ANALYTICAL REASONING SKILLS: IMPROVING PERFORMANCE WITH NOTATIONS

Michelle Eskritt & Carol Arthurs

In this study we explored the effect of notations that undergraduates produce on their reasoning abilities. Participants solved four analytical reasoning problems, making notes to solve two problems and solving two without notations. Most participants produced notes when given the opportunity. The production of higher quality notes did not aid performance when we compared the write and no write conditions and those producing poorer quality notes performed even worse in the write condition. The results suggest that students do not necessarily have adequate note-taking skills and that they may need instruction in the production and use of external representations.

Keywords: problem-solving, notational quality, note-taking, external representations

Dans cette étude, les auteures ont exploré l’effet des notes prises par des étudiants de premier cycle sur leur capacité de raisonnement. Les participants ont résolu quatre problèmes de raisonnement analytique, prenant des notes en vue de résoudre deux problèmes sur quatre. La plupart des participants ont pris lorsqu’ils en avaient l’occasion. La prise de meilleures notes n’a pas contribué à l’obtention de meilleurs résultats si l’on compare les contextes avec ou sans notes ; les personnes qui ont pris de moins bonnes notes ont obtenu un rendement encore plus mauvais dans le contexte avec notes. Les résultats semblent indiquer que les étudiants ne savent pas nécessairement comment prendre de bonnes notes et qu’il faudrait peut-être leur montrer comment prendre des notes et s’en servir.

Mots clés : résolution de problèmes, prise de notes, qualité des notes
As writing systems developed in Western culture, they were accompanied by large social changes including more complex agricultural and trade systems, political and economic advances, and mass education and production (Scribner & Cole, 1981). Although such changes have played an important role in modern society, scholars have long debated the impact of literacy on both the individual and society and its relationship with cognitive processes. For example, Aristotle felt that external representation of information would not affect thought because it was only a transcription of speech (Olson, 1994), a view still held by some theorists today (e.g., Carruthers, 1990). This perspective assumes that literacy does not impact individual cognition, although it can have an influence at the societal level. In contrast, other theorists have suggested that the act of writing may alter thought processes (Harris, 1989). In this view, literacy paved the way for the development of logical, analytical, and scientific thinking (Donald, 1991; Goody & Watt, 1968; Havelock, 1963; McLuhan, 1962; Ong, 1982). It can have an impact at the individual level.

The effect of print exposure on cognitive processes is one area that examines the impact of literacy on the individual. In the field of the cognitive psychology of reading, Stanovich and Cunningham (1992) have attempted to pinpoint the cognitive processes that support effective reading performance. In contrast, there has been little focus on the influence of exposure to print on cognitive processes. Avid readers tend to differ from nonreaders in a variety of ways, including their cognitive skills, behavioural habits, and background variables (Stanovich, 1993). Stanovich, West, and Harrison (1995) also found that print exposure was directly related to measures of crystallized intelligence (i.e., vocabulary, general knowledge), but they did not find a direct influence of print exposure on fluid abilities (i.e., reasoning, working memory).

The debate about the impact of literacy remains, partly because of the way it has been defined. Most often literacy is thought of as the ability to read and write. However, this definition limits the interpretation of literacy and excludes other important functions. A more comprehensive definition of literacy includes the assumptions, expectations, and attitudes that accompany the basic skills of reading and writing (Illich, 1991; McLane & McNamee, 1990; Olson, 1994).
Limiting the definition of literacy to the skills of reading and writing only confirms the implications of literacy (Olson, 1994) and the research that is conducted. Although literacy is often viewed as competence in alphabetic reading (Donald, 1991), a number of other symbol systems exist. A more comprehensive definition of literacy is needed which should include the ability to use external representations in general and not to be restricted to alphanumeric symbols. External representations, such as diagrams, graphs, or pictures are commonly used in problem solving, decision making, or reasoning tasks. It is theorized that external representations are an integral part of these tasks (Zhang, 1997).

However, until recently no distinction has been made between internal and external representations, which work together in many cognitive tasks (Zhang, 1997). This distinction between internal and external representations allows for the study of external representations as a field on their own. The present study examines how undergraduates use external representations to aid cognitive processing, specifically performance on a reasoning task.

Theorists from a number of viewpoints, including historical, evolutionary, and cross-cultural perspectives, have noted that a main property of external representations is to serve as a memory aid. Further research has indicated that external memory aids are used more often and they are perceived to be more reliable than internal memory aids (Harris, 1982; Intons-Peterson & Fournier, 1986). Given this property of notations, several studies have examined the impact of external symbols on memory (for review see Kiewra, 1985, 1989).

Eskritt, Lee, and Donald (2001) conducted a series of experiments examining the impact of producing notations on the memory performance. Participants who played a memory card game were allowed to make notes to help in some conditions but not in others. When given the opportunity to make notes, over three-quarters of the participants chose to do so, and this note production improved memory performance. Furthermore, they found that the unexpected removal of notations led to a decrease in performance, suggesting that the spontaneous production of notes served as external memory storage. However, further experimentation found that the identity of cards was still retained in memory, though notations were used to store location
information for cards (Eskritt et al., 2001). These findings suggest that the relationship between memory and literacy is more complex than typically thought. These findings warrant further study.

Theorists also believe that notations may influence the cognitive process of reasoning (Goody & Watt, 1968; Ong, 1982); however, there has been little empirical research examining the effects of notations on problem solving and reasoning. Goody and Watt (1968) have argued that formal logic and reasoning began in Ancient Greece as a result of the development of a literate society. In fact, Levy-Bruhl (1923) made the assumption that oral societies were “prelogical” in their thinking. A more modern standpoint is that oral cultures use similar logic principles as literate cultures, with the difference being how societies generalize information (Luria, 1976). For example, in his research with the illiterate adult residents of an isolated area of Russia, Luria (1976) found that schooling could have a dramatic effect on the responses given to syllogisms. Luria argued that the introduction of literacy led to an increase in abstract, theoretical, and logical thinking compared to residents who had not received schooling.

Zhang has examined how different types of external representations can affect a literate individual’s problem solving, using variations of two games, Tic-Tac-Toe (Zhang, 1997) and the Tower of Hanoi (Zhang & Norman, 1994). He found that the type of external representation used to characterize a problem could influence participants’ perceived difficulty of the problem as well as their performance on the problem. The form of the external representation can influence what information participants perceive more readily, and what processes they use in solving problems (Zhang, 1997). Furthermore, from the problem solver’s point of view, tasks are completely different with and without external representations (Zhang & Norman, 1994). People often use external representations as support when solving problems (Cary & Carlson, 2001). External representations can be used to maintain and manage information necessary in solving a problem and in clarifying the steps needed in solving the problem (Cary & Carlson, 1999; Cox, 1999). However, Cary and Carlson (2001) have suggested that external representations are helpful only if the benefits of using them outweigh the costs. For example, when using external representations, participants
need to attend not only to the task at hand but also to the external representations that they employed as an aid (Carlson, 1997; Miyake & Shah, 1999).

Another issue is participants’ production of their own notations for reasoning tasks. Cox (1999) found that self-constructed notations can be either graphical or linguistic and that different individuals construct a wide range of representations. Furthermore, Cox and Brna (1995) found that the majority of subjects (> 80%) use some form of notations to aid in solving analytical reasoning tasks. However, when studies have examined how participants use self-generated or presented representations to solve problems, they did not compare participants’ performance on problems solved with and without the use of notations (Cox & Brna, 1995; Zhang, 1997). Conditions where participants do not make notes can serve as a baseline for individual performance. Providing participants with the opportunity to use notes and not use notes allows researchers to make comparisons between the two conditions to see if the use of notations actually aids performance.

THE STUDY

Previous research has shown that schooling may have an effect on reasoning skills (Luria, 1976) and the types of external representations provided by an experimenter may affect problem solving (Zhang, 1997; Zhang & Norman, 1994). The present study explored the effect of notations that undergraduate students produce have on their analytical reasoning abilities. We gave participants four analytical reasoning tasks to complete, two of which they had the option to make notes to solve, and the other two where they did not have the opportunity to produce notations. We then divided participants into groups based on the quality of the notations produced and we compared their performance on analytic questions when they had the opportunity to make notes and when they did not.

Specifically, the study examined three questions. We investigated whether undergraduates produced notes to aid in a reasoning task when given the opportunity, and what types of notations they chose to produce. Based on previous research, we expected that the majority of participants would choose to make notations when given the
opportunity (e.g., Cox & Brna, 1995; Eskritt et al, 2001). Given the results of research on the effects of different types of external representations on problem solving (Zhang, 1997; Zhang & Norman, 1994), we knew that participants might not produce notes of similar quality and that this difference in quality might affect the usefulness of the notations. We also examined whether the use of notations improved the performance of participants as measured by the number of correct responses when they make notes compared to when they do not. We predicted that performance would improve when participants make notations to aid in solving the reasoning problems. Finally, we examined participants' exposure to print. Research has found that familiarity of representations is related to reasoning with notes (Cox 1999) and the type of representation used affects the level of difficulty of a problem (Zhang & Norman, 1994). Thus, we predicted that print exposure would be positively correlated with performance on the reasoning problems because a greater familiarity with reading could lead to a greater familiarity with the possible notations that participants could use to solve problems.

METHOD

Participants

Forty-six undergraduate students enrolled in introductory psychology participated. Forty participants were female and six were male. The mean age of participants was 22.4 (SD = 6.7). They received a bonus mark in the course in return for their participation. Participants were treated ethically according to Tri-Council guidelines (Canadian Public Works and Government Services, 1998).

Materials and Procedure

Each participant was tested individually. We presented participants with four analytical reasoning problems in paper format. Problems were adapted from sample Graduate Record Examination (GRE) questions (Brownstein, Weiner, Green & Hilbert, 1990; GRE practice test, 1996). Each problem consisted of a passage involving six items followed by three restrictions. Participants were required to answer two multiple-choice questions based on the passage and the restrictions. (See
Appendix for a sample problem.) We conducted pilot testing to ensure that responders could complete each question without the use of writing. In the pilot study, seven undergraduate students attempted to solve ten different questions. We chose four questions because pilot test students demonstrated that they could complete them without making notes, although the questions were difficult enough to warrant making notes.

There were two conditions under which participants solved problems: the write condition and the no write condition. This factor was within-subject as participants completed both conditions. In the write condition, we allowed participants to make notes. They were told, “You can use the pencil and paper provided to help you complete the problems.” In the no write condition, we required participants to orally answer the questions and to solve the question without the aid of pencil and paper. The order of questions and writing conditions was counterbalanced.

After the completion of the four questions, participants were given the Author Recognition Test (ART) as a measure of their exposure to print. Stanovich (1993) found that the ART was a valid measure of print exposure because scores on the ART correlated positively with other measures of exposure to print such as interview techniques and activity diary methods. The ART is a simple checklist including the names of 43 authors and 43 foils (Stanovich, 1993). We instructed participants to place a mark next to the names on the list they were sure were authors. We calculated participants’ scores by subtracting the number the foils incorrectly chosen from the number of authors correctly identified.

**Scoring**

We evaluated the quality of notations produced by the participants by examining the amount and type of information contained within the notes. Notes were scored on a scale from 0 to 4. A 0 indicated that only an answer was given, whereas 1 indicated that very little was written and it was not obvious how the information in the notes could help solve the problem. For example, some participants simply rewrote the items contained in the question. A score of 2 was awarded when the restrictions were rewritten in a new format or incorporated into a diagram, but no new information was included. A 3 indicated that the
restrictions were rewritten in a new format or incorporated in a diagram so that some new information could be derived. A 4 indicated that not only was new information provided, but also the question was almost or totally solved within the notations. Accuracy of the notations was also examined but none of the participants made any errors in their notations. Two raters rated all the notations and resolved any differences in ratings. Inter-rater reliability was 89.7 per cent.

Figure 1 illustrates the frequency of scores participants received for their notations. Participants were divided into two groups, those who made poor quality notations and those who made good quality notations, based on the mean quality of the notations scores for both questions in the write condition. Those participants whose mean score was less than 1.5 were placed in the poor notational quality group \((n = 17, M \text{ score} = .93, SD = .6)\) while those participants whose mean score was more than 1.5 were labeled the good notational quality group \((n = 21, M \text{ score} = 2.2, SD = .6)\). The poor notational quality group contained those participants whose notes were minimal (i.e., restated the items in the question) and the information contained within the notes was of questionable aid for solving the problem. The good notational quality group contained those participants who produced more substantial notations to aid in solving the problems. See Figure 2 for examples of notations for each group.

RESULTS

Five participants chose not to make notes when given the opportunity (10.9%) and three participants (6.5%) chose to make notes only for one of the two questions in the write condition. We excluded these participants from further analysis, leaving us with 38 participants from whom we used data in the analyses.

Performance Measures

We conducted two ANOVAs to examine performance on the analytical reasoning questions. Preliminary analyses found no difference in the order of the write and no write conditions \((F(1,36) = 1.27, \text{ n.s, for number correct and } F(1,36) = 0.15, \text{ n.s, for time); therefore we collapsed the data across this variable. In the first analysis, we examined the number of} \)
Figure 1 Frequency of notation scores for participants who wrote

A.

B.

Figure 2 Example of: a) a poor quality notation and b) a good quality notation. Poor notations contained information that was of questionable use for solving the problem, while good notations contained more substantial information that could aid in problem solving.
correct responses participants made. The good notational quality group had a mean of 3.2 (SD = .9) questions correct in the write condition and 3.0 (SD = 1.0) questions correct in the no write condition. For the poor notational quality group, the mean was 2.4 (SD = 1.4) questions correct in the write condition and 2.9 (SD = 0.8) in the no write condition. We performed a 2 (condition) X 2 (notational quality group) ANOVA to examine whether the number of correct responses in the write condition differed significantly from the no write condition depending on notational quality. Results indicated a significant interaction between the good and poor notational quality groups and condition on the number of correct responses, $F(1, 36) = 5.03, p < .05, \eta^2 = .12$. An analysis of simple effects revealed a statistically significant difference in the number of correct responses in the write condition between the two groups ($p < .05$). The good notational quality group ($M = 3.1, SD = 1.0$) did significantly better than the poor notational quality group ($M = 2.7, SD = 1.2$). Although those who made poor quality notations tended to have fewer questions correct when given the opportunity to write, this difference was not significant, $p = .07$. However, notational production did not appear to aid the good notational quality group because the number of correct responses for these participants did not differ between conditions (write condition: $M = 3.2, SD = 0.9$; no write condition: $M = 3.0, SD = 1.0$). There was also no statistically significant difference found between groups for the no write condition.

The second analysis was a 2 (write versus no write condition) X 2 (poor versus good notational quality groups) ANOVA with time taken to complete the questions as the dependent variable. We found a significant main effect for condition, $F(1, 36) = 4.71, p < .05, \eta^2 = .12$. Regardless of notational quality, participants took significantly longer to complete the questions in the write condition (good quality: $M = 8.6$ min, $SD = 3.6$; poor quality: $M = 7.6$ min, $SD = 3.4$) compared to the no write condition (good quality: $M = 6.7$ min, $SD = 2.5$; poor quality: $M = 6.7$ min, $SD = 2.3$). No other significant effects were found.

**ART scores**

The mean score of participants on the ART was 11.3 ($SD = 7.2$) with a range of 2 to 31. We computed correlation coefficients between the ART
and quality of notations, number correct in the write condition, number correct in the no write condition, time to complete write condition and time to complete the no write condition. No significant correlations were found, suggesting that exposure to print was not related to the quality of notes produced by participants or their performance on the reasoning task.

**DISCUSSION**

In this study, we explored the effect of participants’ production of different types of notations on their performance on analytical reasoning problems. As expected, the majority of participants made notes when they were given the opportunity. As Cary and Carlson (2001) have noted, creating notes is influenced by the costs and benefits of using them. The act of making notes takes time and effort as reflected by participants taking significantly longer to complete questions in the write condition than in the no write condition. It could be argued that participants in the present study judged the time and effort required in making notes to be outweighed by the benefits of using notes in solving the analytical reasoning problems. Alternatively, participants may have felt obligated to make notes in the write condition because they were provided with note-taking material in contrast to the no write condition where they were not allowed such materials. Nonetheless, previous research examining the more spontaneous production of notes also found that the majority of participants chose to make notes (Cox & Brna, 1995; Eskritt et al., 2001).

We predicted that making notes would aid in participants’ reasoning skills in the write condition. The influence of notations, however, was not straightforward. Participants who made notes could be divided into those who made poor quality or good quality notes. The poor notational quality group was not as accurate in answering the analytical questions when given the chance to make notes compared to good notational quality group. However, this difference was not the result of good notational quality group doing better in the write condition compared to the no write condition. Instead, poor notational quality group tended not to perform as well in the write condition as compared to the no write condition, though this difference was not statistically significant. It is
important to note that no significant differences were found between good and poor notational quality groups in the no write condition implying that performance differences in the write condition were not related to general reasoning abilities. Because there was also no significant difference between these two groups on time taken to complete the questions, length of time cannot be the explanation. These findings suggest that the production of notations may not always be beneficial. Participants’ notational production could even be viewed as detrimental to performance in the present study because participants took significantly more time to complete the problems in the write condition than in the no write condition.

Several possible factors explain why making notes did not appear to aid performance although theorists have argued that external representations can influence reasoning (Cary & Carlson, 1999; Cox, 1999; Zhang, 1997; Zhang & Norman, 1994). The fact that even the good notational quality group did not benefit from producing notes was a particularly surprising finding. Although the good notational quality group had better quality notes compared to the poor notational quality group, their notes could have included more information. The mean score for the quality of notations in the good notational quality group was 2.2 ($SD = .6$) out of a maximum score of 4 and, as is apparent in Figure 1, the mode was a score of 2. For many in the good notational quality group, their notes only re-represented the information contained in the question and did not go beyond. Perhaps students with notes receiving a score of 3 or 4 may have found their notes more beneficial in solving the problems; however, too few students did so to examine this possibility. It has been noted that background knowledge about the use of diagrams is a predictor of success in reasoning with diagrams (Cox, 1999). The adequate representation of the information from the question and constructing an appropriate representation for the specific task are factors that impact reasoning with notes (Cox, 1999). Zhang and Norman (1994) also noted that different representations of problems could have a large impact on perceived problem difficulty and participant performance. The more types of representations a participant is aware of increases the number of possible representations they can construct to solve a problem and increases their chance of picking one
that accurately represents the information necessary to solve the problem in the most efficient way possible (Cox, 1999). Therefore, participants may not have had enough background knowledge about the use of diagrams and external representations to produce more effective notations.

The ART was used as a measure of familiarity with different notational systems and therefore it was predicted that the participants’ level of print exposure would be positively correlated with quality of notes and performance on the task. Unfortunately, the ART was not found to correlate with these measures. The ART is designed to examine participants’ exposure to text because it measures how much reading the participant engages in by their recognition of popular authors. Thus it may not be an effective means of estimating participants’ exposure to other types of external representations such as matrices, which would be much more helpful for the reasoning problems presented in the present study (Schwartz, 1971; Schwartz & Fattaleh, 1972). Therefore, future research needs to use other measures to examine familiarity of notational systems in general and its influence on the production of notes in reasoning tasks.

A number of factors besides background knowledge of external representations may also contribute to the difference in performance between good and poor notational quality groups in the write condition. Because no significant differences occurred between groups in the no write condition, it is unlikely that participants varied on factors that can influence problem solving in general such as verbal ability or working memory. However, variables such as learning strategies and control processes have also been found to influence notational production (Kiewra, 1988). Research is needed to examine how these variables as well as background knowledge may interact to influence participants’ production and use of notations.

The within subjects comparison of performance in the write and no write conditions controls for cognitive variables that may explain individual differences in performance across the groups. The results of the comparison across conditions indicate that notational production, regardless of quality, did not aid reasoning ability. The difficulty of learning to use symbol systems is often underestimated in educational
settings (Lehrer & Schauble, 2002). Students need to be exposed to the basic roles of external representations and they should have exposure to a wide range of external representations that are not limited to a specific area of study (i.e., physics, mathematics). This exposure may lead to an increase in students’ familiarity with representations, allowing them to apply notations more effectively. Differences in the performance of participants in the present study may also be due to differences in their educational experiences. Students might differ in what types of representations they know how to construct and apply. Furthermore, participants may differ in their previous experience using notations to solve problems. Spending adequate time teaching students to construct, interpret, and effectively utilize different forms of external representations is necessary. It should not be assumed that students learn these skills without instruction and practice (De Simone, Schmid, & McEwan, 2001).

ACKNOWLEDGEMENT

This research was supported in part by a Social Sciences and Humanities Research Council of Canada grant to the first author. Correspondence concerning this article should be addressed to Michelle Eskritt, Department of Psychology, Mount St. Vincent University, 166 Bedford Hwy., Halifax, Nova Scotia, B3M 2J6 Email: michelle.eskritt@msvu.ca.

REFERENCES


APPENDIX – SAMPLE PROBLEM

Below is a sample problem given to participants adapted from Brownstein et al. (1990).

A professor offers a six-week course on film genres. Each time the following six genres are covered: adventure films, detective films, fantasy films, horror films, silent films and westerns.

- Westerns are always covered before adventure films.
- Detective films are always covered the week immediately preceding, or the week immediately following adventure films.
- Horror films are covered either the first or the last week.

Question 1: Which of the following is an acceptable schedule?
- a) fantasy films, westerns, adventure films, silent films, detective films, and horror films
- b) horror films, adventure films, detective films, westerns, fantasy films, and silent films
- c) horror films, westerns, fantasy films, adventure films, silent films, and detective films
- d) westerns, fantasy films, silent films, adventure films, detective films, and horror films
- e) adventure films, detective films, silent films, westerns, fantasy films, and horror films

Question 2: If silent films are shown in week 1 and westerns are shown in week 2, what week(s) can adventure films not be shown?
- a) 6
- b) 5, 6
- c) 4, 5, 6
- d) 3
- e) 4, 5