

---

## The influence of modality on deep-reasoning questions

---

Jeremiah Sullins\*, Scotty D. Craig and  
Arthur C. Graesser

University of Memphis,  
202 Psychology Building,  
Memphis, TN 38152, USA  
E-mail: jsullins@memphis.edu  
E-mail: scraig@memphis.edu  
E-mail: a-graesser@memphis.edu

\*Corresponding author

**Abstract:** This study investigated the influence that modality (print versus spoken text) had on learning with deep reasoning questions. Half the participants were randomly assigned to receive deep-reasoning questions during the learning session. The other half received the same information in the absence of deep-reasoning questions. The participants who received deep reasoning questions were randomly assigned to one of two different groups. One group received deep reasoning questions as on-screen printed text while the other group received deep reasoning questions in a spoken modality via a text to speech engine. Participants who received deep reasoning questions had higher post-test scores than those who did not, a finding that replicated previous research. Additionally, learning was better for the learners who received printed text than spoken messages, a finding that is not compatible with a number of theoretical and empirical claims in the literature.

**Keywords:** vicarious learning; deep reasoning questions; multimedia; tutoring; computer-based learning environment; modality; spoken versus typed.

**Reference** to this paper should be made as follows: Sullins, J., Craig, S.D. and Graesser, A.C. (2010) 'The influence of modality on deep-reasoning questions', *Int. J. Learning Technology*, Vol. 5, No. 4, pp.378–387.

**Biographical notes:** Jeremiah Sullins is currently a Doctoral student in the Institute for Intelligent Systems at the University of Memphis. His research interests include intelligent tutoring environments, student generated questions, and emotions.

Scotty D. Craig is a Research Scientist in the Institute for Intelligent Systems located at the University of Memphis. To date, he has worked on projects in such areas as affect and learning, discourse processing, mechanical reasoning, multimedia learning, vicarious learning environments and intelligent tutoring systems.

Arthur C. Graesser is currently a Full Professor in the Department of Psychology, an Adjunct Professor in Computer Science, and the Co-director of the Institute for Intelligent Systems at the University of Memphis. His primary research interests are in cognitive science, discourse processing, and the learning sciences. More specific interests include knowledge representation, question asking and answering, tutoring, text comprehension, inference generation, conversation, reading, education, memory, expert systems, artificial intelligence, and human-computer interaction.

## **1 Introduction**

Does the modality in which students receive information in a vicarious learning environment significantly affect how much a student learns? In addition,, if deep reasoning questions (described in subsequent sections) are inserted into the content delivered to the student, does that have any impact on student learning gains?

A vicarious learning environment is an environment in which learners are not the addressee of the material and/or they do not have control over the material they are expected to master. The concept of vicarious learning in the psychology of human learning dates back to Bandura's early work modelling aggression with children (Bandura, 1962). It has continued under such labels as observational learning and social learning (Lee et al., 1998; Rosenthal and Zimmerman, 1978). According to current vicarious learning theory, it is not necessary for the learner to be physically active, only cognitively active, for knowledge acquisition to occur. Previous research has found vicarious learning environments to be an effective source of information delivery that significantly increases students' learning when compared to various controls (Gholson and Craig, 2006).

The primary focus of the current study is to explore the possible effects of modality (spoken versus typed) on the content delivered within a vicarious learning environment. Furthermore, we will explore any possible relations that exist between deep reasoning questions inserted into the content and students' learning gains.

### *1.1 Previous findings on deep reasoning questions*

Researchers have investigated the impact of deep reasoning questions on learning in recent years (Gholson and Craig, 2006; Craig et al., 2006). A deep reasoning question is one which integrates content and fosters understanding of the components and mechanisms being covered (Craig et al., 2008). These questions are aligned with the higher levels of Bloom's (1956) taxonomy and the long-answer question categories in the question taxonomy proposed by Graesser and Person (1994). In order to illustrate the difference between a deep reasoning question and a shallow reasoning question, consider an example of each. An example of a shallow question, according to Graesser and Person (1994), would be "Does the CPU use RAM when running an application?". The reason for categorising this type of question as shallow is because it does not take much intrinsic thought on the student's part. It is a simple 'yes' or 'no' question. In contrast, an example of a deep reasoning question would be "How does the CPU use RAM when running an application?". The reason for categorising this question as 'deep' is because the student must take the knowledge he or she knows about computers, and make a connection to compare the differences between the two.

Researchers have recently tested claims about deep reasoning questions by pedagogical agents and their impact on learning (Craig et al., 2006; Sullins et al., 2007; Sullins and Azevedo, 2007). In one study, Craig et al. (2006) randomly assigned students to one of four different conditions. The interactive condition had students interact with an intelligent tutoring system on 12 subtopics on computer literacy. These students' sessions were recorded and saved as a video file. The *yoked-vicarious* condition had students watch a video file from other student's interaction with the intelligent tutoring session. A condition involving a *deep-level reasoning questions monologue* consisted of students watching an animated pedagogical agent asking deep reasoning questions, and then the

same agent providing the answers to the questions. A fourth *deep-level reasoning questions dialogue* condition had an animated pedagogical agent ask deep reasoning questions; the answers to those questions were then answered by a separate distinct voice (at all times, only one agent was located on the monitor). Analyses revealed that the students who were assigned to the deep-level reasoning questions conditions (both monologue and dialogue) significantly outperformed all other conditions. Furthermore, there were no statistically significant differences between the two deep reasoning questions conditions. These results suggest that by simply lacing in deep reasoning questions into the content, students will show marked improvements in learning gains.

### *1.2 Previous findings with spoken narration vs. on-screen printed text*

Researchers have explored differences between on-screen printed text and spoken narration in multimedia environments (Mayer, 2005; Moreno and Mayer, 1999; Graesser et al., 2003). Potential differences lie in that nature of the mental representations that get constructed through the two media. One example would be the dual-coding hypothesis (Sadoski and Paivio, 2001; Paivio, 1986). The dual-coding hypothesis is based on Paivio's model that postulates that both visual and verbal information are processed differently and along distinct channels. The human mind then creates separate representations for the information processed in each channel. Proponents of the dual-coding hypothesis hold that auditory and visual information (pictures or words) are processed in separate streams (Penny, 1989). The dual-coding hypothesis predicts that information is more effectively processed if both auditory and visual working memories are used. The mechanism of working memory is believed to be an important component for explaining potential differences between spoken and printed media.

A second example would be the working memory hypothesis based on Baddeley's (1992) model that consists of a visuo-spatial sketchpad, and a phonological loop. The visuo-spatial sketchpad is used for such things as diagrams or real objects while the phonological loop is concerned with auditory information and language. According to Baddeley's working memory model, learning is more effective if both channels are utilised resulting in less cognitive load for either channel.

Another area of research that has explored the relationship between spoken and visually presented content is Mayer's modality effect. Studies on the modality effect have shown that when the mode of presentation has been auditory as opposed to visual, there is superior recall for auditory information (Penny, 1989). Over the past decade, Mayer and colleagues have conducted numerous studies examining the effects of spoken narration versus on-screen printed text. For example, Moreno and Mayer (1999) discovered that in six of six experiments students were better able to transfer what they had learned to new problems when animation was accompanied by spoken words (narration) than by printed words (on-screen text). The six experiments involved explanations of how lightning forms, how brakes work, and how plants grow. Mayer's theoretical rationale for this principle is that the learner's visual channel might be overloaded when words and pictures are presented to the learner.

However, not all research supports the claim that presenting auditory information is superior compared to text-based content whether in the presence on a diagram or not (Craig et al., 2002). Craig et al. (2004) conducted a study examining on-screen printed

text vs. spoken narration in the absence of an animated pedagogical agent. Participants viewed a short animation on the formation of lightning and after watching the animation, all participants took a post-test with questions of retention, matching, and transfer. Results revealed that participants who viewed the animation as on-screen printed text outperformed, although not significantly, the participants who viewed the animation as spoken narration.

Furthermore, there is evidence in the literature that suggests, contrary to Mayer's (2001, 2005) work, that students who receive information in a written modality will learn significantly more than students who receive the same information in a spoken modality (Clark, 1996). Clark's theoretical rationale is that if learners miss a word or phrase they are unable to retrieve the information due to spoken narration being ephemeral.

### *1.3 Predictions*

Based on the previously mentioned literature, there are a number of possible outcomes regarding the current study. First, based on the deep reasoning questions effect (Craig et al., 2006), we can predict that adding these type of questions will improve learning compared to students who receive equivalent information minus deep reasoning questions.

Regarding the spoken versus typed condition, there is available research to suggest two potential outcomes. A *spoken narrative hypothesis* is based on Mayer's (2005) modality principle. This principle says that students learn more when receiving narration and animation where students can hear explanations while viewing pictures and video rather than on-screen printed text and animation where students must read text along while simultaneously viewing pictures or video. According to the spoken narrative hypothesis, we would expect in the current study that the participants that receive information as spoken narration will significantly outperform participants that receive the same information as on-screen printed text.

The *visual text hypothesis* which is based on Clark's (1996) work suggests that students who receive information in a text-based format will show greater learning gains because students have a chance to reread text, whereas speech is evanescent. A second prediction, based on this hypothesis, would be that students who receive information as on-screen printed text would significantly outperform all other conditions.

### *1.4 Current study*

The current experiment was designed primarily to compare learning gains obtained using deep reasoning questions presented as spoken narration via a voice engine with the same deep reasoning questions presented via on-screen printed text. More specifically, in this study we were interested in two research questions. First, does presenting students deep reasoning questions facilitate learning when compared to various controls? Second, does the modality in which deep reasoning questions are presented to students affect learning gains?

## 2 Methods

### 2.1 Participants

The 120 participants were undergraduate students from the University of Memphis enrolled in an introductory psychology course. Students who had low domain knowledge (i.e., below 40% on a pre-test) were selected for inclusion into the study. Low knowledge students were included in this study because previous research has shown that design principles are most pronounced for low knowledge learners (Mayer, 2005).

### 2.2 Materials and design

During the study, students were tutored on the topic of computer literacy. Throughout the session, learners received information on 12 subtopics regarding hardware (e.g., “How is RAM related to the CPU?”). A 2 (Dialogue type) X 2 (Modality type) factorial design was implemented for this study. Our dialogue type factor consisted of two levels: deep-reasoning questions versus monologue. The modality factor’s levels for this study consisted of materials being presented as an auditory narration or printed on the computer screen. In the *deep questions dialogue* condition, participants were given deep reasoning questions laced into the content throughout the learning session. The *monologue* condition, presented participants the exact same content as the learners in the deep questions condition minus the deep reasoning questions. The *spoken* text type presented learners content in a spoken modality via a voice engine, but no animated agent was present during the presentation. A standard Microsoft voice engine was used in the study. The *written* text type presented learners the same content as on-screen printed text.

### 2.3 Knowledge tests

Two sets of multiple choice questions were used to assess the learning gains of the participants. The multiple choice tests, which tapped deep knowledge on the domain of computer literacy, consisted of 24 questions (see the Appendix for example questions). The two tests were counterbalanced among participants in each condition as pre-test and post-test.

### 2.4 Procedure

Participants were first given the computerised pre-test consisting of 24 multiple choice questions on computer literacy. The participants completed the test at their own pace, but were not allowed to go back to previously answered questions. Students were selected only if they had low domain knowledge on the topic of computer literacy, based on their pre-test scores (below 40%). Following the pre-test, the participants were then seated at a computer for approximately one hour. During the presentation students were presented with information on computer literacy via a computer monitor. Students were instructed to pay attention during the presentation and not to touch the keyboard. Immediately, following the computerised presentation, participants were then given a computerised post-test consisting of 24 multiple choice questions assessing their knowledge on computer literacy.

## 2.5 Results

Table 1 presents data on the pre-test and post-test scores as a function of the conditions that manipulated dialogue type and presentation modality. A 2 X 2 ANCOVA was conducted on the post-test scores, using pre-test scores as a covariate. It was predicted based on previous research (Craig et al., 2006) that students who received the deep reasoning questions would outperform the students who did not receive deep reasoning questions. Because of this prediction, a one-tailed test was used in the question analysis. The results indicated that participants who received deep reasoning questions laced into the content yielded higher post-test scores on the multiple choice test than did the participants who did not receive deep reasoning questions, with means of .50 and .45, respectively,  $F(1, 115) = 2.71, p < .06$ , one tailed,  $\eta^2 = .023$ . The ANCOVA showed a significant main effect of modality, with lower scores for spoken messages than printed text, within means of .44 and .52, respectively,  $F(1, 115) = 5.46, p < .05$ ,  $\eta^2 = .047$ . There were no significant interactions between the presence versus absence of deep reasoning questions and modality. Furthermore, analyses revealed large effect sizes for all learning conditions. Effect sizes were computed even though the participants were selected from a restricted range, i.e., only low knowledge students participated. Because of the restricted range effect sizes were calculated as  $[(M1 - M2)/\text{pre-test standard deviation}]$ .

**Table 1** Pre-test and post-test means, standard deviations and effect sizes for the different conditions

Condition	Pre-test		Post-test		Effect sizes	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Spoken	Questions	.27	.09	.47	.15	1.58
	Monolog	.27	.08	.42	.18	1.03
Written	Questions	.28	.07	.54	.16	2.02
	Monolog	.29	.06	.49	.14	1.81

## 3 Discussion

Results are consistent with past research showing that deep reasoning questions positively influence learning (Craig et al., 2006). However, it should be noted that the results were only marginally significant in a one-tailed statistical test. Moreover, results revealed that participants who received the answer content in a written modality scored significantly higher on post test scores than participants who received the same information in a spoken modality. Furthermore, no significant interactions were discovered.

Results appear to give support to the visual text hypothesis that predicted students who received information in a written modality would outperform students who received the same information in a spoken modality. These results contribute to the already rich multimedia research community (Mayer, 2001). At first glance our results would appear to contradict the previous findings presented by Mayer. More specifically, according to Mayer, students who received the on-screen printed text should have shown lower learning gains due to their working memory becoming overloaded. However, there were

differences between Mayer's methodology and the methodology used in the current study that warrant mentioning. First, in the current study there were a limited number of on-screen pictures that occurred throughout the presentation (one picture per subtopic). In Mayer's studies, pictures were visibly located on the screen at all times. A second difference was the time on task between the two studies. In the current study, participants observed the presentation for approximately 40 to 45 minutes. In Mayer's study, participants watched instruction on the formation of lightning for 180 seconds. An additional difference between our study and the work of Mayer is the content of the experiments (i.e., computer literacy versus lightning). Because of these differences, it becomes difficult to directly compare the two studies. Further research is required for direct comparisons. One possible reason that the text facilitates learning might be because the student has a chance to re-read the content whereas spoken discourse is evanescent (Clark, 1996). Research has shown that re-reading can be an effective meta-cognitive reading strategy (Rawson et al., 2000). For example, Johnson-Glenberg (2005) conducted a study in which 20 undergraduates read through expository text. Participants in the experimental group read the expository text while incorporating meta-cognitive strategies (e.g., building mental models, generating questions). Participants in the control group read texts and solved word anagram problems. Participants in both conditions were allowed to go back and re-read the text as much as they wanted. Results revealed that participants in the experimental condition re-read the text significantly more than students in the control condition. More specifically, the relatively poor comprehenders utilised this technique more often than the relatively high comprehenders.

Conversely, when students received the information in a spoken modality via a voice engine, if they missed a sentence or did not understand what was said they did not have the option to go back and listen again. Future research is needed to both replicate this finding and explore why and under what conditions presenting content in different modalities facilitates knowledge acquisition for students.

The results from the present study have many applicable implications for the design of intelligent tutoring systems and computer-based learning environments. These results suggest that when presenting multimedia information to students, it is beneficial to incorporate deeper level questions into the curricula. Furthermore, results suggest that by presenting information as on-screen printed text, students showed comparable learning gains to various controls. This is good news for researchers in the area of distant learning who might not have the resources to incorporate animated agents and text to speech engines. We have also shown that in the context of vicarious learning environments, that animated agents and text to speech engines are not needed. By simply lacing in deep reasoning questions and presenting content in a written modality, learning gains can be significantly increased.

Additionally, a growing number of universities are incorporating online courses (e.g., web-based statistics). Our results suggest encouraging possibilities for these types of courses. For example, many of these online courses consist of a large amount of text for students to read and comprehend. Based on our findings, by using on-screen printed text and incorporating deep reasoning questions, learning gains can significantly increase. Furthermore, our results found no benefit of using text to speech engines over traditional on-screen printed text. These are exciting results considering that not everyone has the technology or funding to incorporate this technology into their online curriculum.

As previously mentioned, future research is needed on a number of different issues. As mentioned in the results section our eta-squared for both independent variables was relatively small indicating that only a small amount of the variance was actually explained using the current manipulations. Future studies need to try to account for the larger unexplained variance. Additionally, in the current study, participant covered information on computer literacy which is a topic with a low degree of abstractness. Is it possible that these results will replicate in different domains that are considered to be highly abstract (e.g., algebra and statistics)? Furthermore, what effect does our intervention have on short term memory versus long term memory? More specifically, our post test materials were given immediately following the intervention. Will our results replicate on delayed post test materials? Although there are a number of different issues that require further investigation, our results contribute valuable information to the distance learning and web-based learning communities.

### **Acknowledgements**

This research was supported by the National Science Foundation (REC 0106965, ITR 0325428, REESE 0633918) and the Institute of Education Sciences (R305H050169, R305B070349). Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF or IES. I would also like to thank Joshua Brittingham for his help with the endless hours of data collection.

### **References**

- Baddeley, A.D. (1992) 'Working memory', *Science*, Vol. 255, pp.556–559.
- Bandura, A. (1962) 'Social learning through imitation', in Jones, M.R. (Ed.): *Nebraska Symposium of Motivation*, University of Nebraska Press, Lincoln, pp.211–269.
- Bloom, B.S. (1956) 'Taxonomy of education objectives: the classification of educational goals', *Handbook I: Cognitive Domain*, McKay, New York.
- Clark, H. (1996) *Using Language*, Cambridge University Press, Cambridge.
- Craig, S.D., Driscoll, D. and Gholson, B. (2004) 'Constructing knowledge from dialog in an intelligent tutoring system: interactive learning, vicarious learning, and pedagogical agents', *Journal of Educational Multimedia and Hypermedia*, Vol. 13, pp.163–183.
- Craig, S.D., Gholson, B. and Driscoll, D. (2002) 'Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features, and redundancy', *Journal of Educational Psychology*, Vol. 94, pp.428–434.
- Craig, S.D., Sullins, J., Witherspoon, A. and Gholson, B. (2006) 'Deep-level reasoning questions effect: the role of dialog and deep-level reasoning questions during vicarious learning', *Cognition and Instruction*, Vol. 24, No. 4, pp.563–589.
- Craig, S.D., VanLehn, K. and Chi, M.T.H. (2008) 'Promoting learning by observing deep-level reasoning questions on quantitative physics problem solving with Andes', in McFerrin, K., Weber, R., Weber, R., Carlsen, R. and Willis, D.A. (Eds.): *The proceedings of the 19th International Conference for the Society for Information Technology & Teacher Education*, AACE, Chesapeake, VA, pp.1065–1068.

- Gholson, B. and Craig, S.D. (2006) 'Promoting constructive activities that support vicarious learning during computer-based instruction', *Educational Psychology Review*, Vol. 18, pp.119–139.
- Graesser, A.C. and Person, N. (1994) 'Question asking during tutoring', *American Educational Research Journal*, Vol. 31, pp.104–137.
- Graesser, A.C., Moreno, K., Marineau, J., Adcock, A., Olney, A. and Person, N. (2003) 'AutoTutor improves deep learning of computer literacy: is it the dialog or the talking head?', in Hoppe, U., Verdejo, F. and Kay, J. (Eds.): *Proceedings of Artificial Intelligence in Education*, Finalist for Outstanding Paper Award, IOS Press, Amsterdam, pp.47–54.
- Johnson-Glenberg, M.C. (2005) 'Web-based training of metacognitive strategies for text comprehension: focus on poor comprehenders', *Reading and Writing*, Vol. 18, pp.755–786.
- Lee, J., Dineen, F. and McKendree, J. (1998) 'Supporting student discussions: it isn't just talk', *Education and Information Technologies*, Vol. 3, pp.17–29.
- Mayer, R.E. (2001) *Multimedia Learning*, Cambridge University Press, New York.
- Mayer, R.E. (2005) *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press, New York.
- Moreno, R. and Mayer, R.E. (1999) 'Cognitive principals of multimedia learning: the role of modality and contiguity', *Journal of Educational Psychology*, Vol. 91, pp.358–368.
- Paivio, A. (1986) *Mental Representations: A Dual-coding Approach*, Oxford University Press, New York.
- Penny, C.G. (1989) 'Modality effects and the structure of short-term verbal memory', *Memory and Cognition*, Vol. 17, No. 4, pp.398–422.
- Rawson, K.A., Dunlosky, J. and Theide, K.W. (2000) 'The rereading effect: metacomprehension accuracy improves across reading trials', *Memory and Cognition*, Vol. 28, pp.1004–1010.
- Rosenthal, R.L. and Zimmerman, B.J. (1978) *Social Learning and Cognition*, Academic Press, New York.
- Sadoski, M. and Paivio, A. (2001) *Imagery and Text: A Dual Coding Theory of Reading and Writing*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Sullins, J. and Azevedo, R. (2007) 'Developmental differences in self-regulated learning and question asking during learning with hypermedia', in McNamara, D.S. and Trafton, J.G. (Eds.): *Proceedings of the 29th Annual Cognitive Science Society*, August, Cognitive Science Society, Austin, TX, p.1862.
- Sullins, J., Azevedo, R., Trousdale, A. and Scott, J. (2007) 'The influence of self-regulation on student-generated questions during hypermedia learning', in Luckin, R., Koedinger, K. and Greer, J. (Eds.): *Artificial Intelligence in Education: Building Technology Rich Learning Contexts that Work*, IOS Press, Amsterdam, The Netherlands, pp.187–194.

## **Appendix**

### *Example computer literacy test questions*

Example multiple choice questions

- 1 Why is RAM considered to be volatile memory?
  - a *Its contents can be changed*
  - b Its contents are easily lost
  - c The CPU can read its contents
  - d It can send error messages to computer users
  
- 2 What does the CPU use RAM for when executing an application programme?
  - a Long term storage
  - b To bypass the operating system
  - c *Short term storage*
  - d To hold the instructions necessary to reboot
  
- 3 What will happen if the CPU fails to perform mathematics?
  - a It will not communicate with peripherals
  - b It will only process some of the data
  - c *It will not process any of the data*
  - d It will crash the operating system
  
- 4 How does the CPU obtain information from the keyboard?
  - a Information goes through RAM to the CPU
  - b Information goes from the monitor to RAM to the CPU
  - c Information goes directly from the monitor to the CPU
  - d *Information goes directly from the keyboard to the CPU*