Journal of Educational Multimedia and Hypermedia (2014) **23**(4), 385-400

Are Educational Shows Teaching Our Children to Become Life-Long Learners?

JEREMIAH SULLINS
Harding University, USA
jsullins@umobile.edu

TIFFANY HOWARD
University of Mobile, USA
thoward1126@gmail.com

KIMBERLY GOZA
Middle Tennessee State University, USA
kgoza14@gmail.com

The purpose of this study was to investigate various textual characteristics of popular children television shows. More specifically, researchers examined both the quantity and quality of question asked (i.e., question training). Furthermore, several readability components among the different shows (e.g., narrativity, syntactic simplicity, referential cohesion, deep cohesion and word concreteness) were explored. Results revealed that not all shows are equal when it comes to the amount of question training that is available during an episode. It was also discovered that most of the shows do share similar, although counterintuitive, patterns regarding the different readability components.

INTRODUCTION

Are "educational" shows providing training on the learning strategies that children need in order to become life-long learners? Lapierre, Piotrowski, and Linebarger (2012) recently revealed that the average child age eight months to eight years is exposed to nearly four hours of background television a day. This finding is even more prominent among infants and toddlers, who can watch as much as six hours of background television per day. Other research has also revealed startling statistics when it comes to the amount of television that the average adolescent views by the time they graduate from high school. More specifically, research shows that by the time the average student finishes high school they have watched approximately 15,000 hours of television (compared to 11,000 hours in the classroom).

Additionally, research has shown that for children aged 3-5, which is the target audience for the majority of educational television programs, watch on average 2 or more hours of television per day. It has also been revealed that 59% of children younger than two years old watch on average 1.3 hours of television every day (Rideout, Vandewater, & Wartella, 2003; Huston, Wright, Marquis, & Green, 1999; Christakis, Ebel, Rivara, & Zimmerman, 2004). Regardless of which statistic is correct related to the number of hours of television watched by children of various ages, the important question is what kind of cognitive impact is this television watching having on the cognitive faculties of children.

There are mixed results regarding the exact effects of television watching on the cognitive faculties of children. However, much of the available research does suggest that watching an excessive amount of television can have detrimental effects on cognitive growth. More specifically, Gadberry (1980) revealed that a six week reduction in television viewing among six year olds showed significant improvements in their performance IQ and attention time on cognitive tasks. Furthermore, Lillard and Peterson (2011) were interested in the effects of fast paced television on children's executive functions (e.g., self regulation and working memory). The researchers had sixty 4 year olds watch either a faced past television show for nine minutes or an educational show for nine minutes. Following the completion of the television viewing, the researchers tested the children's executive functioning through a variety of tasks such as delay of gratification and the Tower of Hanoi. Results revealed that the children who watched the faced paced television shows performed significantly worse on tasks of executive function compared to the children that watched the educational television shows.

Lastly, various research has discovered that watching an excessive amount of television is related to decreased performance in areas such as reflection and interpretation in addition to stunting the growth of imagination and creativity. Because of these results, the American Academy of Pediatrics has issued guidelines urging parents to avoid any television viewing before age two.

There is little doubt that the potential benefits and drawbacks of childrens' television shows are many. It is beyond the scope of this article to discuss all the specific factors that can increase or decrease certain cognitive abilities as a result of watching television. Of particular interest in the current study are the effects of question asking. The focus of this paper is to explore the research question: How often and to what extent is educational television teaching effective "question training" in children?

QUESTION GENERATION

Question generation has received a great deal of attention in recent years from researchers in the fields of computer science (Heilman & Smith, 2010), psychology (Graesser, Ozuru, & Sullins, 2009; Rus & Graesser, 2009; Sullins & McNamara 2009) and education. Question generation is believed to play a crucial role in a variety of cognitive faculties, including comprehension (Collins, Brown, & Larkin, 1980; Graesser, Singer, & Trabasso, 1994) and reasoning (Graesser, Baggett, & Williams, 1996; Sternberg, 1987). Asking good questions has been shown to lead to improved memory and comprehension of material in school children and adult populations (Rosenshine, Meister, & Chapman, 1996). Available research suggests that learning how to ask good questions should be taught at an early age but all ages benefit from question generation training (Wisher & Graesser, 2007).

Sadly, it is well documented that the ideal scenario of a curious question asker does not match reality. Students are unspectacular at monitoring their own knowledge deficits and their question generation is both infrequent and unsophisticated (Baker, 1979; Dillon, 1988; Graesser & Person, 1994; Van der Meij, 1988). Graesser and Person (1994) reported that an individual student asks approximately 1 question in 7 hours of class time (around 1 question per day). Most of these questions are not good questions, so the quality is also disappointing.

It is well documented that students and adults have trouble generating questions (Dillon, 1988; Graesser & Person, 1994; Wisher & Graesser, 2007). Of the questions that are generated, the majority are shallow questions rather than questions that require deep reasoning. A deep reasoning question is one which integrates content and that fosters understanding of the components and mechanisms being covered (Craig, Vanlehn, & Chi, 2009). Deep reasoning questions are questions that typically invite lengthier answers (usually around a paragraph in length) and often start with words such as *why, how,* or *what-if* (Graesser, Ozuru, & Sullins, 2009). These questions are aligned with the higher levels of Bloom's taxonomy (1956) and the long-answer question categories in the question taxonomy proposed by Graesser and Person (1994).

In order to illustrate the difference between shallow reasoning questions and deep reasoning questions, consider an example of each. An example of a shallow reasoning question, according to Graesser and Person (1994), would be "Does the CPU use RAM when running an application?". The reason for categorizing this type of question as "shallow" is because it does not require substantial thought on the student's part; indeed, the student could answer it by simply guessing yes or no. In contrast, a deep reasoning question would be "How does the CPU use RAM when running an application?". The reason for categorizing this question as "deep" is because the student must use the knowledge known about computers to articulate the causal mechanisms that relate two components in the operating system. They not only need to generate a nontrivial amount of content, but must be able to reason about complex causal mechanisms.

Graesser and Person (1994) estimated that a typical student asks only .11 questions per hour in a traditional setting, such as a classroom. There are several possible explanations as to why students do not ask many questions. These include the lack of prior domain knowledge, high social editing, and insufficient training/modeling. The first explanation for the lack of student questions might be due to insufficient prior knowledge so they are incapable of monitoring the fidelity of knowledge. For example, Miyaki and Norman (1979) posit that students need a large amount of knowledge to detect when they do not understand something. Because of this, students simply do not know that they do not understand and therefore do not ask questions. The second possible explanation for a low amount of student questioning is due to social editing. Students may not ask questions because they are afraid of looking ignorant in front of their peers and losing social status. The third reason for a low number of student questions has to do with the training they receive. Graesser and Person (1994) point out that 96% of questions that occur in the classroom come from the teacher and most of the questions are shallow. Therefore, students in a typical classroom are not provided with examples of good deep-reasoning questions from the teachers. And of course, given the above statistics on student question asking, students rarely observe other students asking questions. Good student role models are essentially absent.

Based on these findings it is possible that students need to be receiving question-asking training outside of the classroom. One area that is ripe for exploration is the type of question training that is provided to children through educational programming.

VICARIOUS LEARNING

According to constructivist epistemology, learners actively create meaning and knowledge by interacting with people and other objects. Rather than simply delivering information, learning environments should stimulate the learner to actively construct knowledge. This traditional view of constructivism has focused on keeping the learner physically active usually by interactivity. However, Mayer (2002) has made the claim that the learner does not have to be physically active in order for constructivist learning to occur. Under this view, the learner must only be cognitively active during knowledge acquisition.

In this situation, the physically passive learner would engage in a form of vicarious learning. For our purposes, vicarious learning is defined as learning in multimedia environments under conditions in which the user is passive, in that they do not physically interact in any way with the source of the information. Historically, the term vicarious learning was frequently used synonymously with observational learning, social learning, or modeling (e.g., Bandura, 1962; Rosenthal & Zimmerman, 1978). According to this perspective, by simply observing activities carried out by others, learners can master those activities without overt practice or direct incentives (Rosenthal & Zimmerman, 1978).

Current trends in educational technology such as computer-based courses (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995; Mayer, 2009), and distance learning (Barker & Dickson, 1996; Bourdeau & Bates, 1997; Moore & Kearsley, 1996) have created situations in which learners are more and more likely to find themselves trying to gain knowledge as observers (Cox, McKendree, Tobin, Lee, & Mayers, 1999), rather than active participants. Because of these technologies further empirical understanding of the conditions that promote learning among relatively isolated observers is required.

Available research has compared student learning gains in the context of vicarious learning environments versus interactive learning environments. For example, Craig, Sullins, Witherspoon and Gholson (2006) conducted two experiments in order to compare student learning gains between interacting and observing. In Experiment 1 students were randomly assigned to one of five different conditions (one interactive and four vicarious). Students in the interactive condition interacted with an intelligent tutoring system called AutoTutor. The learners in this condition used a dialogue box and a keyboard to respond to AutoTutor's spoken questions, assertions, hints, prompts, pumps, back-channel feedback and gestures. The video and audio of each interactive session was recorded. In one vicarious learning condition, each recorded interactive session was presented to a yoked participant who simply watched and listened to it. This condition was known as the yoked-vicarious. A second vicarious condition was simply a monologue that contained the same information as in the interactive and yoked vicarious condition. The information was presented using the same voice engine and agent. This condition was known as monologue-vicarious. In a third vicarious condition, half of the "main points" were preceded by a deep level reasoning question. In the context of AutoTutor, these "main points" are known as ideal answers and expectations. In this vicarious condition, only the ideal answers were preceded by deep level reasoning questions. This condition was known as half-question vicarious. A fourth and final vicarious condition included deep level reasoning questions that preceded every ideal answer and expectation in the monologue. This condition was known as full-questions vicarious. Results revealed that learners in the full-questions vicarious condition significantly outperformed learners in each of the other four conditions.

In Experiment 2, participants were randomly assigned to one of four different conditions: interactive, yoked-vicarious, full-questions vicarious presented as a monologue, and full-questions vicarious presented as a dialogue. As in Experiment 1, participants in the interactive condition directly interacted with AutoTutor on 12 topics concerned with computer literacy. The video and audio of the interactive condition was recorded and showed to participants in the yoked-vicarious condition. The full-questions vicarious presented at monologue condition contained deep level reasoning questions before every ideal answer and expectation. In this condition, the same agent and voice engine used in the interactive and yoked vicarious condition spoke the question and content. In the full-questions vicarious presented as a dialogue condition every ideal answer and expectation were preceding by a deep level reasoning question. However, the deep level reasoning question was asked by a separate distinct voice. Only the agent from the previous three conditions was present on the screen. Results revealed that both vicarious deep level reasoning question conditions significantly outperformed both the yoked-vicarious and interactive conditions.

Although according to constructivist epistemology, interaction needs to occur in order for deep learning to take place, the results from these two experiments suggest that this way of thinking may not be completely accurate. More specifically, learners can achieve deep learning in a vicarious learning environment if designed properly (i.e., if it contains deep level reasoning questions). Related to the current study, this is as a crucial point that needs to be emphasized. By their very nature, children's educational shows are vicarious learning environments. Unfortunately, little is currently known regarding the content structure of children's educational shows. In other words, what specific learning strategies are being included in these "educational" shows? Furthermore, although beyond the scope of the current paper, if these educational shows (i.e., vicarious learning environments) contain deep level reasoning questions how much learning is acquired by observers compared to other interactive learning environments (e.g., school or tutoring)?

Based on the previous two sections (question asking and vicarious learning) research shows that question asking can be a beneficial learning strategy. We also know that unfortunately, learners have difficulty generating good questions. One potential reason could be that students are not receiving good question training inside a traditional classroom setting. This presents an opportunity to explore other ways in which students could receive the training they need outside of the classroom. One readily available opportunity is through the use of educational television programs (vicarious learning environments). However, in order for vicarious learning environments to be effective, there needs to be an inclusion of deep level reasoning questions. With the addition of deep level reasoning questions to the vicarious learning session, does this in itself help train observers to become better question askers?

QUESTION TRAINING IN THE CONTEXT OF A VICARIOUS LEARNING ENVIRONMENT

Within the context of a vicarious learning environment, is it possible to train learners to become better question askers in a relatively short period of time (e.g., 30 minutes) by simply including deep level reasoning questions into the content? There is available research suggesting that this is a possibility. More specifically, Craig, Gholson, Ventura, Graesser, and the Tutoring Research Group (2000) used a vicarious learning environment in an attempt to not only increase domain knowledge but to also increase the quality of questions asked during the learning session. Learners were randomly assigned to one of two different conditions: monologue or dialogue.

In both conditions, each subtopic was introduced by a brief information delivery presented by a virtual tutor. In the monologue condition, the virtual tutor then asked one question, and the virtual tutor followed this with a monologue presentation of the tutorial content for that subtopic. In the dialogue condition, after each information delivery a virtual tutee asked a series of questions, ranging from five to 14 across the eight subtopics for a total of 66. The exact words, phrases, and sentences used by the virtual tutor in response to the virtual tutee's questions were identical in the dialogue and monologue conditions in each content domain (Craig et al., 2000).

Following the learning session, participants were allowed to ask the experimenter any follow up questions they wanted regarding the topics covered. Additionally, participants took a retention test regarding the material they had covered during the learning session. Results revealed that students in the dialogue condition generated more questions than the participants in the monologue condition which could have in turn caused the significant difference in retention scores (dialogue significantly outperformed monologue).

The results from this study suggest that it is possible to teach learners how to become better question askers in a relatively short amount of time by simply observing good questions being asked. Consistent with previous findings, the participants that viewed the vicarious presentation that contain more deep level reasoning questions (dialogue condition) scored significantly higher on tests of retention than did the participants that viewed fewer deep level reasoning questions (monologue condition). The three previous sections provide justification for the exploration in current study. More specifically 1) are children's television shows designed in a way to maximize learning and 2) are children's television shows being designed in a way so that learners are receiving good question asking training?

CURRENT STUDY

The current study explored various textual characteristics of popular children television shows. More specifically, we examined both the quantity and quality of questions asked (i.e., question training). Furthermore, we explored several readability components among the different shows (e.g., narrativity, syntactic simplicity, referential cohesion, deep cohesion and word concreteness).

Procedure

Researchers first decided what shows were to be used for the analysis. Shows were selected that aired with regularity on popular television channels that are marketed towards children (e.g., Sprout and Disney Junior). Researchers then gathered all available transcripts from the web (livedash.com). In order to have time equivalency, researchers made sure that all transcripts for a specific show fell into the 230-290 minute range. Once all transcripts had been collected and checked for time equivalency, three independent researchers went into all transcripts and marked where a question had been asked during the episode. Researchers collected approximately 4,200 questions from 32 hours of shows. All questions were then analyzed as "deep" or "shallow" based on the Graesser and Person (1994) question taxonomy.

Results

Results revealed a healthy distribution of both deep-level reasoning and shallow-level reasoning questions. However, not all shows are created equal when it comes to the availability of question training techniques. More specifically, it was discovered that the three worst shows for teaching children good question asking skills were Handy Mandy, Cyberchase, and Caillou. Handy Mandy asked a total of 649 questions in approximately four hours of programming. However, 81% (525) of those questions were shallow-level reasoning questions whereas only 19% (124) were considered deep-level reasoning questions. This breaks down to about 71 questions during a 30 minute show (which is the average length of a children's show). Of those 71 questions 57 are shallow-level reasoning whereas 14 are deep-level reasoning on average. Cyberchase asked a total of 588 questions in approximately four hours of programming. However, 80% (473) of those questions were shallow-level reasoning questions whereas only 20% (115) were considered deep-level reasoning questions. This breaks down to about 69 questions during a 30 minute show. Of those 69 questions, 55 are shallow-level reasoning whereas 14 are deep-level reasoning on average. Lastly, Caillou asked a total of 494 questions in approximately four hours of programming. However, 86% (424) of those questions were shallow-level reasoning questions whereas only 14% (70) were considered deep-level reasoning questions. This breaks down to about 68 questions during a 30 minute show. Of those 68 questions 58 were shallow-level reasoning whereas 10 would be deep-level reasoning on average.

The three best shows for teaching children good question asking skills were Sid the Science Kid, Mickey Mouse Clubhouse, and Jungle Junction. Sid the Science Kid asked a total 505 questions in approximately four hours of programming. The good news is that 60% (302) of those questions were considered to be deep-level reasoning questions. Only 40% (203) were shallow-level reasoning questions. This breaks down to about 59 questions during a 30 minute show (which is the average length of a children's show). Of those 59 questions, 35 would be deep-level reasoning questions whereas 24 would be shallow level-reasoning questions on average. Mickey Mouse Clubhouse asked a total of 662 questions in approximately four hours of programming. The good news is that 50% (328) of those questions were considered to be deep-level reasoning questions. Only 50% (328) of those questions were considered to be shallow-level reasoning questions. This breaks down to about 62 questions during a 30 minute show. Of those 62 questions, 31 would be deep-level reasoning questions whereas 31 would be shallow-level reasoning questions on average. Lastly, Jungle Junction asked a total of 795 questions in approximately four hours of programming. The good news is that 62% (496) of those questions were considered to be deep-level reasoning questions. Only 38% (299) were shallow-level reasoning questions. This breaks down to about 101 questions during a 30 minute show. Of those 101 questions, 63 would be deep-level reasoning questions whereas 38 would be shallow-level reasoning questions on average. For a full list of all the shows, please see Table 1.

Table 1

Frequency of Questions among Educational Programs.

Title of Show	Raw Number of Questions	Shallow	Deep	Per Episode
Handy Manny	649	525 (81%)	124 (19%)	71 (57/14)
Sesame Street	701	508 (72%)	193 (28%)	91 (66/25)
Sid the Science Kid	505	302 (60%)	203 (40%)	59 (35/24)
Mickey Mouse	662	328 (50%)	334 (50%)	62 (31/31)
Cyberchase	588	473 (80%)	115 (20%)	69 (55/14)
Caillou	494	424 (86%)	70 (14%)	68 (58/10)

In addition to exploring the type of question asking skills that were being provided during these eight popular children shows, as a secondary analysis, we were also interested in the textual characteristics of each episode. The five textual characteristics that we were interested were: Narrativity, Syntactic Simplicity, Word Concreteness, Referential Cohesion, and Deep Cohesion. A *narrative* text is a text that tells a story with characters, events, places and things that are familiar to the reader. *Syntactic simplicity* refers to sentences with few words in addition to simple, familiar structure that is easy to process and understand. *Word concreteness* relates to words that evoke mental images and are more meaningful to the reader compared to abstract words. *Referential cohesion* refers to texts that contain words and ideas that overlap across sentences and the entire text, forming threads that connect the textbase together for the reader. Finally, *deep cohesion* relates to causal, intentional and temporal connectives that help the reader form a more coherent and deeper understanding of the text. As can be seen in Figure 1, there does seem to be a somewhat consistent pattern across most shows. More specifically, this analysis reveals that most shows tend to be high in narrativity and syntactic simplicity and relatively low in deep cohesion, referential cohesion, and word concreteness. The exception is the show Sid the Science Kid which shows higher levels of word concreteness and deep cohesion.

DISCUSSION

As the results revealed, not all shows are created equal when it comes to promoting learning and teaching children effective question training skills. More specifically, Sid the Science Kid, Mickey Mouse Clubhouse and Jungle Junction provide more deep-level reasoning question training than do the other shows that were analyzed for this study. Additionally, the shows that contained the least amount of deep-level reasoning question training were Handy Manny, Cyberchase and Caillou.

As previously stated, the lack of question asking by both children and adults is near epidemic proportion. Learners are not asking questions, and the questions they do ask are unspectacular. It is beyond the scope of this paper to go into every hypothesis as to why learners are not asking questions, but according to available research (Graesser & Person, 1994), one possible reason could be due to the lack of question training students are receiving. For example, as mentioned earlier, the majority of questions that are posed by the teacher in the classroom are classified as shallow-level reasoning questions.

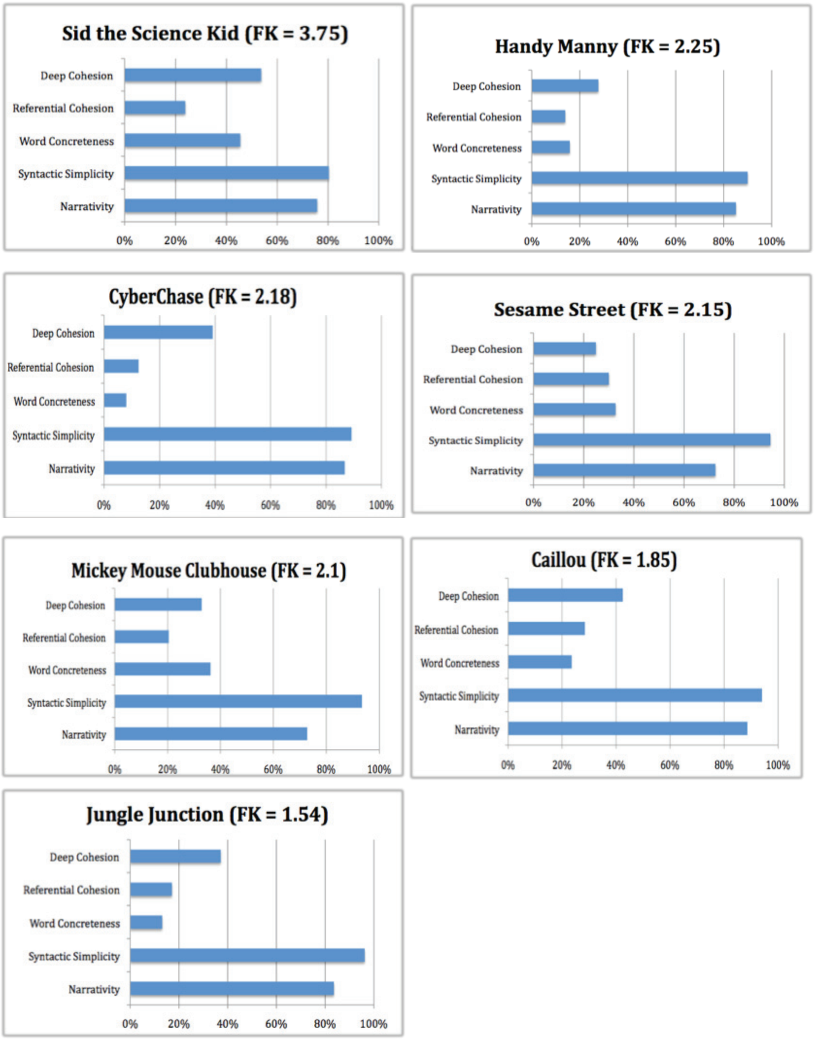

Figure 1. Coh-Metrix output for educational television.

This suggests that there is a significant need for good question training skills outside of the classroom. One of the most readily available ways for children to learn additional skills outside of the classroom is through the use of educational programming. As previously mentioned, educational shows are one specific type of vicarious learning environment which research has shown can actually be more effective than actual interaction with the to be learned material. One caveat of these vicarious learning environments however is that just observation alone without the proper instructional design will not lead to learning. More specifically, vicarious learning environments need to be embedded with deep level reasoning questions in order to be most effective. As these results suggest, some educational shows contain more deep level reasoning questions than others which in turn could not only lead to more retention on the material being presented but also with the inclusion of more deep level reasoning questions, could actually train learners to become better questions askers. As previously mentioned, question asking is a beneficial learning strategy that leads students to become more effective learners.

Overall, the results from the current study suggest that there is the opportunity for children to receive proper instruction through the use of these vicarious multimedia learning environments in addition to receiving question training outside of the classroom. These findings warrant further research due to the fact that the current study did not examine a direct link between children that watch the "high quality" question training shows and question asking ability.

As a secondary analysis, we examined specific textual components of each of the children's television shows to see if any commonalities existed. The Coh-Metrix analysis revealed a similar pattern among all of the shows except one (i.e., Sid the Science Kid). More specifically, the results revealed that all children shows are relatively high in Syntactic Simplicity and Narrativity. This seems appropriate given the age range that these shows are appropriate for (which are children ages 6-9 according to Flesch-Kincaid grade estimates). For example, it is expected that a show geared towards kids would be one that tells a story with familiar places, characters, events and things (Narrativity). Additionally, it is expected that shows aimed at children would have sentences with few words along with simple and familiar structure that is easy to process and understand (Syntactic Simplicity).

Results also found that the shows seem to be relatively low in Deep Cohesion, Referential Cohesion, and Word Concreteness. These findings are not as clear cut as with narrativity and syntactic simplicity. Regarding referential cohesion, low cohesion texts are typically more difficult to process because there are fewer connections that tie the ideas together for the reader. However, if the reader has sufficient knowledge, then the required inferences in low cohesion text may benefit comprehension. It would have been expected to find the dialogue from the shows to contain more cohesion so that the children do not have to make inferences. However, the more connectives that are made to increase cohesion, the higher the syntactic difficulty. The question becomes which is the lesser of the two evils?

Regarding deep cohesion, when a text contains many relationships but does not contain connectives, then the reader must infer the relationships between the ideas in the text. If the text is high in cohesion, then those relationships and global cohesion are more explicit. As with referential cohesion, an increase in cohesion would lead to an increase in syntactic difficulty. Again, the question becomes which aspect of the text is more important to the researcher, the syntactic simplicity or the cohesiveness of a text.

Lastly, the relatively low scores of word concreteness across the majority of the texts is another finding that is somewhat counterintuitive. It would be expected that these shows would use words that are more easily visualized (e.g., truck, ball, playhouse, etc…). However these results suggest that the use of concrete words is infrequent. The results show that word concreteness is one of the lowest (if not the lowest) component analyzed on every text. One possible explanation for this relatively low occurrence of concrete words could be due to the nature of the topics discussed in these shows. For example, it could be that the message conveyed in these episodes has to do with abstract concepts such as "friendship" or "sharing".

References

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutor: Lessons learned. *The Journal of the Learning Sciences, 4***,** 167-202.

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior, 11,* 363-374.

Bandura, A. (1962) Social learning through imitation. In M.R. Jones (Ed.), *Nebraska symposium on motivation,* (pp. 211-269), University of Nebraska Press, Lincoln.

Barker, B. O., & Dickson, M. W. (1996). Distance learning technologies in K-12 schools: Past, present, and future practice. *Techtrends, 41***,** 19-22.

Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain*. New York: McKay.

Bourdeau, J., & Bates, A. (1997) Instructional design for distance learning. In R. D. Tennyson, S. Dijsktra, N. Steel, & F. Schott (Eds.) *Instructional design: International perspectives,* (pp. 369-397), Erlbaum, Hillsdale, NJ.

Christakis D. A., Ebel B. E., Rivara F. P., & Zimmerman F. J. (2004) Television, video, and computer game usage in children under 11 years of age. *Journal of Pediatrics, 145*(5), 652-656.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.) *Theoretical issues in reading comprehension* (pp. 385-407). Hillsdale, NJ: Erlbaum.

Cox, R., McKendree, J., Tobin, R., Lee, J., & Mayes, T. (1999). Vicarious learning from dialogue and discourse. *Instructional Science, 27* , 431–458.

Craig, S. D., Gholson B., Ventura, M., Graesser, A. C., & the Tutoring Research Group. (2000). Overhearing dialogues and monologues in virtual tutoring sessions: Effects on questioning and vicarious learning. *International Journal of Artificial Intelligence in Education (Special Issue: Analyzing Educational Dialogue Interaction), 11*, 242–253.

Craig, S. D., Sullins, J., Witherspoon, A. & Gholson, B. (2006). Deep-level reasoning questions effect: The role of dialog and deep-level reasoning questions during vicarious learning. *Cognition and Instruction, 24(4),* 563-589.

Craig, S. D., VanLehn, K. & Chi, M. T. H. (2009). Improving classroom learning by collaboratively observing human tutoring videos while problem solving. *Journal of Educational Psychology*, *101*, 779-789.

Dillon, J. T. (1988). *Questioning and teaching: A manual of practice*. New York: Teachers College Press.

Gadberry, S. (1980). Effects of restricting first graders' TV-viewing on leisure time use, IQ change, and cognitive style. *Journal of Applied Developmental Psychology*, *29*(*1*), *45*–57.

Graesser, A. C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology, 10*, 17-32.

Graesser, A., Ozuru, Y., & Sullins, J. (2009). What is a good question? In M. G. McKeown & L. Kucan (Eds.), *Threads of coherence in research on the development of reading ability* (pp. 112-141). NY: Guilford.

Graesser, A. C, & Person, N. (1994). Question asking during tutoring. *American Educational Research Journal, 31*, 104-137.

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 3*, 371-395.

Heilman, M., & Smith, N. A. (2010). Extracting Simplified Statements for Factual Question Generation. In K. E. Boyer & P. Piwek (Eds.), *QG2010: The Third Workshop on Question Generation* (11-20). *Springer:Berlin / Heidelberg*.

Huston, A. C., Wright, J. C., Marquis, J., & Green, S. (1999). How young children spend their time: Television and other activities. *Developmental Psychology, 35*, 912–925.

Lapierre M, Piotrowski M, & Linebarger D. (2012, May). *Background television in the homes of American children.* Paper presented at the International Communication Association's Annual Conference, Phoenix, AZ, USA.

Lillard, A. S., & Peterson, J. (2011). The immediate impact of different types of television on young children's executive function. *Pediatrics, 128*, 644–649. doi:10.1542/peds.2010-1919

Mayer, R. E. (2002). Using illustrations to promote constructivist learning from science text. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The Psychology of Science Text Comprehension* (pp. 333–356). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, R.E. (2009) *Multimedia learning*. Cambridge University press, Cambridge, UK.

Moore, M. G., & Kearsley, G. (1996). *Distance education: A systems view*. Wadsworth:Albany, NY.

Miyake, N., & Norman, D.A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18,* 357-364.

Rideout, V.J., Vandewater, E.A., & Wartella, E.A. (2003). *Zero to six: Electronic media in the lives of infants, toddlers, and preschoolers.* Menlo Park, CA: Kaiser Family Foundation.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66,* 181-221.

Rosenthal, R.L., & Zimmerman, B.J. (1978). *Social leaning and cognition*. Academic Press, New York.

Rus, V., & Graesser, A. C. (Eds.). (2009). *The question generation shared task and evaluation challenge.* Available at: http://www.questiongeneration.org/.

Sullins, J., & McNamara, D.S. (2009). *iSTART question training module: Training students efficient questioning skills.* Presented at the 2nd Workshop on Question Generation, Brighton, United Kingdom.

Sternberg, R. J. (1987). Questioning and intelligence. *Question Exchange, 1*, 11-13.

Wisher, R. A., & Graesser, A. C. (2007). Question asking in advanced distributed learning environments. In S. M. Fiore & E. Salas (Eds.), *Toward a science of distributed learning and training* (pp. 209–234). Washington, DC: American Psychological Association.

Van der Meij, H. (1988). Constraints on question asking in classrooms. *Journal of Educational Psychology, 80*, 401-405.