044/0161-1461.

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 explicit print referencing. *Language, Speech, and Hearing Services in Schools, 40*(1), 67–85. https://doi.org/10.1044/0161-1461 (2008/07-0098).

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*(1), 18–26. https://doi.org/10.1177/0165025410367443.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3.

Kalyuga, S. (2011). Cognitive load theory: how many types of load does it really need? *Educational Psychology Review, 23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7.

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

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning, 13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y.

Leppink, J., Paas, F., van der Vleuten, C. P., van Gog, T., & van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1.

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & De Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cognitive Science, 36*(4), 698–713. https://doi.org/10.1111/j.1551-6709.2011.01228.x.

Loman, N. L., & Mayer, R. E. (1983). Signaling techniques that increase the understandability of expository prose. *Journal of Educational Psychology, 75*(3), 402–412. https://doi.org/10.1037/0022-0663.75.3.402.

Macken, L., & Ginns, P. (2014). Pointing and tracing gestures may enhance anatomy and physiology learning. *Medical Teacher, 36*(7), 596–601. https://doi.org/10.3109/0142159X.2014.899684.

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

Martin, A. J., & Evans, P. (2018). Load reduction instruction: exploring a framework that assesses explicit instruction through to independent learning. *Teaching and Teacher Education, 73*, 203–214. https://doi.org/10.1016/j.tate.2018.03.018.

McDonald, R. P. (1999). *Test theory: a unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.

McNeill, D. (1992). *Hand and mind: what gesture reveals about thought*. Chicago: University of Chicago Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158.

Montessori, M. (1912). *The Montessori method*. London: William Heinemann.

Montessori, M. (1914). *Dr. Montessori's own handbook*. London: William Heinemann.

Newcombe, R. G. (2006a). Confidence intervals for an effect size measure based on the Mann–Whitney statistic. Part 1: general issues and tail-area-based methods. *Statistics in Medicine, 25*(4), 543–557. https://doi.org/10.1002/sim.2323.

Newcombe, R. G. (2006b). Confidence intervals for an effect size measure based on the Mann–Whitney statistic. Part 2: asymptotic methods and evaluation. *Statistics in Medicine, 25*(4), 559–573. https://doi.org/10.1002/sim.2324.

Nowicki, S. (2012). Holt McDougal biology: Interactive reader. **Houghton Mifflin Harcourt**.

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*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2.

Pefani, N. (2005). *Excel Science Study Guide Years 7–8*. NSW: Pascal Press.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*(3), 193–198. https://doi.org/10.1037/h0049234.

Ping, R., & Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science, 34*(4), 602–619. https://doi.org/10.1111/j.1551-6709.2010.01102.x.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*(1), 61–86. https://doi.org/10.1016/S0959-4752(01)00016-0.

Pouw, W. T., Mavilidi, M. F., van Gog, T., & Paas, F. (2016). Gesturing during mental problem solving reduces eye movements, especially for individuals with lower visual working memory capacity. *Cognitive Processing, 17*(3), 269–277. https://doi.org/10.1007/s10339-016-0757-6.

Povinelli, D. J., Reaux, J. E., Bierschwale, D. T., Allain, A. D., & Simon, B. B. (1997). Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees. *Cognitive Development, 12*(4), 423–461. https://doi.org/10.1016/S0885-2014(97)90017-4.

Reed, C. L., Grubb, J. D., & Steele, C. (2006). Hands up: attentional prioritization of space near the hand. *Journal of Experimental Psychology: Human Perception and Performance, 32*(1), 166–177. https://doi.org/10.1037/0096-1523.32.1.166.

Sharpe, K., & Van Gelder, L. (2006). Evidence for cave marking by Paleolithic children. *Antiquity, 80*(310), 937–947.

Sheng, Y., & Sheng, Z. (2012). Is coefficient alpha robust to non-normal data? *Frontiers in Psychology, 3*, 1–13. https://doi.org/10.3389/fpsyg.2012.00034.

Steinbach, M. J., & Held, R. (1968). Eye tracking of observer-generated target movements. *Science, 161*(3837), 187–188.

Sweller, J. (2008). Instructional implications of David C. Geary's evolutionary educational psychology. *Educational Psychologist, 43*(4), 214–216. https://doi.org/10.1080/00461520802392208.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5.

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation (Vol. 55, pp. 37–76)*. Academic Press.

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

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

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*(3), 705–722. https://doi.org/10.1111/j.1467-8624.2007.01025.x.

Tseng, P., & Bridgeman, B. (2011). Improved change detection with nearby hands. *Experimental Brain Research, 209*(1), 257–269. https://doi.org/10.1016/j.scient.2012.11.006.

Van Gelder, L. (2015). Counting the children: the role of children in the production of finger flutings in four Upper Palaeolithic caves. *Oxford Journal of Archaeology, 34*(2), 119–138. https://doi.org/10.1111/ojoa.12052.

Vygotsky, L. S. (2004). Mastering attention. In R. W. Rieber & D. K. Robinson (Eds.), *The essential Vygotsky* (pp. 375–400). Boston, MA: Springer.