

# Pointing and tracing gestures may enhance anatomy and physiology learning

LUCY MACKEN AND PAUL GINNS

The University of Sydney, Australia

#### **Abstract**

**Background**: Currently, instructional effects generated by Cognitive load theory (CLT) are limited to visual and auditory cognitive processing. In contrast, "embodied cognition" perspectives suggest a range of gestures, including pointing, may act to support communication and learning, but there is relatively little research showing benefits of such "embodied learning" in the health sciences.

Aim: This study investigated whether explicit instructions to gesture enhance learning through its cognitive effects.

**Methods**: Forty-two university-educated adults were randomly assigned to conditions in which they were instructed to gesture, or not gesture, as they learnt from novel, paper-based materials about the structure and function of the human heart. Subjective ratings were used to measure levels of intrinsic, extraneous and germane cognitive load.

**Results**: Participants who were instructed to gesture performed better on a knowledge test of terminology and a test of comprehension; however, instructions to gesture had no effect on subjective ratings of cognitive load.

**Conclusions**: This very simple instructional re-design has the potential to markedly enhance student learning of typical topics and materials in the health sciences and medicine.

#### Introduction

Cognitive load theory (CLT) is a contemporary educational psychology theory applying cognitive science principles to instructional design (Sweller et al. 2011). CLT foregrounds the importance of working memory as a "bottleneck" for learning because of its limitations both in terms of capacity and duration, especially in processing novel information. CLT differentiates three types of cognitive load that may draw on working memory resources during learning (Sweller 2010). Intrinsic load is driven by the inherent complexity of the material experienced by a learner, and is a function of the degree of interactivity between to-be-learned elements of information relative to the learner's prior knowledge. Extraneous load is a function of the unnecessary processing of interacting elements not relevant to learning, because cognitive limitations have not been taken into account in design. Germane load is mental effort directed towards intrinsic interacting elements leading to schema construction and automation through appropriately challenging material designed to encourage accommodation of new information. By extension, the aim of this instructional design should be to balance competing demands by reducing extraneous load and increasing germane load, wherever possible, if learning is to be maximised.

Through a strong focus on human cognitive architecture, instructional design theories such as CLT have the potential to markedly enhance student learning in medical and health sciences (for a review, see Van Merriënboer & Sweller 2010).

### **Practice points**

- Cognitive load theory (CLT) emphasises the strengths and limitations of human cognitive architecture when designing instructions.
- Recent embodied cognition perspectives suggest new directions for CLT based on gesturing for learning.
- This randomised experiment demonstrated explicit instructions to point to and trace out elements of complex anatomical instructions, improved performance on terminology and comprehension tests.

While the "cognitive revolution" in psychology has led to the development of many theories emphasising the role of cognition in learning, the focus has remained, perhaps too narrowly, on "in-head" processes. Over the past decade, findings from basic research examining the neural basis of multi-modal perception and attention (Abrams et al. 2008; Cosman & Vecera 2010; Spence 2010) suggest contemporary educational theories concerned with attentional resources during learning may benefit from a greater consideration of "embodied cognition" perspectives. For example, Abrams et al. (2008) examined participants' performance across three classic visual attention tasks (visual search, inhibition of return and attentional blink), varying whether participants held their hands close to the stimulus display or not. They found participant's vision was enhanced for objects near the hands,

whether the hands were visible or not, and suggested such a mechanism "could facilitate the detailed evaluation of objects for potential manipulation, or the assessment of potentially dangerous objects for a defensive response" (p. 1035).

One of the earlier instructional re-designs generated by CLT was the split-attention effect (Tarmizi & Sweller 1988; Ginns 2006). Split-attention occurs whenever a learner is required to mentally integrate two or more disparate but related sources of information (e.g. diagram and text) in order to permit understanding. This splitting of attention acts as a source of extraneous cognitive load by requiring allocation of working memory resources to search for and match the separate sources of instruction; resources allocated to search and match are thus unavailable for learning. Designing materials differently may ameliorate the split-attention effect; for example, reducing the need for search and match by placing text on the related diagram rather than separately on the page, or through the use of colour coding or arrows and numbers to link visual and textual elements (for a review of signaling methods, see de Koning et al. 2009). Pointing may also act to reduce split-attention by enhancing the capacity to search and match-related referents in instructional text (Alibali & Nathan 2012), and lead to enhanced germane cognitive load through the recruitment of cognitive resources for learning. Pointing gestures have been found to play a role in attentional guidance, communication and vocabulary development from as young as seven months (Colonnesi et al. 2010). Justice and Ezell (2004) advocate "print referencing" methods based on pointing when helping students learn to read, including pointing to specific words on a page, or tracking under words as the text is read. The benefits of pointing have recently been found to extend well into highschool learning; Moreno et al. (2010) found students whose computer-based lesson on electrical circuits included an animated agent that pointed to the various elements of ongoing instruction outperformed those whose materials included an animated "arrow" cursor, or no signalling method.

Fingers - primarily the index finger - can be used by learners not only to point, but to trace out visuo-spatial materials such as diagrammatic elements. Montessori (1912) developed a system of sandpaper letters and numbers which were traced out while a teacher verbalised the letter or number; the efficacy of this teaching approach has subsequently been verified for letters (Hulme et al. 1987) and geometric shapes (Kalenine et al. 2011). Van Gog et al. (2009) speculate that in addition to auditory and visual working memory processing channels, the haptic system may have its own working memory processing channel dedicated specifically to haptic (including tactile) processing. Incorporating haptic elements into medical teaching activities and materials through gesturing could further expand working memory capacity available for learning, and build straightforwardly on the basic impulse of medical teachers and students to use their bodies while communicating and learning.

In CLT terms, then, the use of gestures – including pointing and tracing – to assist search and match is of particular interest, as there may be fundamental limitations to the extent to which visually presented text and associated referents within diagrams can always be closely integrated. Thus, pointing at

related referents may act to further enhance learning from already highly integrated instructional materials. The current study provides an initial test of these conjectures using the CLT framework to investigate gesturing for learning, and proposes the following hypotheses.

### Research hypotheses

- Participants who are instructed to point and trace while learning novel materials will outperform participants who are instructed not to gesture, on both terminology and comprehension tests.
- (2) Participants in the two conditions will not differ in the degree of intrinsic cognitive load reported, as this aspect of the experimental design is held constant.
- (3) Participants who are instructed to point and trace while learning the novel materials will rate the materials as less difficult (i.e. lower in extraneous cognitive load) than participants who are instructed not to gesture, as pointing and tracing will assist in searching for and matching related parts of the diagram and text.
- (4) Participants who are instructed to point and trace while learning the novel materials will report concentrating more during learning than participants who are instructed not to gesture; i.e., instructions to gesture will generate higher levels of germane cognitive load, as gesturing will facilitate more detailed evaluation of the materials (cf. Abrams et al. 2008).

#### Method

#### Participants

Forty-two university-educated adults (31 females, 11 males) with a mean age of 26 years 11 months (SD=5.51 years) participated in this study. Half of the participants served in the gesture group (15 females, six males), and half served in the non-gesture group (16 females, five males).

#### Materials

Paper-based experimental materials were presented in a display folder and consisted of: (1) a physiology prior knowledge multiple choice test, (2) a single page of initial learning instructions detailing how the 12 page document should be learnt, (3) 12 pages of instructional materials detailing the structure and functions of the human heart, (4) a single page of subjective ratings of intrinsic, extraneous and germane cognitive load, consisting of three questions on a 7-point Likert scale, (5) a terminology multiple choice test and (6) a comprehension multiple choice test. Participants were given a separate answer sheet for each test.

A single page detailing how the material about the heart should be learnt preceded the instructional materials. This page of learning instructions was structured similarly for both conditions and the following identical statement appeared at the top of the page in both cases, "In the next 25 min I am asking you to read through, understand, and learn about some aspects of human anatomy". Instructions to participants in the gesture condition were as follows: "Please use your hands

where you need to make a link between text and an associated part of the diagram. Some ways you might like to do this are:

- Point at the word in the text, then point at the corresponding location on the diagram
- Leave your finger on the diagram as you read about the corresponding element in the text
- Use more than one finger to simultaneously point to parts of text and the diagram that are related
- Where you see arrows indicating blood flow of the heart, use your hands to trace along the arrows."

In contrast, participants in the non-gesture condition received the following instructions: "Please do not use your hands while you learn this material. To assist you keeping in your hands still, please:

- Sit on your hands
- Only use your hand to turn the page."

Participants in both conditions were then given the opportunity to practice either gesturing or not gesturing while learning, by studying a diagram of the human eye accompanied by 57 words of expository text.

The instructional material consisted of a 12 page document, approximately 2000 words long, giving an introduction to the structure and function of the human heart. A black and white diagram accompanied the text on each page, detailing the particular structures and blood flows through the heart. This material was originally developed as part of the Programme of Systemic Evaluation (Dwyer 1972). Across the 12 pages, the potential for split-attention was minimised wherever possible with standard methods used in split-attention effect experiments. These included positioning explanatory text related to a specific aspect of heart structure or function as close as possible to its diagrammatic referent; numbers, e.g. to mark where particular structures or processes mentioned in the text appeared in diagrams; and the use of arrows, to direct attention to associated referents.

Subjective ratings were used to measure aspects of cognitive load, in the form of a 7-point Likert scale, based on the scales developed by Cierniak et al. (2009). This scale was used to measure each of the three types of cognitive load: intrinsic, extraneous and germane cognitive load. The scales had three labels: 1 = not at all, 4 = somewhat, 7 = extremely. The scale measuring intrinsic cognitive load asked "How complicated was the learning content for you?" This scale was designed to measure intrinsic cognitive load through a focus on the content of the instructional material. The extraneous cognitive load scale asked "How difficult was it for you to learn with the material?" This scale was designed as a measure of extraneous cognitive load through a focus on the material itself, rather than the content of the material. The scale measuring germane cognitive load asked, "How much were you able to concentrate on learning?" This scale was adapted from Salomon (1984), and is assumed to be a measure of germane cognitive load through the question's focus on the participant's ability to concentrate, reflecting the direction of attention and investment of effort towards schema construction.

A Physiology Prior Knowledge test was used to measure students' general knowledge of physiology and anatomy. This test, developed by Dwyer (1972), has been utilised as a covariate in previous experiments using these Programme of Systemic Evaluation materials (e.g. Joseph & Dwyer 1984). The Physiology Prior Knowledge test consisted of 30 multiple-choice questions.

The Terminology test and Comprehension test were used to measure participants' learning from the instructional materials. Dwyer (1972) developed these instruments as criterionreferenced tests appropriate for testing learning from the 12 page document about the structure and functions of the human heart. The Terminology test consisted of 20 multiplechoice items. These items took the form of sentences with one word missing (e.g. "The \_\_\_\_ is the common opening between the right auricle and the right ventricle"); designed to measure participants' knowledge of specific definitions and labels. Participants were required to choose one word to complete the sentence from five alternatives. The Comprehension test consisted of 20 multiple-choice items. These items took the form of sentences that gave one piece of information about the heart and, based on that information, participants were required to make judgements about some other aspect of the heart (e.g. "When the blood in the aorta is exerting a superior pressure on the aortic valve, what is the position of the mitral valve?"), by choosing one of the four options. The comprehension questions were designed to measure participants' understanding of the structures and functions of the heart, including processes that occur simultaneously; to answer such questions correctly, the learner must have both understood the instructions, and be able to use the information to explain other phenomena (Dwyer 1972).

#### Procedure

Participants were tested individually, with each participant randomly assigned to either the gesture or non-gesture condition. Participants first received an information sheet and consent form to read and complete. Participants then completed a Physiology Prior Knowledge pre-test, which lasted for 20 min. The researcher then presented participants with the page of learning instructions and asked them to attempt to understand and remember the heart materials by following the instructions on that page. Participants were advised that they had 25 min to learn the materials about the heart and that they should continue to go through the materials until the full 25 min had elapsed. Participants were able to move through each page of instructions at their own pace. After the 25-min learning session, participants rated the degree of intrinsic, extraneous and germane cognitive load experienced. Participants then completed the terminology and comprehension tests, with a maximum of 10 min to complete each test. Finally, participants were debriefed and thanked for their participation in the study.

#### Results

The reliability of the Prior Knowledge, Terminology and Comprehension tests was assessed using the Kuder–Richardson (KR-20) measure of internal consistency for dichotomously scored items. For the Physiology Prior Knowledge test, KR-20 = 0.64; for the Terminology test, KR-20 = 0.81; and for the Comprehension test, KR-20 = 0.71.

The reliability of each of the tests was judged suitable for further analyses, although it should be noted that the Prior Knowledge test had an estimate of reliability lower than the 0.70 benchmark often used in social science research.

Group differences on the variates described above were analysed using independent t-tests, accompanied with an estimate of the effect size, the standardised mean difference, Cohen's d, and its 95% confidence interval. Although Cohen's (1988) benchmarks for effect sizes are widely used (small d=0.20, medium d=0.50, and large d=0.80 and above), Cohen himself cautioned against blanket adoption of these recommendations if benchmarks within a substantive field are available. The present study adopted benchmarks recommended by Hattie (2009) (small d=0.20, medium d=0.40, and large d=0.60). Based on his review of over 800 meta-analyses in the field of education, Hattie (2009) argues that these benchmarks provide a more accurate continuum for understanding the meaningfulness of experimental results in education.

## Were the groups equivalent in prior knowledge?

The difference in performance on the Physiology Prior Knowledge test between the gesture condition ( $M\!=\!18.38$ ,  $SD\!=\!3.26$ ) and the non-gesture condition ( $M\!=\!17.76$ ,  $SD\!=\!4.36$ ) was not statistically significant,  $t(40)\!=\!0.52$ ,  $p\!=\!0.605$ ,  $d\!=\!0.16$ , 95% CI [-0.46, 0.78], indicating the random assignment of participants to conditions was successful in distributing levels of prior knowledge equally across conditions. However, scores on the Physiology Prior Knowledge test did not correlate sufficiently highly with scores on the Terminology test ( $r\!=\!0.13$ ,  $p\!=\!0.403$ ) or the Comprehension test ( $r\!=\!0.26$ ,  $p\!=\!0.097$ ) to justify its use as a covariate in subsequent analyses.

# Does the instruction to gesture enhance learning?

We hypothesised that participants who were instructed to gesture while learning the novel materials would outperform participants who were instructed not to gesture, on both the Terminology and Comprehension tests. As predicted, participants who were instructed to gesture performed significantly better on the Terminology test (M=15.90, SD=2.41) than participants who were instructed not to gesture (M=12.62, SD=4.65), t(40) = 2.88, p=0.006, d=0.89, 95% CI [0.25, 1.52]. The difference in the Comprehension test between participants who were instructed to gesture (M=13.95, SD=2.60) and those instructed not to gesture (M=11.86, SD=3.97), was also statistically reliable, t(40) = 2.03, p=0.050, d=0.62, 95% CI [0.01, 1.25].

# Does the instruction to gesture affect self-reported cognitive load?

As hypothesised, for intrinsic cognitive load, the difference between the gesture condition (M=3.90, SD=1.41) and the non-gesture condition (M=4.52, SD=1.75) was not statistically significant, t(40)=-1.26, p=0.214, d=-0.39, 95% CI

[-1.00, 0.22]. However, contrary to our hypotheses, participants who were instructed to gesture did not report lower extraneous cognitive load (M=3.62, SD=1.40) than participants who were instructed not to gesture (M=4.00, SD=1.82), t(40)=-0.76, p=0.450, d=-0.23, 95% CI [-0.84, 0.37]. Participants who were instructed to gesture also did not report higher germane cognitive load (M=5.05, SD=0.86) than participants who were instructed not to gesture (M=4.67, SD=1.59), t(40)=0.96 (unequal variances assumed), p=0.343, d=0.30, 95% CI [-0.33, 0.92].

#### Discussion

This study was a novel investigation of the impact of instructions to gesture on learning from an extended expository text with diagrams, of the kind regularly used by students in the health sciences and medicine. We drew on CLT and embodied cognition perspectives to hypothesise that particular gestures – using the fingers to point and trace while studying – would enhance learning, because such actions would assist in the "search and match" process underlying the split-attention effect (Ginns, 2006), as well as enhance attentional processes for visual targets (i.e. diagrams and associated text) (Abrams et al. 2008). Results across two tests of learning were consistent with this hypothesis. Using Hattie's (2009) suggested benchmarks for experimental effects in education, we found a large effect (d=0.89) of gesturing on learning terminology (i.e. specific heart structures), and a large effect (d=0.62) of gesturing on comprehending the instructional materials. Importantly, these results were generated using instructional materials that had already been modified to represent "best practice", from a cognitive load perspective, in instructional design of paper-based materials. Wherever possible, standard techniques (e.g. integration, numbering, arrows) for reducing extraneous cognitive load from split-attention had been employed. Thus, the gesture effect investigated in the present study may be a means for enhancing learning over and above the already substantial improvements in learning possible through CLT-generated instructional redesigns.

While the above results for tests of learning are consistent with a CLT interpretation, our hypotheses regarding self-reports of extraneous and germane cognitive load were not supported. Cognitive load self-reports used in our study were based on those used by Cierniak et al. (2009). In their study, participants rated their intrinsic, extraneous and germane cognitive load after a 16.5-min learning phase on the structure and function of the kidney, presented across four computer screens. Our study followed this procedure, with participants rating the three aspects of cognitive load after a 25-min learning phase, presented across 12 pages. Given the extended instruction time and number of pages studied, it is possible that the measures used in our study were not sensitive enough to detect group differences in self-reports of cognitive load. More sensitive measures might be gained in future studies by using a multi-item cognitive load scale (Leppink et al. 2013). Alternatively, future studies might test hypotheses about relations between pointing and visual attention using eyetracking methods (e.g. Johnson & Mayer, 2012). Future studies using materials unfamiliar to participants might also measure participants' test anxiety as a control variable, given the potential for its "worry" component to consume working memory resources (Zeidner 1998).

The present study has some limitations which should be addressed in future research. Firstly, as the study was planned as the first in a series on the gesture effect, we used an extreme-groups design (see Mayer 2008, for a discussion) to maximise the difference between groups on the independent variable. Thus, the no-gesture participants were instructed to sit on their hands, to ensure maximal differences between the conditions. This focus on internal validity comes at the expense of the ecological validity of the findings - students rarely attempt to learn sitting on their hands. In future research, we plan to relax this constraint and include a control group that is merely instructed to study the materials but not gesture. Secondly, while the point estimates of the effect of gesturing on terminology and comprehension learning can be classed as high, the confidence intervals around these effects were wide, reflecting the relatively small sample (N=42). These confidence intervals represent a range of plausible population values of the gesture effect; thus, at present, it might be argued that there is substantial uncertainty about the magnitude of this effect, necessitating further replications of this study. On the other hand, if the gesture instructions used in the present study represent a method to reduce extraneous cognitive load by assisting in searching and matching for associated referents, then the magnitude of the effects seen here are consistent with those found for complex learning materials in Ginns' (2006) meta-analysis of the splitattention effect: meta-analytic d=0.78, 95% CI [0.69, 0.87]. Recall, however, that the results of the present experiment were obtained using materials in which standard methods to reduce search and match processing (e.g. text and diagram integration, arrows, numbering) had already been used. Thus, we would expect the effect of gesturing to be even greater if used with high split-attention materials. Lastly, the participants in this study were not enrolled in health science and medicine studies, raising potential questions about the generalisability of these results. From a CLT perspective, however, these participants are functionally equivalent to the student or professional who is placed in a situation where s/he must learn an unfamiliar topic. Given the ongoing expansion of knowledge related to human anatomy and physiology, the present results suggest an active learning strategy that should be easy to implement across a career, whenever novel text-based and diagrammatic information needs to be comprehended and learnt.

In summary, this study found explicit instructions to gesture enhances learning of complex materials of the kind regularly studied in the health sciences and medicine; however, the hypothesised role of cognitive load remains unclear. Further work is required to develop more effective ways to objectively measure cognitive load during learning and directly test whether learners who gesture find the material less cognitively demanding. The incorporation of instructions to gesture while students are learning novel materials may have a large positive effect on learning outcomes.

### Glossary

**Cognitive load**: Mental demands placed on a learner by a particular task.

**Gesture**: Motions made by the limbs, head or body to express thoughts or emphasize speech.

#### Notes on contributors

LUCY MACKEN, BA, BPsych, MTeach (Hons I), is a Secondary School Teacher at Kambala School, Sydney, Australia. The study described here was undertaken as part of an Honours degree in Education.

PAUL GINNS, BA (Hons I), MPsych (App), Grad Cert (Ed Studies), PhD, is a Senior Lecturer in Educational Psychology. His research applies cognitive psychology to instructional design.

**Declaration of interest**: Neither author has a declaration of interest to report.

#### References

Abrams R, Davoli C, Du F, Knapp III W, Paull D. 2008. Altered vision near the hands. Cognition 107:1035–1047.

Alibali M, Nathan M. 2012. Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. J Learn Sci 21: 247–286.

Cierniak G, Scheiter K, Gerjets P. 2009. Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? Comput Human Behav 25:315–324.

Cohen J. 1988. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: L. Erlbaum Associates.

Colonnesi C, Stams G, Koster I, Noom M. 2010. The relation between pointing and language development: A meta-analysis. Dev Rev 30: 352–366.

Cosman J, Vecera S. 2010. Attention affects visual perceptual processing near the hand. Psychol Sci 21:1254–1258.

de Koning B, Tabbers H, Rikers R, Paas F. 2009. Towards a framework for attention cueing in instructional animations: Guidelines for research and design. Educ Psychol Rev 21:113–140.

Dwyer F. 1972. A guide for improving visualized instruction. Pennsylvania: State College.

Ginns P. 2006. Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. Learn Instruc 16:511–525.

Hattie J. 2009. Visible learning. London: Routledge.

Hulme C, Monk A, Ives S. 1987. Some experimental studies of multisensory teaching: The effects of manual tracing on children's pairedassociate learning. Br J Dev Psychol 5:299–307.

Johnson C, Mayer R. 2012. An eye movement analysis of the spatial contiguity effect in multimedia learning. J Exp Psychol Appl 18: 178–191.

Joseph JH, Dwyer FM. 1984. The effects of prior knowledge, presentation mode and visual realism on student achievement. J Exp Ed 52:110–121.

Justice L, Ezell H. 2004. Print referencing: An emergent literacy enhancement strategy and its clinical applications. Lang Speech Hear Serv Sch 35:185–193.

Kalenine S, Pinet L, Gentaz E. 2011. The visual and visuo-haptic exploration of geometrical shapes increases their recognition in preschoolers. Int J Behav Dev 35:18–26.

Leppink J, Paas F, Van Der Vleuten C, Van Gog T, Van Merriënboer J. 2013.
Development of an instrument for measuring different types of cognitive load. Behav Res Meth 45:1058–1072.

Mayer R. 2008. Old advice for new researchers. Educ Psychol Rev 20:19–28. Montessori M. 1912. The Montessori method. London: Heineman.

Moreno R, Reislein M, Ozogul G. 2010. Using virtual peers to guide visual attention during learning. J Med Psychol 22:52–60.

- Salomon G. 1984. Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributes. J Ed Psych 78:647–658.
- Spence C. 2010. Crossmodal spatial attention. Ann N Y Acad Sci 1191: 182–200.
- Sweller J. 2010. Element interactivity and intrinsic, extraneous, and germane cognitive load. Educ Psychol Rev 22:123–138.
- Sweller J, Ayres P, Kalyuga S. 2011. Cognitive load theory. New York: Springer.
- Tarmizi R, Sweller J. 1988. Guidance during mathematical problem solving. J Educ Psychol 80:424–436.
- 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. Educ Psychol Rev 21:21–30.
- Van Merriënboer J, Sweller J. 2010. Cognitive load theory in health professional education: Design principles and strategies. Med Educ 44: 85–93.
- Zeidner M. 1998. Test anxiety: The state of the art. New York: Plenum Press

Copyright of Medical Teacher is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.