Measuring the Power of Learning.®

Research Memorandum

ETS RM–18-07

Facilitating Graph Comprehension Using a Cognitive Model of Successful Solution Processes

Jung Aa Moon

Gary Feng

Jennifer Lentini

Diego Zapata-Rivera

May 2018

ETS Research Memorandum Series

Facilitating Graph Comprehension Using a Cognitive Model of Successful Solution Processes

Jung Aa Moon, Gary Feng, Jennifer Lentini, and Diego Zapata-Rivera
Educational Testing Service, Princeton, New Jersey

May 2018

Corresponding author: Jung Aa Moon, E-mail: jmoon001@ets.org

Suggested citation: Moon, J., Feng, G., Lentini, J., & Zapata-Rivera, D. (2018). *Facilitating graph comprehension using a cognitive model of successful solution processes* (Research Memorandum No. RM-18-07). Princeton, NJ: Educational Testing Service.

Action Editor: John Sabatini

Reviewers: Edith Aurora Graf and Gabrielle Cayton-Hodges

Abstract

This study investigated the effect of a model-based learning intervention in the math domain. A cognitive model was developed to represent expert solution processes for a range-based graph comprehension task. Based on the model, a learning intervention was designed to target cognitive steps that pose major challenges for students. We present preliminary results that suggest positive effects of the intervention in middle school students.

Key words: cognitive model, graph comprehension, learning

Cognitive models have been used as the basis of intelligent tutoring systems (ITS) to guide and support student learning (Anderson, Corbett, Koedinger, & Pelletier, 1995). Some of these systems involve expert models that contain correct and incorrect solution steps to diagnose student errors and provide appropriate feedback (Heffernan, Koedinger, & Razzaq, 2008). The current study used a cognitive model to address major cognitive challenges in a graph comprehension task. The expert solution processes represented in the model guided development of a learning intervention. We present the results from a pilot study in which we tested this intervention with middle school students. This work shows how cognitive models can be used to gather evidence of “cognitive bottlenecks” that students encounter and to guide the development of learning materials. These learning materials could potentially be integrated into an ITS or other formative assessment system.

Modeling Graph Comprehension

One of the major challenges in graph comprehension is interpretation of ranges (Bell, Brekke, & Swan, 1987; Swan & Phillips, 1998). When asked to identify a range that satisfies a given condition (e.g., “When is B greater than A?”), many students make a pointwise interpretation (e.g., 8 p.m.) instead of a rangewise interpretation (e.g., 7–9 p.m.). While the former involves extracting localized information, the latter involves more complex processes of identifying and comparing overall patterns of data. The “T-shirt problem” (Figures 1 and 2), developed as a part of the conversation-based assessment research (Zapata-Rivera, Jackson, & Katz, 2015), illustrates such challenges in range-based graph comprehension. It involves a graph of three linear functions that plot the total cost charged by three companies as a function of the number of T-shirts. Students are asked to help the student council find the company that minimizes the cost of making T-shirts. Students are also told that there are about 300 students in the class, but the student council does not yet know how many students will buy a T-shirt.

Student responses in this problem can be classified as (a) answers based on range-based reasoning and (b) answers based on point-based reasoning. Using range-based reasoning, one can recommend different companies for different ranges of T-shirts (e.g., Company C for 0–100 T-shirts and Company B for 100–200 T-shirts). Using point-based reasoning, one can recommend one company that minimizes the cost for a certain number of T-shirts. Although there are multiple correct answers for this problem, our view is that the more desirable answer involves range-based reasoning. Range-based reasoning takes into account the constraint that the student council does not yet know how many students will buy a T-shirt. It involves examining all possible T-shirt numbers within the given range (0–300), which requires more comprehensive graph processing than making a recommendation based on specific T-shirt numbers. In addition, range-based reasoning involves higher order cognitive skills of identifying global patterns of data, as compared with point-based reasoning, which mainly involves extracting localized information.

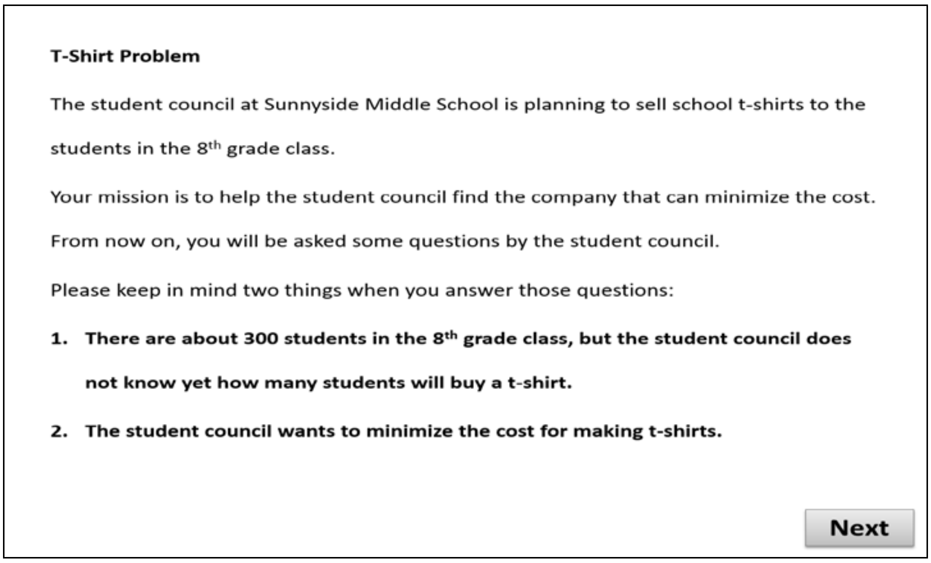

Figure 1. T-shirt problem introduction.

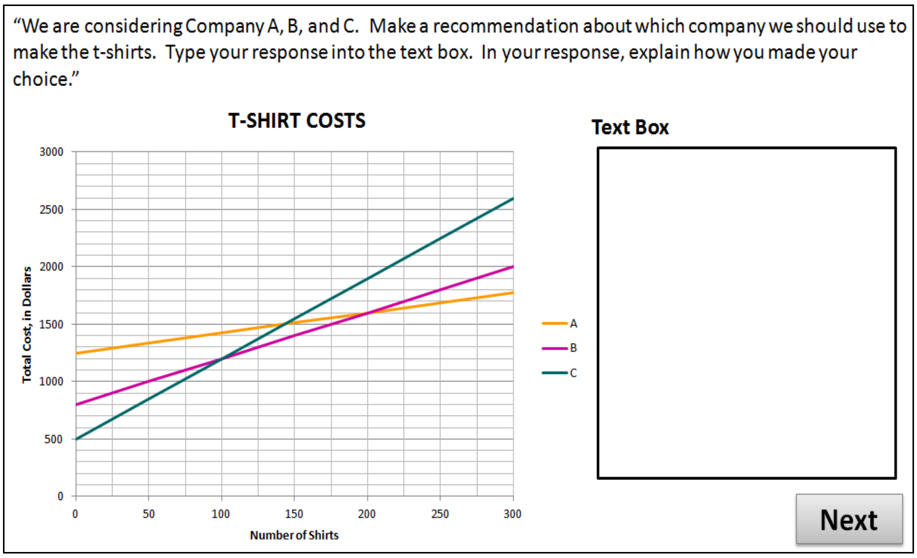

Figure 2. T-shirt problem prompt.

A conceptual cognitive model of the task was developed based on the ACT-R approach (Anderson et al., 2004) to specify the solution steps involved in range-based reasoning. These steps were identified from the adults who solved the T-shirt problem. The following steps were identified as cognitive bottlenecks that pose major challenges for students:

(a) find that the cost-minimizing company differs for different T-shirt numbers and (b) use intersecting points to identify the ranges for which the cost-minimizing company differs. Prior research has shown students' misconceptions and errors in these steps (Bell et al., 1987; Mevarech & Kramarsky, 1997). We predicted that supporting students in these two steps would lead to positive learning outcomes. To test this prediction, we developed learning materials that specifically targeted those bottlenecks and examined their effects.

Study

A pilot study was conducted with 10 students in Grades 7–8. The students were assigned to two between-subject conditions. Students in the experimental condition ($N = 6$) received video tutorials that demonstrated solution processes represented in the cognitive model. Students in the control condition ($N = 4$) received video tutorials that demonstrated solution processes for comparing slopes and intercepts of functions. The control tasks are appropriate for the target grade levels, yet their solution steps are not relevant to our cognitive bottlenecks. The experiment consisted of two tutorial sessions and three tests (pre-, mid-, & posttest) in which students solved different versions of the T-shirt problem. In the experimental condition, the Session 1 tutorial demonstrated that the cost-minimizing company differs for different T-shirt numbers. The Session 2 tutorial demonstrated how to identify appropriate ranges. The control students received tutorials that taught how to compare intercepts (Session 1) and slopes (Session 2). To prevent memory carryover, each of the tests presented a new graph that had intersecting points different from those in the previous test. The graphs in the tests were also different from the graph used in the video tutorials.

Thirty student responses (three tests x 10 students) were coded by two raters based on a rubric (Table 1) with nine codes. For each response and code combination, the raters assigned *0* or *1* to indicate whether the response satisfied the description of the given code. Codes 1 through 6 were intended to capture what kind of information from the given graph was used to make a recommendation (e.g., slope, point, and range). Codes 7 through 9 were to code how many companies are recommended (*0*, *1*, and *3*[1]). Codes 7 through 9 are mutually exclusive. Each student response belonged to only one of the three categories. Codes 1 through 6 are not mutually exclusive because students could make a recommendation using more than one kind of information—for instance, using both points and ranges. The interrater reliability measure (kappa) was calculated for each code. The resulting kappas ranged from .71 to 1.0.

Table 1. Rubrics for Coding Constructed Responses

ID	Description	Sample response (verbatim)
Code 1	Student uses slope (or cost for each additional T-shirt) to make a recommendation.	"I think that the student council should use company A because the slope of the line is less, that means that the price for many shirts won't be too high."
Code 2	Student uses only one specific T-shirt number to make a recommendation.	"I would reccomend choosing company A because the cost for 300 shirts would be less than the other companys."
Code 3	Student uses more than one specific T-shirt number to make a recommendation.	"If 50 students wanted a shirt I would reccomend K if 100 students wanted a shirt I would choose J and if 300 I would reccomend L."
Code 4	Student uses one or more ranges of T-shirts to make a recommendation, but the ranges do not accompany specific numbers or they do not correspond to the intersecting points.	"I recommend Company N because it has the cheapest prices for the longest range of T-shirt amounts and is never the most exspensive. This way, for most amounts of T-shirts Company N will have the cheapest price."
Code 5	Student uses one or more ranges of T-shirts to make a recommendation. The ranges correspond to the intersecting points at a reasonable accuracy.	"I would recommend Company O for 0–50 and than Company N for 50–200 and than Company N for 215–300."
Code 6	Student uses other criteria to make a recommendation.	"Company C has the most number of T-shirts for the cheapest price. I looked at Company A and saw that it was very expensive, so I didn't even put them into consideration. Than, I looked at Company B and saw that thier shirts were cheaper but still more expensive than Company C."
Code 7	Student does not recommend any company.	"I would recommend none of the companies because all of them charge you money to buy 0 shirts."
Code 8	Student recommends one company.	"I recommend Company N because it has the cheapest prices for the longest range of T-shirt amounts and is never the most exspensive."
Code 9	Student recommends three companies.	"If the amount was less than 150, I would recommend company K, but if it was more than 150 and less than 250, I would recommend company J, and if it was more than 250, I would recommend company L."

Results

Tables 2, 3, and 4 show the number of students who recommended different numbers of companies in the three tests. All control students recommended one company in all tests. Five out of six experimental students recommended one company in the pretest, and this number dropped to three in the midtest. In contrast, the number for three companies increased from pre- (0) to midtest (3).

Table 2. Number of Students Who Recommended Different Numbers of Companies in the Pretest Control and Experimental Conditions

Condition	0C	1C	3C
Control	0	4	0
Experimental	1	5	0

Note. 0C = no recommendation; 1C = a one-company recommendation; 3C = a three-company recommendation. No student recommended two companies.

Table 3. Number of Students Who Recommended Different Numbers of Companies in the Midtest Control and Experimental Conditions

Condition	0C	1C	3C
Control	0	4	0
Experimental	0	3	3

Note. 0C = no recommendation; 1C = a one-company recommendation; 3C = a three-company recommendation. No student recommended two companies.

Table 4. Number of Students Who Recommended Different Numbers of Companies in the Posttest Control and Experimental Conditions

Condition	0C	1C	3C
Control	0	4	0
Experimental	0	3	3

Note. 0C = no recommendation; 1C = a one-company recommendation; 3C = a three-company recommendation. No student recommended two companies.

Tables 5, 6, and 7 show the number of students who demonstrated point-based (Codes 2 and 3) and range-based (Codes 4 and 5) reasoning. In the control condition, the number with point-based reasoning dropped from pre- to posttest, whereas the number with range-based reasoning did not change. In the experimental condition, the number with point-based reasoning decreased from pre- to posttest, whereas the number with range-based reasoning increased. Further, four out of the five experimental students who showed range-based reasoning in the posttest were classified as Code 5 (i.e., correct ranges). This indicates that most of them were able to correctly identify ranges in a new graph instead of simply giving any answer that involved ranges. Overall, our results suggest some positive effects of the Session

1 intervention on the first cognitive bottleneck (i.e., the cost-minimizing company differs for different T-shirt numbers) and the Session 2 intervention on the second cognitive bottleneck (i.e., how to identify appropriate ranges).

Table 5. Number of Students With Point-Based and Range-Based Reasoning in the Pretest Control and Experimental Conditions

Condition	Point	Range
Control	2	2
Experimental	5	0

Note. “Point” indicates point-based reasoning. “Range” indicates range-based reasoning.

Table 6. Number of Students With Point-Based and Range-Based Reasoning in the Midtest Control and Experimental Conditions

Condition	Point	Range
Control	3	1
Experimental	5	1

Note. “Point” indicates point-based reasoning. “Range” indicates range-based reasoning.

Table 7. Number of Students With Point-Based and Range-Based Reasoning in the Posttest Control and Experimental Conditions

Condition	Point	Range
Control	0	2[a]
Experimental	2	5

Note. “Point” indicates point-based reasoning. “Range” indicates range-based reasoning.
[a]Both of these students are the ones who showed range-based reasoning in the pretest.

Discussion

The results of the current study suggest that (a) the cognitive bottlenecks in the model represented major challenges in solving the T-shirt problem and (b) our interventions had preliminary, yet positive, effects on overcoming those challenges. Two main limitations of the current study are (a) the small sample size and (b) the lack of a control condition in which students perform no particular problem-solving activities between the tests. Additional data collection is necessary to address these limitations.

Our results provide implications for the use of cognitive models in learning and assessment research. As demonstrated in this study, ideal solution processes of cognitive models can be used to develop teaching and instructional interventions and help students overcome their cognitive challenges. These models could potentially be used as part of an ITS or a formative assessment system to assess changes in student cognition, update the student model based on this information, and select appropriate feedback/learning materials. Future directions for this research involve applying this model-based intervention approach to different kinds of graph comprehension problems, evaluating the effectiveness of different types of interventions (e.g., learning materials/feedback), and extending the approach to various other problems in the math domain.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036–1060. https://doi.org/10.1037/0033-295X.111.4.1036

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*, 167–207. https://doi.org/10.1207/s15327809jls0402_2

Bell, A., Brekke, G., & Swan, M. (1987). Diagnostic teaching: 5 graphical interpretation teaching styles and their effects. *Mathematics Teaching*, *120*, 50–57.

Heffernan, N. T., Koedinger, K. R., & Razzaq, L. (2008). Expanding the model-tracing architecture: A 3rd generation intelligent tutor for algebra symbolization. *International Journal of Artificial Intelligence in Education, 18*, 153–178.

Mevarech, Z. R., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, *32*, 229–263. https://doi.org/10.1023/A:1002965907987

Swan, M., & Phillips, R. (1998). Graph interpretation skills among lower-achieving school leavers. *Research in Education, 60*(1), 10–20. https://doi.org/10.1177/003452379806000102

Zapata-Rivera, D., Jackson, T., & Katz, I. R. (2015). Authoring conversation-based assessment scenarios. In R. A. Sottilare, A. C. Graesser, X. Hu, & K. Brawner (Eds.), *Design recommendations for intelligent tutoring systems: Vol. 3. Authoring tools and expert modeling techniques* (pp. 169–178). Orlando, FL: U.S. Army Research Laboratory.

Notes

[1]A code for two-company recommendations is not available because none of the students recommended two companies.