Utilizing Cognitive Interviews to Improve Items That Measure Mathematical Knowledge for Teaching Community College Algebra

Bismark Akoto	Dexter Lim	Irene Duranczyk
University of Minnesota	University of Minnesota	University of Minnesota

The Algebra Instruction at Community Colleges: Validating Measures of Quality Instruction project (VMQI) developed an instrument to measure mathematical knowledge for teaching community college algebra (MKT-CCA) and conducted cognitive interviews with 12 community college instructors of College Algebra. Thirty-six drafted MKT-CCA test items were reviewed with each item being reviewed by two instructors. This paper presents the lessons learned from this preliminary analysis of the cognitive interview transcripts.

Keywords: Cognitive interviews, mathematical knowledge for teaching, community college, college algebra

Introduction

The Algebra Instruction at Community Colleges: Validating Measures of Quality Instruction (VMQI) is a National Science Foundation-funded project that is developing an instrument to (1) measure mathematical knowledge for teaching community college algebra (MKT-CCA) and (2) identify the dimensionality of this knowledge. We began by situating our conceptualization of mathematical knowledge for teaching vis-a-vis one of the most common conceptualizations produced by Deborah Ball and colleagues (Ball et al., 2008; Hill et al., 2008). Ko and Herbst (2020) further proposed organizing the subject matter knowledge in Ball et al.'s framework by tasks of teaching. To move this work further in community colleges (CC), we have hypothesized to measure two main constructors, Tasks of Teaching (Choosing Problems and Understanding Student Work) and Function Types (linear, exponential, and rational functions that are foundational for advanced work in calculus) (see Figure 1). This conceptualization of algebra knowledge for teaching framed the instrument's development. Cohen and Swerdlik (2009) outline the development of an instrument in five main stages: conceptualization, construction, tryout, analysis, and revision. In the construction stage, cognitive interviews (CIs) play a significant role.

Cognitive interviews are an important qualitative method used by researchers to discern interviewees' understanding of survey items (Meadows, 2021; Willis, 2005). As part of the development of our MKT-CCA instrument, we conducted CIs with CC college algebra instructors to understand whether the items we developed to assess the dimensions (see Figure 1) are interpreted by respondents as intended and whether the participants used the intended knowledge (Mesa et al., 2020-2023). We used the following research questions to guide the analysis.

1. How do CC college algebra instructors interpret our developed items in relation to the dimensions of our blueprint?
2. What knowledge do CC college algebra instructors use to respond to the MKT-CCA instrument?

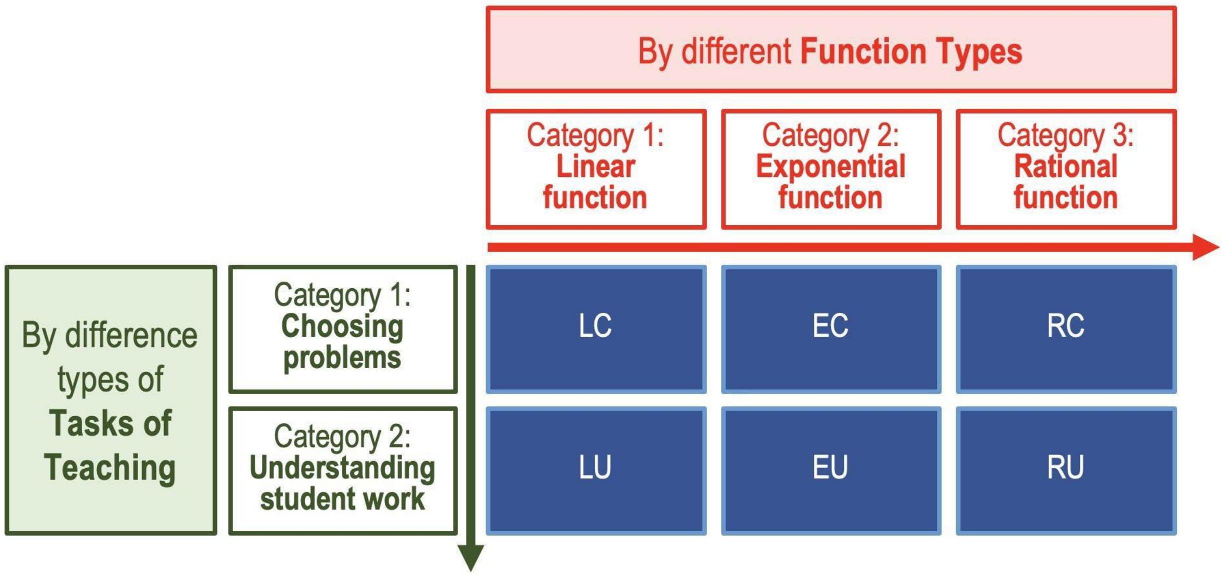

Figure 1. MKT-CCA blueprint showing its six dimensions.

Method

Materials and Procedures

Research on CIs suggest conducting "rounds" that consist of 5 to 15 interviews, which are ideally repeated following efforts to revise questions and eliminate problems (McColl, 2001; Willis, 1994, 2005). After the initial instrument construction and before piloting the instrument, CIs provide information to evaluate whether the items proposed uses the knowledge that is intended. Typically, CIs are planned to take a 3:1 ratio (Willis, 2015) of time (three times more than it would take to answer the item without think-aloud or probing). As a team, we decided to limit interviews to at most 90 minutes to eliminate fatigue (D'Ardenne, 2015). For this reason, we sent two items out of six planned for each respondent, ahead of time so that respondents could think about them and take notes beforehand. By doing this, we anticipated that we could average 12 minutes per item during the CIs. At the time of our CIs, our item pool consisted of 60 draft items that we had deemed ready for review. The 36 items selected (six items per dimension) would provide us with the information to understand whether the draft items were interpreted by respondents as intended, used the knowledge we thought was needed, and targeted the dimensions of the MKT-CCA blueprint.

We developed a protocol that we agreed to use for all interviews. Three teams of two researchers each performed the interviews (one person would be asking the questions, whereas the other was in charge of taking field notes and managing technical aspects). Each team interviewed four participants. We selected 36 items with the goal of having each item be reviewed by two participants. All interviews were conducted online in the last week of July 2021 using Zoom with screen sharing, and audio and video recording with transcription. We used Jamboard to present the items one at a time since it allows for real time documenting the participants' item solving process. We gave each participant two items to work on ahead of the interview and asked them to send us the documentation of their work before the scheduled interview. We asked them to keep track of how long it took them to complete each item. The interview protocol asked for their impressions of the content of the item, whether the content was relevant to their work, the answer they chose and their reasoning, and the reason(s) for eliminating other options.

Analysis

In this paper, we use the interview response from a respondent on an item as our unit of analysis. This gave us a total of 72 (one missing data point due to time restraint) interviews on 36 items. All the interviews were fully transcribed and summarized. The field notes were also coded. Analysis was done in two phases. The first phase focused broadly on answering whether respondents (a) choose the correct answer, (b) use correct reasoning in choosing the answer, (c) choose an incorrect option with correct reasoning, and (d) refer to the dimensions of the blueprint. This phase of coding was done by five members of the research team.

The second phase involved adapting Conrad and Blair's (1996) cognitive coding scheme and was done by the three authors. The coding scheme proposed attending to three categories: *Understanding the Intent*, *Performing the Primary Task,* and *Formatting the Response*. Conrad and Blair used *Understanding the Intent* to code the respondents' comprehension and appreciation of the researcher's objectives; *Performing the Primary Task* was used to code how well the respondents were able to retrieve information from memory, do mental arithmetic and evaluate a response. *Formatting the Response* was used to map how respondents responded to prespecified answer options. We adapted this coding scheme for our purposes as follows:

Understanding the Intent. This category examined whether the respondent was able to place the item in a dimension of our blueprint. Did the respondent recognize both the task of teaching and the function type? We coded a yes if the respondent recognized both the function type and the task of teaching, a partial if only one was identified and a no if neither was present.

Using Knowledge. This category examined whether the respondent understood the instructional goal embedded in the stem. Did the respondent reason about the stem in ways that conformed with our reasoning? Did the respondent use mathematical knowledge to make sense of the stem? Did the respondent's response indicate whether the stem was clear? Did the respondent give any indication of confusion of the stem? Did the respondent understand the context presented in the stem? This included instances where the wording was not clear but did not include situations in which the respondent did not understand the mathematical ideas or terminology in the stem. We coded the response as yes if it indicated a clear understanding of the stem, partial if the understanding was not clear and no if the response indicated that the respondent did not understand the stem.

The stem is the part of the item that defines the mathematical task. It is followed by possible solutions (optional answers).

Formatting the Response. This category examined whether the respondent's reasoning on the options was mathematically correct or aligned with our template. Was the respondent's correct or incorrect selection of options and elimination of distractors expected? We coded the formatting as expected if the justification matched our template intent or was found to be mathematically reasonable. It was coded as problematic when the respondent gave a correct answer or eliminated an option for the wrong reasons, when the reasons could not be mathematically justified, or when the justification given did not conform to our intent. It was coded as partial if the response was a mix of expected and problematic.

Findings

We present the preliminary findings of our analysis in this section. Table 1 shows the summary of the analysis using the three categories. Overall, 88.73% of participants' responses showed a recognition of the function type and task of teaching, 85.92% of responses showed an understanding of the stem and 61.97% of participants responses to the items had mathematically appropriate justifications.

Table 1: Percentage of Responses to the Individual Items (N=72).*

Dimension of Blueprint	Understanding the Intent			Using Knowledge			Formatting the Response		
	Yes	Partial	No	Yes	Partial	No	Expected	Partial	Problematic
EC (n=12)	100.0	0.0	0.0	91.7	0.0	8.3	66.7	8.3	25.0
EU (n=12)	91.7	8.3	0.0	91.7	8.3	0.0	66.7	0.0	33.3
LC (n=11*)	100.0	0.0	0.0	90.9	9.1	0.0	45.5	0.0	54.5
LU (n=12)	83.3	16.7	0.0	91.7	0.0	8.3	100.0	0.0	0.0
RC (n=12)	66.7	25.0	8.3	83.3	8.3	8.3	41.7	0.0	58.3
RU (n=12)	91.7	8.3	0.0	91.7	0.0	8.3	50.0	8.3	41.7

*One data point is missing. One participant only responded to five items.

Research Question One

Research question one sought to answer how CC college algebra instructors interpret our developed items in relation to the dimensions of our blueprint. This will be addressed using the *Understanding the Intent* code.

Understanding the Intent. As expected, preliminary analysis shows that participants mostly indicated a recognition of the function type of the items with only one exception. However, it was noted that the function type was mentioned in the stem of the item and is likely to be the reason the participant did not mention it in the interview. One main observation noted is that participants mostly were able to identify the task of teaching except in the case of choosing problems for rational functions where we observe a 33.3% of responses not indicating the task of teaching.

Research Question Two

We wanted to know what knowledge CC college algebra instructors will use to respond to the MKT-CCA instrument for research question two. We will apply the *Using Knowledg*e and *Formatting the Response* codes for this research question.

Using Knowledge. Participants demonstrated an understanding of the instructional goal in the stem for most of the items. With the exception of choosing an exponential function problem, we recorded high percentages for this category. We observed that the participants were distracted from the objectives of the items by the context of the item. For example, on an item to introduce students to interpreting the meaning of vertical asymptotes in a real world applied context, Justina (pseudo) stated, "I did not get where those numbers are coming from, 0.5 and 6380. I did not get . . . those assumptions being made about the surface or something". Justina was clearly focused on the context rather than the mathematics in the item and was distracted by the source of numbers and assumptions.

Formatting the Response. Only a little over half the reasons given for the choice of options were expected or mathematically justified. This was concerning since we had recorded high percentages for the *Understanding the Intent* and *Using Knowledge* codes. Further analysis revealed that out of the 25 problematic codes recorded, 12 (about half) of them involved rational functions. About 39% of the reasons (involving 17 items) given were deemed problematic for varied reasons. The top reasons were (1) lack of knowledge needed to correctly answer items (especially with rational functions), (2) selecting familiar options over mathematically sound options, and (3) some also were distracted by multiple contexts in one item and tend not to reason through such items mathematically. Additionally, a few items were noted to be complicated.

Emmanuel (pseudo) gave the following reason for selecting an option: "I go with [option] A only because I already do it with my students. And it's one that they enjoy and it's one that I have expanded on over time." Emmanuel does not offer a mathematical justification for his choice.

Conclusion

We learned some lessons from conducting these CIs. Respondents comprehend an item better when the item contains fewer applied contexts and confusions associated with parts of an item distract respondents from the objective. Additionally, items need to have more precise language when describing the stem and options.

Based on the feedback received from the CIs, four complicated items were discarded, the stems of seven items were revised, and the options of 14 items were revised to make them more focused, independent from each other, reasonable for the context, and possible to assess mathematically. Ideas collected from the respondents were also used to create new items and revise items that were not part of the cognitive interviews.

Questions for Discussion

1. Based on our review of the literature, isolating factors of mathematical knowledge for teaching has been elusive? What affordances or limitations do you see in our design to isolate two tasks of teaching? What else should be considered?
2. Is there another cognitive interview coding scheme that you have used to extract useful, reliable, objective, and quantifiable data from Cognitive Interviews? We may be doing cognitive interviews after our next version of the instrument. Any advice?

Acknowledgement

Funding for this work was provided by the National Science Foundation (EHR awards 1561436, 20006021, 20006442, 20005273, 20005664). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank our participants; without them this work would not have been possible.

References

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389-407.

D'Ardenne, J. (2015). Planning a cognitive interviewing project. In D. Collins (Ed.), *Cognitive interviewing practice* (pp. 59-79). SAGE Publications Ltd. https://www-doi-org.ezp2.lib.umn.edu/10.4135/9781473910102

Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430-511.

Cohen, R. J., & Swerdlik, M. E. (2009). *Psychological testing and assessment: An introduction to tests and measurement* (7th ed.). New York: McGraw-Hill.

Conrad, F., & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1–9.

Ko, I., & Herbst, P. (2020). Subject matter knowledge of geometry needed in tasks of teaching: Relationship to prior geometry teaching experience. *Journal for Research in Mathematics Education, 51*(5), 600-630. https://doi.org/10.5951/jresematheduc-2020-0163

Ko, I., Mesa, V., Duranczyk, I., Herbst, P., Kohli, N., Ström, A., & Watkins, L. (2021). Understanding the characteristics of mathematical knowledge for teaching algebra in high schools and community colleges, *International Journal of Mathematical Education in Science and Technology*, https://doi.org/10.1080/0020739X.2021.2006348

Meadows, K. (2021). Cognitive interviewing methodologies. *Clinical Nursing Research, 30*(4), 375-379. https://doi.org/10.1177/10547738211014099

Mesa, V., Watkins, L., Duranczyk, I., Beisiegel, M., Kimani, P., & Ström, A. (2020-2023). *Algebra instruction at community colleges: Validating measures of quality instruction (AI@CC 2.0)*. National Science Foundation (EHR #2000602, 2000644, 2000527, 2000566).

Miller, E. R. (2018). *High cognitive demand examples in precalculus: Examining the work and knowledge entailed in enactment* [Unpublished doctoral dissertation]. The University of Nebraska-Lincoln.

Ryan, K., Gannon-Slater, N., & Culbertson, M. J. (2012). Improving survey methods with cognitive interviews in small and medium-scale evaluations. J*ournal of Evaluation, 33*(3), 414-430.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks: SAGE Publications, Inc.

Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford: Oxford University Press.