Learning interaction patterns using diagrams varying in level and type of interactivity

Nuno Otero*, Yvonne Rogers§ and Benedict du Boulay±

* INESC-ID, email: nuno.otero@gaips.inesc-id.pt

§ University of Indiana, email: yrogers@indiana.edu

± University of Sussex: B.du-Boulay@sussex.ac.uk

Abstract

An experiment was conducted to investigate the differences between learners when using computer based learning environments (CBLEs) that incorporated different levels of interactivity in diagrams. Four CBLEs were created with combinations of the following two interactivity properties: (a) the possibility to rotate the whole diagram (b) the possibility to move individual elements of the diagram in order to apprehend the relationships between them. We present and discuss the qualitative findings from the study in terms of the learners' interaction patterns and their relevance for the understanding of performance scores. This supports our previous quantitative analysis showing an interaction between cognitive abilities and interactivity. Based on our findings we reflect on the possibilities to inform CBLEs with relevant information regarding learners' cognitive abilities and representational preferences.

Introduction

Interactive graphical representations have been used and promoted to facilitate learning (Cheng, 1999). However, not enough empirical evidence has been accumulated regarding the benefits of generally adding interactivity to graphical representations or the conditions under which beneficial effects emerge (Otero, Rogers, & Du Boulay, 2001; Rogers, 1999). Furthermore, several authors have also pointed out the need to understand better the influence that individual differences have on problem solving and learning with external representations (ERs) (Cox, 1996, 1999; Hegarty & Steinhoff, 1997; Kozhevnikov, Hegarty, & Mayer, 2002, 2002; Otero et al., 2001; Scaife & Rogers, 1996).

This paper focuses on the description and analyses of learners' interaction patterns with CBLEs that implemented diagrams with distinct levels of interactivity to teach a geometry concept - the stereographic projection. This concept involves the visualisation of 3D objects. More specifically, we wanted to find out if specific patterns of interaction were associated with increased learning performance scores, levels of spatial visualisation ability and domain knowledge. The interactivity properties were: (a) the possibility to rotate the whole diagram - whole diagram manipulation and (b) the possibility to move individual elements of the diagram in order to apprehend the relationships between them - manipulation of the diagrams' elements. The "whole diagram manipulation" functionality aimed at facilitating the perception of the objects' 3D properties, reducing the need to apprehend the 3D form through the decoding of perspective cues. The "manipulation of the diagrams' elements" functionality enables the learner to manipulate elements of a graphical representation allowing the simulation of alternatives. The learner can read off directly from the diagram how altering a certain parameter affects other parameters.

Previously reported analyses found a complex interaction between levels of interactivity, levels of spatial visualisation ability and performance scores (Otero et al., 2001). This paper discusses the interaction patterns found and possible ways to relate these with other information concerning learners' cognitive abilities and representational preferences. Three questions come to mind: (i) How to gain knowledge about the learners' cognitive abilities when they are interacting with a CBLE? (ii) Can the knowledge acquired regarding learners' cognitive abilities be effectively used? (iii) What should we do to help them use effectively the interactivity provided?

The paper is organised as follows. First we review research concerning external representations (ERs). Secondly, we describe the CBLEs developed for this investigation. Third, we illustrate the experiment run. Fourth, the learners' interaction patterns with the CBLEs are presented. In the Discussion section we review the findings and reflect on how these illuminate the three questions referred to above. Finally, we suggest future research.

Background

Different types of external representations (ERs) impact decisively on task performance and how a task is performed (Larkin & Simon, 1987; Scaife & Rogers,

1996). Larkin and Simon (1987) showed that diagrams seem to help problem solving in geometry. Using a geometry problem the authors concluded that a diagram was advantageous to the problem solving at hand. Its utilisation may be preferable to the use of a sentential representation by easing search, decreasing cognitive load associated with labelling and easing recognition of necessary elements for a correct inference.

Cheng (1999) suggests that adding interactivity can reduce the learner's effort when constructing and/or manipulating a graphical representation for problem solving. Furthermore, interactive graphical representations can offload low-level cognitive activities allowing the learner to concentrate on high-level ones. However, Rogers (1999) states the need for a better understanding of how interactive graphical representations can facilitate problem solving and, in doing so, what are the learning benefits of making the task easier. The main problem lies, therefore, in determining the optimal balance between how much of the solution is readily available from the representation and how much is left for the learner to do.

Learners' levels of cognitive ability interact with the way they use ERs. Hegarty and Kozhevnikov (1999) and Kozhevnikov, Hegarty and Mayer (2002) found a significant a correlation between spatial visualisation test scores, using the Paper Folding Test, and tasks that demand the ability to mentally animate several components of a diagram at the same time. Kirby and Boulter (1999) report that previous geometry knowledge and level of spatial ability are important variables in the prediction of geometry test scores.

Learners might also show a preference for a certain type of external representation modality (Cox, 1999). Many studies classify learners according to one dimension: visualisers and verbalisers. Visualisers tended to choose graphical representations while verbalisers choose sentential ones. However, Cox (1999) considers this distinction too simplistic. Instead another dimension might be the ability to translate between different modalities or to stick to just one. Moreover, Hegarty et al. (Hegarty & Kozhevnikov, 1999; Kozhevnikov et al., 2002) consider the existence of two distinct types of visual-spatial representations users. Some persons are good at pictorial imagery, building detailed visual images. Others are good at schematic imagery, meaning that they are able to represent accurately the spatial relationships between objects, and imagining spatial transformations. Hegarty and Kozhevnikov (1999) found that schematic visualisers were good performers of mathematical problems. However, no significant correlation was found between pictorial visualisation and performance on mathematical problems.

CBLEs usually employ more than one type of external representation. There has been considerable research into whether providing information in different modalities simultaneously can hinder or facilitate learning (Ainsworth, Wood, & Bibby, 1996; Ainsworth, Wood, & O'Malley, 1998). Several studies report learners'

difficulties in translating between different types of representations (Ainsworth et al., 1996; Ainsworth et al., 1998; Cox, 1996; Cox, Stenning, & Oberlander, 1995). Ainsworth (1999) provides a taxonomy of the different functions of different forms of external representations. Her taxonomy addresses the important issue of the distinct types of relationships that different external representations might establish between themselves in a CBLE and how these might be used to foster learning.

Methodology

The CBLEs

Four CBLEs were created with different graphical representations (thus exploring the combinations of the two interactivity properties referred to in the introduction space - see Table 1 and Figure 1). Each CBLE tackled the same concept in the same manner, except for differences in the diagrams' interactivity. In total, the CBLEs had 15 diagrams each.

		Manipulating the diagrams' elements (ED)	
		Yes	No
Manipulating the whole diagram (WD)	Yes	*WD+ED*	*WD*
	No	*ED*	*Control*

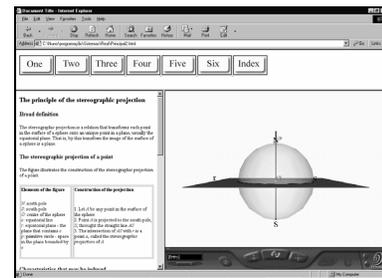Table 1 - the different CBLEs under scrutiny

Figure 1 - Snapshot of the WD+ED CBLE, showing the different frames and their content

The concept of the stereographic projection was broken into six explanatory steps, each containing a piece of text and corresponding diagram(s). In some cases there was some overlap in the information displayed in each of the CBLEs between the diagrams and the text. The text employed a non-quantitative approach to the explanation as there were no algebraic or analytical expressions present. The design of the CBLEs and writing of the content was informed by a maths teacher, a geology teacher and a maths post-graduate. Prior to the experiment, a pilot test with 3 subjects was run focusing on possible design and usability issues.

The procedure

Forty two first year undergraduates studying a geology course and thirty eight from a mathematics course

participated in the experiment. They were randomly assigned to one of the four CBLEs. For the qualitative analysis a sub-sample was formed of eighteen learners, from an initial sample of 40 videoed participants.

The experimental design comprised three stages. The first stage involved the participants answering a spatial visualisation ability test, the Paper Folding Test (Ekstrom, French, & Harman, 1979), and a geometry test. The geometry test used to assess the level of learners' domain knowledge was an adaptation of the GRE geometry sub-tests (Stewart & O'Toole, 1999). The standard procedures for the administration of these two tests were followed.

For the second stage the participants were randomly assigned to one of the four different conditions. The participants were told that the session would entail their learning about the geometry concept and answering a few questions displayed at the end of each text frame on a separate sheet - the multiple-choice test. The content of the CBLEs was available for consultation during the answering. The experimenter would then read the instructions for the session: (a) no constraints were imposed concerning the time available for completion; (b) no order of the exploration was forced; (c) the interactivity available was explained, briefly demonstrated and a small practice period was allowed; (d) the questions' format of the multiple-choice test was explained and (e) the participants were told that they could ask questions related to general issues of the interaction whenever they needed. Forty participants were recorded using a video camera with an open microphone aimed at the computer screen.

In the final stage, the participants were tested on the knowledge acquired in stage 2 through a questionnaire containing 16 sentences to be answered true or false - the post-test. The participants answered the post-test individually. On average the interval between answering the multiple-choice test and the post-test was 10 days[1].

Describing the analysis

The analysis of the interaction patterns concentrates on a video analysis and follows a qualitative approach. It was based on a two-stage process for the organisation and representation of the initial raw data (the videos). The preliminary work of the video analysis involved coding the observed activity, which was followed by the build up of a time based diagrammatic representation for each learner in the second stage.

The first stage involved two types of coding. The first type of coding was called Action Coding and coded actions where no special interpretative rules or skills were needed. The second type of coding was called Interpretative Coding and its main goal was to understand on which part of the CBLE the learner was focusing his/her attention,

giving the possibility to analyse the transitions between the different representations - text and diagrams.

In the second stage, a time-based diagrammatic representation of each learner's activity was constructed in order to facilitate the extraction of information to help identify activity patterns.

The basic categories and coding scheme used

The Action Coding categories were divided into two main groups. Diagram actions, where the learners' actions over the diagrammatic representations were recorded - moving the diagrams' elements or rotating them. Text actions were the learners' actions done over textual representations - scrolling the text. The actions were also coded taking into account if they were done before or after reaching the questions part o the CBLEs.

In relation to the Interpretative Coding, the main objective was to help understand learners' transitions between representations (for example, moving from a diagram to text). The Interpretative Coding categories are: *in diagrams*, when the learner is focusing on the diagrams; *in text*, when the learner is focusing on the text; *neutral*, when the interpreter is unable to decide which of the previous categories apply due to lack of evidence.

The use of the Interpretative Coding involved the application of a set of rules of thumb in order to maintain consistency in the coding process. The rules of thumb used were: (a) novelty effect - if the learner clicks on a hyperlink one can assume that he will attend to the new information available; (b) consecutive actions on a representation - if a learner is interacting with a representation repeatedly with short time delays one might assume that there is a certain degree of focused attention on that representation; (c) mouse activity - if the user is tracking parts of the representation with the mouse then it can be assumed that his/her attention is being focused on that representation.

In order to guard against observer biases, studies that use direct observation utilise more than one person for the coding procedure. However, due to limited research resources an alternative method was used. The author explained the coding procedure to a PhD student and asked her to code 15 minutes of two participants. The two participants chosen were using the control and the WD CBLE. The coding was then compared with the author's own coding. In relation to the control condition and the Action Coding, from the 31 actions coded by the DPhil student 25 were in accordance with the author's coding - approximately 80%. The rate for the Interpretative Coding was 7 out of 10 - 70%. Considering the WD ILE, the rate for the Action Coding was 23 out of 29 (79%) while for the Interpretative Coding it was 11 out of 17 (70%).

The interaction patterns

This analysis identifies patterns of activity and their relationships with test scores. We start by considering time-based measures and unambiguous actions. After, we

---

[1] It should be noted, however, that due to participants' time availability the lag between interacting with the CBLEs and answering the PT was not equal for all the participants. This was taken into account in the analysis.

turn our attention to the patterns of activity that demand some degree of interpretation in order to be identified - the transitions between representations. The analysis is supported by a series of correlations to help tease out the patterns.

Can we identify a relationship between the test scores and the way learners used the time when interacting with the CBLEs?

Previous quantitative analysis suggested that the time spent interacting with the CBLEs was not a good predictor of success on the multiple-choice test and post-test. The more detailed measures considered in the present analysis provide additional information on the actual impact of time on learning performance. The analysis focuses on the time spent on each step and how it is allocated between "before" and "after" reaching the questions part of the CBLE. Furthermore, two different time measures are used: (a) the time before reaching the questions (TIME BEFORE QUESTIONS) and (b) the ratio between the TIME BEFORE QUESTIONS and the total time spent in the step (RELATIVE TIME BEFORE QUESTIONS).

Considering the nature of the two types of time measures used, Table 2 suggests that learners who took longer to reach the questions had, on average, better scores in the multiple-choice test and post-test. Such results hint that jumping into the questions is not a good strategy and that answering the questions needs a certain previous degree of exploration of the content (bear in mind that, for the multiple-choice test, the content was available for consultation during the answering).

	TIME BEFORE QUESTIONS	RELATIVE TIME BEFORE QUESTIONS
MCT	r(18)=.413, sig.=.088	r(18) =.471, sig.=.048
PT	r(14) =.259, sig.=.371	r(14) =.499, sig.=.069
GT	r(18)=-.350, sig.=.154	r(18) =-.535, sig.=.022
PFT	r(14) =.175, sig.=.488	r(14) =-.142, sig.=.574

Table 2 - Correlations between time measures with multiple choice test (MCT), post-test (PT), geometry test (GT) and Paper Folding Test (PFT)

Table 2 suggests that the learners with higher scores on the geometry test were jumping into the questions part since the geometry test correlates negatively and significantly with the time measures. From the previous table we can see this strategy did not seem to be successful.

Can we identify relationships between the test scores and basic recorded actions?

The basic recorded actions are the ones considered unambiguous. In relation to scrolling the text, a significant association found was between scrolling the text before reaching the questions and the multiple-choice test scores (r(18)=.493, sig.=.037). This result suggests that reading the test before starting to answer the questions is preferable to reading the text while trying to answer the questions.

Moving the diagrams' elements did not correlate significantly with any of the tests used. However, rotating the whole diagram before reaching the questions was positively and significantly correlated with the multiple-choice test scores (r(10)=.682, sig.=.30). Once more we can see the benefits of exploring the content in more depth before starting to answer the questions.

	N	Mean of actions performed per user	Std. Deviation
move before questions	9	1.81	0.68
rotate before questions	10	3.63	2.60
move after questions	9	1.69	1.10
rotate after questions	10	1.20	0.84

Table 3 - Descriptive statistics of "move" and "rotate" before and after reaching the questions

On the one hand, Table 3 shows that the average use of "Move" does not seem to be very impressive, less than two per step. This helps explain the results referred to above. Maybe the learners did not see the need to move the diagrams' elements or were not comfortable with the functionality. On the other hand, the average use of "Rotate" is much more intense during the "before the questions" period being as much as three times higher than after the questions, clearly indicating that this functionality was especially used during the exploratory phase.

Can we identify relationships between the test scores and transitions to the representations?

The different combinations of transitions between text and diagrams were aggregated into four different types: (a) transitions to a different modality made before reaching the questions part- text to diagrams or diagrams to text; (b) transitions to the same modality made before reaching the questions; (c) transitions to text made after reaching the questions; and (d) transitions to diagrams done after reaching the questions.

Transitions occurring before and after reaching the questions have a different impact on the multiple-choice test scores: performance is more positively affected by transitions that happen before reaching the questions. Considering the period before reaching the questions, both the transitions to representations of a different modality (r(18) = .592, sig.= .005) and transitions made to representations of the same modality (r(18)= .545, sig.= .01) correlate positively and significantly with the multiple-choice test scores. For the period after reaching the questions, no transitions seem to correlate significantly with the multiple-choice test scores. The scores on the post-test, however, did not show a significant association with any of the four transitions referred to.

Considering the geometry test and the Paper Folding Test, no significant correlations were found except for one case. The transitions occurring before reaching the questions to representations of the same modality were positively associated with the Paper Folding Test scores

(r(18) = .486, p = .041) but were negatively related to the geometry test scores (r(18) = -.458, p = .058).

Aditional findings concerning the best and worst scorers on the more interactive CBLEs

Considering a comparison between pairs of the best and worst scorers on interaction patterns, the following points emerged. Differences between the patterns of the best and worst scorers were clearly discernible. The best scorers tended to spend more time exploring the content before starting to answer the questions and their explorations involved transitions between representations. Not switching between the questions and supporting representations in order to check the alternative answers was not an efficient strategy to answer the multiple-choice test. Furthermore, still regarding the best scorers' explorations, two distinct strategies were found: some learners focused on each type of representation at a time and while others preferred to switch more frequently between diagrams and text.

Additionally, we were able to observe that learners did not tend to maintain their general interaction patterns across all steps. Some of the changes might be related to the content itself but some came with familiarity with the functionality provided. Moreover, sometimes the learners did not explore the informative representations at all.

Discussion

In previous studies (Otero et al., 2001) we reported the inability of the two interactivity properties, "whole diagram manipulation" and "manipulation of the diagrams' elements", to facilitate performance on the multiple-choice test. In fact, the additional interactivity had a negative effect on the learners with more geometry knowledge and greater spatial visualisation ability. In its turn, the analysis of the post-test revealed that the learners with less geometry knowledge and lower spatial visualisation ability appear to have benefited from using the "manipulation of the diagrams' elements" property.

We would expect the "manipulation of the diagrams' elements" property to help the learners with lower levels of spatial visualisation ability. Thus, it came as a surprise to find the "manipulation of the diagrams' elements" affected only the post-test. Furthermore, the present analysis did not find an association between the action of "moving the diagrams elements" and the learning performance scores. Rotating the whole diagrams, however, was positively related with the multiple-choice test. The question left unanswered is: is there a positive relation between moving the diagrams' elements and scores on the learning tests for the low spatial visualisation ability learners? Unfortunately, the data does not include enough observations to investigate this issue. Furthermore, following Kozhevnikov, Hegarty and Mayer's (2002) distinction between pictorial and schematic visualisers, the "manipulation of the diagrams' elements" property might have helped the former. However, we would like to investigate further if the pictorial visualisers really improved their understanding of the relationships between the diagrams' elements or just memorised a set of cases.

From the findings concerning the time based measures and the transitions, one can summarise by saying that: the best scorers tended to spend more time exploring the content before starting to answer the questions and their explorations typically involved transitions between representations.

To some extent, the distinction between learners who performed transitions between the two types of representations and the others who did not seems similar to what was found by Oberlander, Cox, Monaghan and Stenning (1996) regarding frequent and less frequent modalities translators. Furthermore, the worst scorers frequently would go to search for information to support their answers in the available representations but instead of exploring in depth the information found, as the best scorers did, they seemed overconfident, soon returning to answer the questions. These two patterns of rushing back to the questions and in-depth exploration of information have similarities to what Cox (1996) has termed "judicious switching" and "thrashing". Judicious switching occurs when learners systematically take advantage of the expressive powers of each type of representation. Thrashing characterises learners who just jumped from wrongly built representations and attempted something different with other types. Another curious finding concerns the high geometry test scorers. These learners jumped more frequently into the questions part and this pattern was associated with lower learning performance. Maybe these learners were confident of their domain knowledge and did not want to spend time exploring the content.

Another topic that we think needs further investigation is the possible relationship between learners' preferences for a certain type of ER, the addition of interactivity to the preferred type of ER, the use of CBLEs with multiple representations and learning performance. It can be the case that adding interactivity to an already preferred type of ER might make its use even more compelling at the expense of the other available and informative ERs. If this is the case then strategies to promote translation between representations are crucial.

Finally, we would like to go back to the three questions posed at the beginning of this paper. First, how to gain knowledge about the learners' cognitive abilities when they are interacting with a CBLE? Making the learners answer a battery of tests whenever they want to interact with a CBLE does not seem realistic. One possibility is to investigate interaction patterns and see if we could relate these with learners' strengths and weaknesses. However, we think that the more interactive a CBLE is, the more difficult this task is since the possibilities for action increase and "successful" interaction patterns can be diverse. An alternative is to focus only on clearly unsuccessful interaction patterns. Another possibility is to devise questions about the way learners are using the

different ERs while they are using them. The validity of this approach could be researched through a psychometric study focusing on the answers to these questions and their relationships with tests of cognitive abilities as well as different types of learners (for example, the distinction between pictorial and schematic visualisers or translators and non-translators).

Second, can the knowledge acquired regarding learners' cognitive abilities be effectively used? Considering the work done regarding the distinction between types of learners we could investigate further how the different types answer the questions referred to above. The next step would be to try to devise different instructional strategies to cater for the distinct groups.

Third, what should we do to help them use the interactivity provided effectively? One way is to provide feedback and help, focusing on fostering clusters of "efficient" learners' actions, enabling the learners to extract important information to reason. From the literature review and empirical evidence it seems important to consider three distinct aspects: (a) how the learner is using the interactivity of a certain ER; (b) how the use of a specific ER is affecting the use of the other available ERs and interactivity; and (c) think of ways to tackle (a) and (b) separately but also how the two issues might be related.

Acknowledgements

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2-3), 131-152.

Ainsworth, S., Wood, D., & Bibby, P. A. (1996). Co-ordinating multiple representations in computer based learning environments. In P. Brna, A. Paiva & J. Self (Eds.), *Euro ai-ed: European conference on artificial intelligence in education* (pp. 336-342). Lisbon: Edicoes Colibri.

Ainsworth, S., Wood, D., & O'Malley, C. (1998). There is more than one way to solve a problem: Evaluating a learning environment that supports the development of children's multiplication skills. *Learning and Instruction, 8*(2), 141-157.

Cheng, P. C. H. (1999). Interactive law encoding diagrams for learning and instruction. *Learning and Instruction, 9*(4), 309-325.

Cox, R. (1996). *Analytical reasoning with multiple external representations.* Unpublished Ph.D., University of Edinburgh, Edinburgh.

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction, 9*(4), 343-363.

Cox, R., Stenning, K., & Oberlander, J. (1995). The effect of graphical and sentential logic teaching on spontaneous external representation. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, 2*(4), 1-20.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). *Cognitive factors: Their identification and replication.* Forth Worth, Texas: Texas Christian University Press (for) The Society of Multivariate Experimental Psychology.

Hegarty, M., & Kozhevnikov, M. (1999). Spatial abilities, working memory and mechanical reasoning. In J. Gero & B. Tversky (Eds.), *Visual and spatial reasoning in design. Pre-prints of the international conference in design.* (pp. 221-241). Cambridge: MIT.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684-689.

Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences, 9*(1), 19-42.

Kirby, J. R., & Boulter, D. R. (1999). Spatial ability and transformational geometry. *European Journal of Psychology of Education, XIV*(2), 283-294.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*(1), 47-77.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Spatial abilities in problem solving in kinematics. In M. Anderson, B. Meyer & P. Olivier (Eds.), *Diagrammatic representation and reasoning.* London: Springer-Verlag.

Larkin, J., & Simon, H. (1987). Why a diagram (sometimes) worth ten thousand words. *Cognitive Science*(11), 65-99.

Oberlander, J., Cox, R., Monaghan, P., Stenning, K., & Tobin, R. (1996). Individual differences in proof structures follwing multimodal logic teaching. In *Proceedings of the 18th annual conference of the cognitive science society* (pp. 201-206). Hillsdale, NJ: Lawrence Erlbaum Associates.

Otero, N., Rogers, Y., & Du Boulay, B. (2001). Is interactivity a good thing? Assessing its benefits for learning. In M. Smith & G. Salvendy (Eds.), *Systems, social and internationalization design aspects of human-computer interaction* (pp. 790-794). New Jersey: Lawrence Erlbaum.

Rogers, Y. (1999). What is different about interactive graphical representations? Commentary. *Learning and Instruction, 9*(4), 419-425.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*(2), 185-213.

Stewart, M., & O'Toole, F. (1999). *Arco teach yourself the gre in 24 hours.* New York: IDG Worldwide.