Dwyer, C., Hogan, M. J., & Stewart, I. (2010). 'The evaluation of argument mapping as a learning tool: Comparing the effects of map reading versus text reading on comprehension and recall of arguments'. *Thinking Skills And Creativity, 5*(1), *16-22.*

Abstract

The current study compared the effects on comprehension and memory of learning via text versus learning via argument map. Argument mapping is a method of diagrammatic representation of arguments designed to simplify the reading of an argument structure and allow for easy assimilation of core propositions and relations. In the current study 400 undergraduate volunteers were presented with a color map, black-and-white map, or text version of a large (50 bit) or small (30 bit) argument that centred on the question: Can Computers Think? Argument comprehension and memory was tested immediately after a 10 minute study period. Results indicated that participants who studied the argument maps scored more highly than those who studied text, on tests of memory, though not comprehension, and that participants had more difficulty assimilating the large (50 bit) argument in the time allotted. Results are discussed in light of research and theory on human learning and memory.

The purpose of educational text material is to be read and understood. In addition, in educational contexts, the material must often be memorized as well; more specifically, it must be encoded and stored and be available for recall in an exam context. However, for long pieces of text, learning can be difficult because the creation of an integrated representation in long-term memory is constrained by ongoing storage limitations in working memory (Cowan, 2000; Miller, 1956).

Limitations in working memory may constrain long-term memory, but working memory is also thought to be very flexible. Classic definitions of working memory describe a space for both the storage *and* manipulation of information (Baddeley, 1986). It has long been argued that the ways in which information is manipulated during encoding is critical for later recall (cf. Shimmerlik, 1978). Organizational strategies during encoding account for one such set of manipulations. Organizational strategies are believed to be beneficial in the sense that they provide a retrieval scheme that facilitates recall. For example, by 'chunking' related bits of information, recall of one part of a chunk aids retrieval of other parts of the same chunk, thus serving as a retrieval cue (Miller, 1956; Shimmerlik, 1978).

Various organizational strategies have been devised to enhance long-term retention of information, including, for example, summarisation (Kintsch & van Dijk, 1978), drawing inferences, generating questions, monitoring for understanding (Dole et al., 1991), constructing representational imagery (Pierce, 1980), and hierarchical summarisation (Taylor, 1982). The bulk of research on organization and memory has examined story memory, and in this regard, evidence suggests that when to-be-remembered story events are presented in a well-organized manner, the level of free recall is better than when the events are presented in a random order (Bower et al., 1969; Myers, 1974). Also, readers who are sensitive to text structure recall more information than readers who are not (Meyer, Brandt & Bluth, 1980).

Some researchers have suggested that because it is too memory intensive to remember everything from a passage of text, a macrostructure, or the ‘gist’ of the text, is stored in long-term memory, and this represents the summary information a reader considers important (Kintsch & van Dijk, 1978). Hence, it is this macrostructure, and not the original text that the reader remembers when later asked to recall the text (Kintsch & van Dijk, 1978). The problem with this learning strategy is that although the formulation of a macrostructure presumably facilitates recall of information, it is likely that information is not encoded at a very deep level of specificity; in other words, the detail of propositions and of relations between propositions will probably not be remembered.

Hierarchical summarisation is a more explicit, active organisational strategy. It involves extracting and summarising the key themes and sub-themes in a text. Taylor (1982) found that the use of hierarchical summarisation increased recall of text in students who were trained in the use of the technique. A similar study by Berkowitz (1986) provides a rare example of how organizational strategies can be used to facilitate learning of prose arguments. Berkowitz taught students to construct maps of prose passages. Using this mapping strategy, the main ideas from the passages were summarised in separate boxes and supporting claims were listed as bullet points beneath each of the main ideas. The boxes were organized in a radial structure (i.e., around a central claim). Berkowitz found that for students who used this technique overall recall of passages was significantly improved relative to students who used traditional study techniques (i.e., question-answering and rereading procedures).

Although Berkowitz described her maps as a graphic representation of the superordinate and some of the more important subordinate ideas in a passage, organized in a manner similar to the way the author organized them in the original selection, the propositional content of the radial maps did not represent fully planned arguments. Also, although Berkowitz attempted to construct maps that corresponded to the way the author organized ideas in the original selection, the radial structures in no way reflected the structure of the argument (see Twardy, 2004 for a discussion of the text to argument map translation process). Finally, Berkowitz provided extensive training to students in the use of the mapping technique. Although training in the use of organization techniques and active manipulation of arguments is a crucial part of learning (Van Gelder, Bissett, & Cumming, 2004), a critical question is whether or not more explicit, complete, logical, hierarchically structured argument maps that faithfully represent the structure of an argument can be packaged as an optimal method of *presenting* information to students who are asked to read, understand, and remember. To date, no study has examined this issue.

Building a mental representation of an argument

When it comes to analysing arguments, the problem with traditional text-based learning is that it does not allow one to readily connect statements that support and dispute specific reasons. The learner must engage in a cognitively demanding process of linking propositions that are located in different paragraphs, on different pages, and so on. When reading a text-based argument, the reader must mentally construct the argument, thus switching attention away from the information presented in the text. In a series of seminal studies, Pollock, Chandler & Sweller (2002) found that learning is impeded when instructional materials require a high degree of attention switching, for example, between text and figures. They concluded, more generally, that encoding environments that increase the cognitive burden (or load) placed on the reader tend not only to slow the learning process, but also reduce overall levels of learning. It is an untested assumption, but, presumably, the provision of a good visual representation of an argument structure will reduce the cognitive burden associated with conjointly reading and mentally representing the structure of an argument. The provision of this critical organizational, structural information should in turn facilitate argument comprehension and memory.

Argument Mapping

An argument map organises the arguments in a text into a hierarchical representation, often pyramid-shaped, with propositions arranged in coloured boxes and connected by arrows that highlight relations between propositions (van Gelder, 2002). Having available the structure of an argument is crucial for many reasons; it facilitates logical reasoning, the ready construction of a 'mental image' of the whole argument, and the answering of specific questions about the relation between one proposition and others. Theoretically at least, representation of an argument using an argument map should remove obstacles to learning related to the need to simultaneously read the text of an argument and mentally visualize the relational structure of the argument being presented. However, there has been very little research into argument mapping as a learning aid, and, in particular, no study has examined whether or not the reading of argument maps facilitates argument comprehension and recall.

The Current Study

Based on the research conducted to date, we hypothesized that, given the same set of arguments to read, argument map reading would facilitate superior comprehension and memory when compared with conventional text reading.

Notably, an argument map contains all the propositions contained in a homologue text, but it also highlights the structure of reasoning and indicates explicitly to the reader that one box contains a proposition that supports or refutes the proposition in another box.

A good argument map also makes use of Gestalt grouping principles, and research suggests that when to-be-remembered items are grouped according to Gestalt cues, they are better stored in visual working memory (Woodman, Vecera & Luck, 2003). For example, a good argument map will place related propositions close to one another, thus complying with the Gestalt grouping principle of proximity. A consistent colour system can also be used to highlight propositions that support (green box) or refute (red box) the central claim, thus complying with the Gestalt grouping principle of similarity. In this context, a secondary hypothesis we tested was that argument maps that contain colour (to distinguish 'reasons' from 'objections' in the argument structure) would facilitate better comprehension and recall when compared with argument maps that contain no colour (i.e., black-and-white maps). Finally, we hypothesized that the benefits of argument map reading over text reading would be greater for larger argument structures. To test this hypothesis we had students read and then answer questions in relation to argument maps and homologue texts that contained either 30 or 50 propositions.

Methodology

Design

A 2x3 between subjects MANCOVA was used to assess the effects of study materials on comprehension and recall of arguments, while controlling for baseline verbal and spatial reasoning ability. The first between-subjects factor was cognitive load (30-bit versus 50-bit argument structure). The second between-subjects factor was study materials (text, black-and-white argument map, argument map with colour).

Participants

Participants were first year psychology students (N = 420), aged between 17 and 25 years from the National University of Ireland, Galway. The study was conducted as part of a core module: *Psychology in Practice*. To ensure confidentiality, students were identified by student ID number only.

Materials

The Verbal and Spatial Reasoning subtests of the Differential Aptitude Test (Bennett, Seashore & Wesman, 1947) were used to assess participants' reasoning ability. The measure was used as a covariate in the analysis of the effects of study materials on comprehension and memory. Experimental reading materials were constructed by extracting a sub-set of the arguments presented in Robert Horn's (1999) argument map: *Can computers think?* Six sets of study materials were developed including: (1) a 30-proposition text; (2) a 30-proposition colour argument map; (3) a 30-proposition monochrome argument map; (4) a 50-proposition text; (5) a 50-proposition colour argument map; and (6) a 50-proposition monochrome argument map. Memory was assessed using a 'fill-in-the-blanks' cued recall test that assessed memory for reasons and objections linked to each of the major arguments supporting or refuting the central claim: *computers can think*. Comprehension was assessed by asking students to decide if a sub-set of 12 of the original propositions (present in all six study conditions) either supported or refuted the central claim.

Procedure

The study took place over three weeks. Each class lasted for 50 minutes. In week 1, participants were provided with an introductory lecture on critical thinking.

The idea of argument mapping was introduced and explained. In week 2, two aptitude tests, the Verbal and Spatial Reasoning subtests of the Differential Aptitude Test (DAT, Bennett, Seashore & Wesman, 1947), were administered. The DAT is often administered in a group testing context, and to ensure optimal test conditions two invigilators were present to monitor performance. When participants returned to class the following week they were randomly assigned to one of the six study conditions outlined above. Study materials were distributed. Participants were allotted 10 minutes to read the materials and were instructed to learn the material with a view to being tested. Pilot research suggested that 10 minutes was sufficient time to read any of the versions of the text/map provided. After 10 minutes had elapsed, study materials were collected and the cued recall test was administered. Finally, participants were given the comprehension test to complete. Total testing time in both weeks 2 and 3 of the study was approximately 40 minutes. Participants were debriefed and thanked at the end of Week 3 and results of the study were reported back to them a month later.

Results

Table 1 lists means and standard deviations for memory and comprehension test performance for each of the six study conditions. In the case of memory, analysis revealed main effects for both study materials ($F[2, 358] = 7.73$, $\eta^2 = 0.041$, $p < 0.001$) and cognitive load ($F[1, 358] = 64.08$, $\eta^2 = 0.152$, $p < 0.001$), whereas in the case of comprehension no significant effects were found for either study materials ($F[1, 364] = 0.69$, $\eta^2 = 0.002$, $p = 0.58$) or cognitive load ($F[2, 364] = 0.56$, $\eta^2 = 0.003$, $p = 0.41$). There were no interaction effects.

Post-hoc analyses in the case of memory revealed a significant difference in favour of both colour argument maps and black-and-white argument maps over standard text ($p < .05$ for both), with no significant difference between black-and-white and colour argument maps. Overall, memory was better for those participants who read the smaller (30-bit) argument structure, and post-hoc analysis revealed that this difference was significant for all three types of study materials: small text versus large text, small black-and-white maps versus large black-and-white maps, and small colour maps versus large colour maps ($p < .05$ for all three).

Insert table 1 around here

Discussion

The findings from this experiment suggest a dissociation between comprehension and recall of arguments as dependent variables in the cases of both stimulus set size and study materials as independent variables. Although the comprehension test was sensitive to individual differences (i.e., no ceiling or basement effects were observed, Mean = 6.78, SD = 2.14, Range = 12) and correlated with verbal aptitude (as assessed using the verbal reasoning subscale of the Differential Aptitude Test, $r = .35$, $p < .001$), the number of arguments a reader was asked to assimilate and the way in which those arguments were presented (map versus text) had no effect on comprehension test performance. However, when participants were asked to recall the arguments that both supported and refuted specific sub-claims, recall memory was poorer when participants were asked to remember from a larger stimulus set (i.e., a 50-bit versus a 30-bit argument structure). Furthermore, the provision of a structural representation of the argument (i.e., argument maps, those both with and without colour cues that were used to distinguish reasons from objections) supported better recall.

These results suggest that, when compared with traditional text-based information delivery methods, argument mapping significantly increases subsequent memory for arguments. Further research is necessary in order to discover the conditions that influence levels of memory performance in this context. Further research is also needed to discover conditions that might produce an effect for comprehension in this context.

In relation to comprehension, it may be that those with higher verbal and spatial reasoning test scores were more likely to engage in some form of critical thinking during the reading of arguments. This would explain the correlation between differential aptitude reasoning ability and both text and map reading comprehension scores, and it would suggest that a delivery method that requires active critical thinking (as opposed to passive reading of material) would enhance comprehension (Taylor, 1982; Berkowitz, 1986). The comprehension test required students to decide if a sub-set of 12 of the original propositions either supported or refuted the central claim *Computers can think*. It was a test that required more than simple memory for individual propositions: it required understanding of the relationships between propositions in the argument structure. Some students found this task surprisingly difficult and scored lower than chance levels (i.e., they made poor guesses and were incorrect more than 50% of the time). Although it is surprising that argument map reading did not facilitate performance in this context, it may be that over-and-above baseline reasoning ability and the spontaneous critical thinking efforts of participants, the reading of argument maps does not motivate an additional tendency to critically engage with an argument, at least not for novice map readers. Some training in the analysis of arguments – using argument maps as study materials – may be necessary to engage students in the deeper relational analysis of maps that is necessary for good performance on tests of comprehension.

Nevertheless, in the presence of a similar significant correlation between verbal reasoning and memory performance ($r = 0.27$, $p < .001$), and controlling for this co-variation, there remained a significant effect of stimulus materials on memory performance. In other words, although verbal reasoning ability also predicted memory performance, unlike for comprehension tests performance the reading of argument maps provided some additional benefits in terms of memory for individual propositions in a cued-recall test. There are a number of possible reasons for this finding. First, it may be that the structure of the cued-recall test provided a visual homologue of the argument mapping encoding condition. Specifically, in the argument maps, sub-ordinate (to-be-remembered) propositions were presented beneath super-ordinate propositions; in the cued-recall test, the super-ordinate proposition was presented and a number of blank spaces (corresponding to the number of sub-ordinate propositions to be remembered) was provided for students to fill in. Second, it may be that a certain amount of information can be remembered in the absence of truly understanding the relational structure of an argument, and that argument maps facilitate memory for discrete propositions in this context by highlighting each proposition in a distinct box. Text does not demarcate discrete propositions in the same way, chunking them together into larger paragraph units instead. Thus, it may be that the benefits for memory of reading argument maps differs significantly when novice and expert map readers are compared. For example, it may be that novice map readers will demonstrate some benefit over text readers when memory for discrete propositions is tested, whereas expert map readers will demonstrate some benefit over text readers for both discrete propositional memory and more complex forms of relational memory (as tapped into by comprehension tests, essay-writing tests, ad lib debating exercises, etc.).

Although both basement and ceiling effects were avoided for the memory test, it is clearly the case that participants responded less well when asked to read and recall a large 50-bit argument. The data suggest that there is an upper limit to the number of arguments that can reasonably be assimilated in a short space of time. Being asked to recall 12 target memories from a set of 50, after a study period of 10 minutes, is much more difficult than being asked, given a similar study period, to recall the same 12 target memories from a set of 30. This is consistent with the vast body of literature pointing to the superiority of distributed over mass learning (Fiore & Salas, 2007). If facts and relations contained in complex argument structures are delivered in manageable chunks, and if each of these smaller chunks is later reviewed before proceeding, it is more likely that complex argument structures will be successfully encoded and later recalled. Consistent with Cognitive Load Theory (Pollock, Chandler & Sweller, 2002) study or work environments that overburden the cognitive system will be associated with poor learning outcomes.

Future research should examine the effects of explicit training in argument mapping (i.e., the generation of argument maps) on comprehension, memory, and general reasoning ability, based on the theoretical suggestion that the active construction of maps can facilitate deeper comprehension of arguments and thus better subsequent memory, and that the development of argument structures can promote a greater appreciation of the nature of reasoning more generally.

Such research would provide an important further test of the effect of argument mapping on learning and, if results favoured this instrument, would constitute strong empirical evidence in favour of van Gelder’s (2002) contention that argument mapping can improve critical thinking. Overall, the results of the current study reveal argument mapping as a very plausible technique for improving memory, and a technique that might prove very beneficial in the educational arena. Further research is needed before we understand the conditions that best foster argument comprehension and memory through argument mapping, and hence the best ways in which to maximise the potential of this methodology with respect to both learning and critical thinking.

References

Baddeley, A.D., Thomson, N. & Buchanan, M., (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behaviour, 14,* 575-589.

Baddeley, A. (1986). *Working Memory*. Oxford: Oxford University Press.

Bennett, G.K., Seashore, H.G. & Wesman, A.G. (1947). *Differential Aptitude Test, verbal reasoning subtest, spatial reasoning subtest.* Windsor: Psychological Corporation.

Berkowitz, S.J. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly, 21,* 2, pp. 161-178.

Bower G.H., Clark, M.C., Lesgold, A.M. & D. Winzenz. (1969). Hierarchical retrieval schemes in recall of categorised word lists. *Journal of Verbal Learning and Verbal Behavior, 8*, 23-343.

Buckingham Shum, S.J., (2002), *The roots of computer supported argument visualization.* Knowledge Media Institute, Open University, UK.

Chi, M.T.H., Glaser, R. & Rees, E. (1982). Expertise in problem solving. In R.S. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (pp.7-77). Hillsdale, NJ: Erlbaum.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences, 24*, 87-185.

Dee-Lucas, D. & Larkin, J.H. (1995). Learning from electronic texts: Effects of interactive overviews for information access. *Cognition and Instruction, 13(*3), 431-468.

Dole, J.A., Duffy, G.G.,Roehler, L.R. & P.D. Pearson. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61,* 239-264.

Duchastel, P.C. (1990). Examining cognitive processing in hypermedia usage. *Hypermedia, 2*, 221-233.

Facione, P.C. & Facione, N.A. (1992). *California Critical Thinking Skills Test*. California Academic Press, California.

Fiore, S. M., & Salas, E. (2007). *Toward a science of distributed learning*. Toward a science of distributed learning. xi, 284 pp. Washington, DC: American Psychological Association.

Hanf, M.B. (1971). Mapping: A technique for translating reading into thinking. *Journal of Reading, 14*, 225-230.

Hogan, M.J., (2006). Against didacticism: a psychologist's view. *Educational Research and Reviews, 1*(7), 206-212.

Horn, R. E. (1999). Can Computers Think?: MacroVU, Inc.

Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kotovsky, H.A., Hayes, K. & J.R. Simon. (1985). Why are some problems hard? Evidence from the Tower of Hanoi. *Cognitive Psychology, 17* (22), 248-294.

Miller, G.A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, *63*, 814-97.

Meyer, B.J.F., Brandt, D.M. & G.J. Bluth. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 16*(1), 72-103.

Myers, J.L. (1974). *Memory for prose material*. Amherst, Mass: University of Massachusetts.

Pierce, J.W. (1980). Field independence and imagery-assisted prose recall of children. *Journal of Educational Psychology, 68*, 355-359.

Pollock, E; Chandler, P. & Sweller, J., (2002) Assimilating complex information. *Learning & Instruction*, *12*, 61-86.

Shimmerlik, S.M. (1978). Organization theory and memory for prose: A review of the literature. *Review of Educational Research, 48*(1), 103-120.

Sweller, J. (1999) Instructional design in technical areas. *Australian Education Review No. 43.* Victoria: Acer Press.

Sweller, J. & Chandler, P., (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351-362.

Taylor, B.M. (1982). Text structure and children's comprehension and memory for expository material. *Journal of Educational Psychology, 74,* 323-340.

Toulmin, S. (1958). *The uses of argument.* Cambridge: Cambridge University Press.

Van Gelder, T., (2002). Argument mapping with Reason!able. *APA Newsletter: Philosophy & Computers*, *2*(1), 85-90.

Van Gelder, T., (2001). How to improve critical thinking using educational technology. *ASCILITE 2001,* (http://www.medfac.unimelb.edu.au/ascilite2001/pdf/papers/vang%/eldert.pdf).

Van Gelder, T., (2003). Enhancing deliberation through computer supported argument mapping. *Computer Simulated Cooperative Work***,** 97-115.

Van Gelder, T; Bissett, M. & Cumming, G. (2004). Enhancing expertise in informal reasoning. *Canadian Journal of Experimental Psychology, 58*(2), 142-152.

Wigmore, J.H. (1913). *The principles of judicial proof: as given by logic, psychology and general experience and illustrated in judicial trials.* Boston: Little, Brown.

Woodman, G.F., Vecera, S.P. & S.J. Luck. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review, 10*(1), 80-87.

Table 1: Descriptive Statistics for Comprehension & Memory

	Small Map		Large Map	
Condition	*M*	*SD*	*M*	*SD*
Comprehension				
Text	7.28	2.12	6.45	2.11
Black & White Map	6.86	2.28	6.86	2.12
Colour Map	6.47	2.16	6.74	1.94
Memory				
Text	6.36	4.39	2.76	2.85
Black & White Map	7.60	4.42	4.28	3.14
Colour Map	7.87	4.48	5.13	2.53