Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 28 Number 12, November 2023 ISSN 1531-7714

Iterative Cognitive Interview Design to Uncover Children's Spatial Reasoning

Elizabeth R. Thomas, *Stephens College*
Robyn K. Pinilla, *The University of Texas at El Paso*
Leanne R. Ketterlin-Geller, *Southern Methodist University*
Cassandra Hatfield, *Southern Methodist University*

Cognitive interviews play an important role in articulating the intended construct of educational assessments. This paper describes the iterative development of protocols for cognitive interviews with kindergarten through second-grade children to understand how their spatial reasoning skill development aligns with intended constructs. We describe the procedures used to gather evidence of construct relevance and improved alignment to task-based interview items through multiple pilot rounds before conducting cognitive interviews. We found improved alignment and reduced construct irrelevant variance after protocol revisions.

Keywords: cognitive interview; learning progression; spatial reasoning; construct irrelevant variance

Introduction

Spatial reasoning is a complex cognitive construct that allows humans to structure ideas spatially and supports many daily activities (National Research Council [NRC], 2006). The construct is comprised of a set of skills individuals use to visually recognize and mentally manipulate the physical properties of objects and the spaces between them (Bruce et al., 2017; Davis et al., 2015) and is often conceptualized as "the ability to interact with, navigate in, and understand one's environment" (NRC, 2001, 2009). Some widely recognized component skills include spatial orientation and visualization (Clements, 2004; Hegarty & Waller, 2005; Linn & Petersen, 1985; Uttal et al., 2013), but other processes, such as mental rotation and perspective-taking, are also thought to contribute to spatial reasoning (c.f., Frick et al., 2014). While spatial reasoning aligns most closely with mathematics in school instruction (e.g., transformations in geometry), it supports learning across science, technology, engineering, and mathematics (STEM) domains (Lord & Rupert, 1995; Newcombe, 2017; Wai et al., 2009). For example, transformations learned in geometry can support students in ascertaining objects' orientation, which is necessary in science domains like chemistry or ideating and building structures as engineering practices (NRC, 2006). In this study, our interest relates to assessment development in mathematics education to highlight young children's development of spatial reasoning skills.

While spatial reasoning is a core component of early mathematics education (Clements, 2004; National Council of Teachers of Mathematics, 2000; NRC, 2009), it receives less curricular attention or direct classroom instruction than numeracy (Bruce et al., 2012; Clements & Sarama, 2011; Davis et al., 2015; Gilligan-Lee et al., 2022). To teach these concepts directly, it is important to understand how children reason spatially and the underlying progression of their skill development. To explore those understandings and conjecture their developmental order, we used cognitive interviews to validate a hypothetical learning progression of spatial reasoning skills for young children.

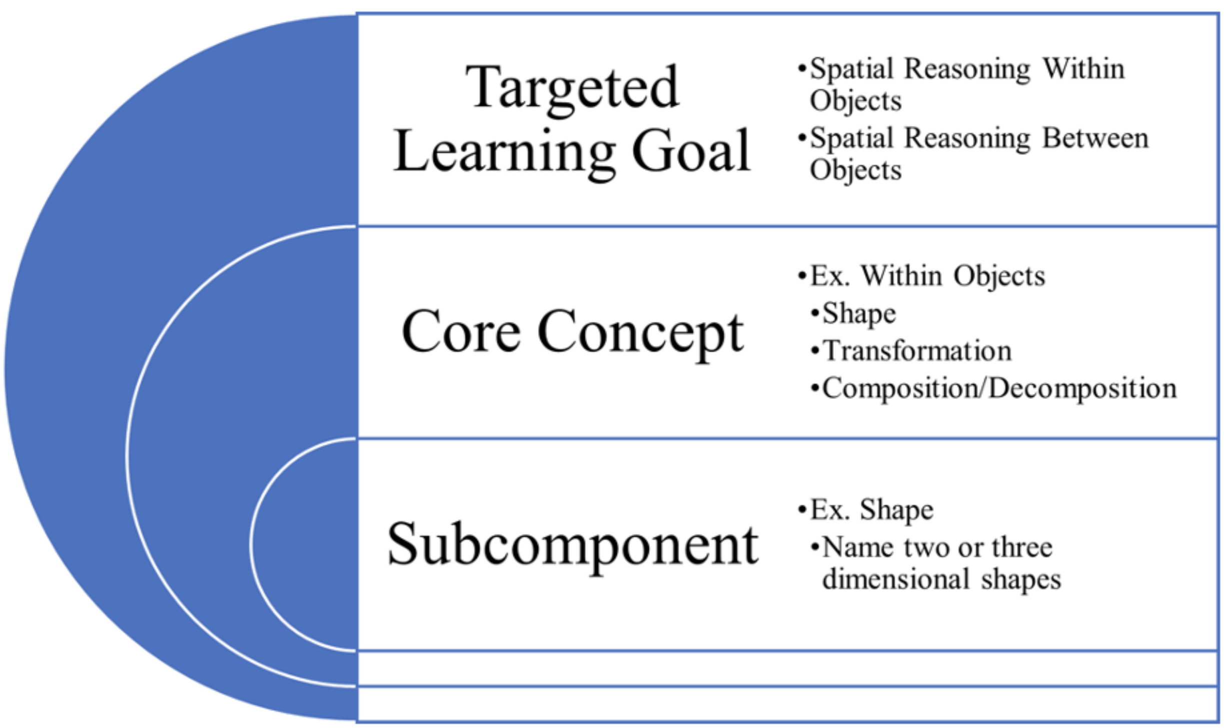
Learning Progressions

Learning progressions are hypothetical models of how individuals learn over time, based on cognitive science and learning theories, and when used as the basis for formative assessment practices, can illuminate children's thinking to help guide teachers' instructional decisions (Alonzo & Steedle, 2006; Duschl et al., 2011). As described by Confrey and Toutkoushian (2019), learning progressions can be conceptualized as climbing walls; they are flexible, non-linear paths toward increasingly more sophisticated knowledge, skills, and abilities within the specified construct domain. The climbing wall analogy illustrates the assertion that there may be multiple ways to traverse the learning process to arrive at the target construct.

As part of a larger project, our research team articulated a learning progression for the development of spatial reasoning in children in Kindergarten through Grade 2. We partition the learning progression into two targeted learning goals: reasoning spatially *within* objects and reasoning spatially *between* objects. Within each of the targeted learning goals are three core concepts, which can be thought of as sets of skills. For reasoning spatially *within* objects, the core concepts are shape, transformations, and composition and decomposition; for reasoning spatially *between* objects, the core concepts are spatial language, maps and models, and perspective taking. Drilling down further, each core concept consists of subcomponents or progressively sophisticated skills subsumed within the set (see Figure 1 for an illustration of this nesting structure). For example, in the perspective taking core concept, students may begin by taking an egocentric perspective and recognizing the view of an object or scene from their own perspective. This skill may progress to where students use photos from allocentric perspectives to recreate an object or scene. This learning progression served as the content framework for a set of classroom assessment resources to support teachers' instructional decision making.

To validate the spatial reasoning learning progression developed as part of this project, we integrated multiple sources of evidence. Following the process outlined by Ketterlin-Geller and colleagues (2013), we conducted a thorough review of research

Figure 1. Structure of the Spatial Reasoning Learning Progression



literature on how children learn to reason spatially, engaged with mathematics education researchers and mathematicians to synthesize the literature into the spatial reasoning learning progression, solicited independent reviews from mathematics education researchers, and finally, collected empirical evidence from teachers and students regarding their cognitive processing. Cognitive interviews contributed important evidence from Kindergarten through Grade 2 students about their early development of spatial reasoning skills that would ultimately contribute to the validation of the learning progression (Graf & van Rijn, 2015).

Cognitive Interviews

Cognitive interviews are a type of task-based interview that are designed to elicit participants' cognitive processing while they complete a task. They are often used when developing educational assessments to study examinees' response processes before scrutinizing item models through think-aloud interviews (Ericsson & Simon, 1993; Padilla & Leighton, 2017). Cognitive interviews are also an established technique for leveraging provisional learning progressions or trajectories to qualitatively analyze students' mathematical thinking (c.f., Luo et al., 2020). However, similar to other instrument development processes, care must be taken when developing the cognitive interview protocol to ensure that the resulting data accurately reflect students' cognitive processes.

A potential threat that may undermine our confidence in the data we obtain from cognitive interviews is construct-irrelevant variance (CIV). CIV is a type of variance that results from systematic error, or a variable that is unrelated to the construct being measured (Haladyna, 2004). When using cognitive interviews to understand students' cognitive processes, CIV in the interview protocol itself may lead to inaccurate inferences about the cognitive processes' (e.g., knowledge, skills, and abilities) underlying construct.

Construct irrelevance relates to how processes that were not targeted impact correctness on the task (American Educational Research Association et al., 2014); in turn, construct relevance can be gained through improved construct alignment to elicit intended processes. For example, if an interview question is intended to elicit a Kindergarten student's reasoning about the transformation of a polygon and the question uses vocabulary that is not common in early grades, the student's performance may not be accurately interpreted, as the understanding of the complex vocabulary is irrelevant to the intended construct. Revising the cognitive interview protocol to elicit the construct more clearly by reducing the vocabulary comprehension barrier increases the construct relevance.

When using cognitive interviews to refine the construct, ensuring the data reflect construct-relevant processes is imperative. In this study, we describe the iterative design process used to refine the task-based cognitive interview protocol to align with our early spatial reasoning learning progression. We used this process to identify and ameliorate potential sources of CIV that would lead to meaningful interpretations about students' spatial reasoning. Toward that aim, we investigated the following research questions:

1. How does iteratively designing a cognitive interview protocol facilitate eliciting observable evidence aligned to the learning progression subcomponents?
2. How does iteratively designing task-based cognitive interview items reduce construct-irrelevant variance (CIV)?

Methods

The cognitive interview protocol was associated with subcomponents that progressed from least to most sophisticated within each core concept. The development team worked carefully through the learning progression to create tasks and interview questions that included content and reasoning prompts to elicit students' thinking that assessed the designated skills and fit within the overall context of the protocol as it was developed. We developed the protocols based on our original hypothetical ordering of subcomponents and iterated them following pilot interviews. See the phases of iteration in Figure 2.

Each task on the protocol included the written subcomponent, necessary materials, interviewer script with action prompts, questions for the interviewer to ask, and scoring instructions to be used by the observer in the moment and scorer after the interview. The protocol script for each task included content questions (i.e., questions aligned to learning progression subcomponents that we scored for correctness in response or task completion); scaffolding prompts if students demonstrated that they did not understand the questions or were unable to proceed with the task; and reasoning questions to delve deeper into student' knowledge, help us understand how students thought about solving the problem and why they engaged with the tasks in certain ways. See Figure 3 for an example of an excerpt from the protocol.

Figure 2. Phases of Iteration

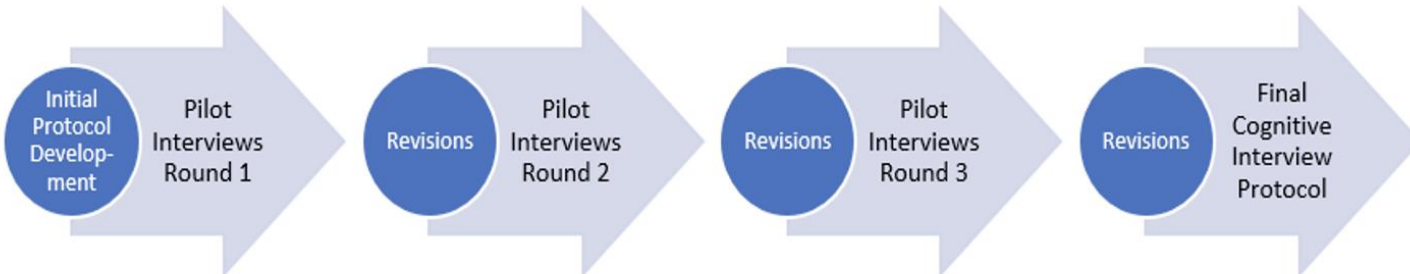
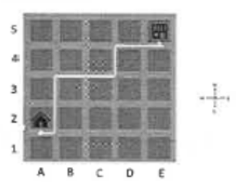
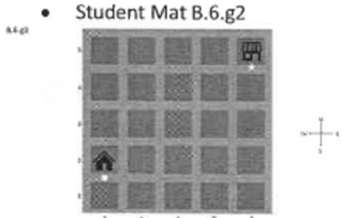


Figure 3. Sample Excerpt of the Cognitive Interview Protocol

SR.B.6.g		Describes and follows **routes** on maps.
Actions	**Questions**	**Student Response**
Materials: • Student Mat B.6.g1 should be copied on the back of Student Mat B.6.f • Sharpie marker with car sticker attached  • Student Mat B.6.g2 	*Julie starts at her house and drives to the store. Her map says to take this yellow route.* [Point to the yellow route on map in student book, but do NOT run your finger along the route to show the path explicitly.] CQ1: *This marker has a car on it so that you can pretend to drive, just like Julie did. Use this marker on this larger map to show how Julie would drive to the store using her route.*	**DURING:** nothing
Scaffolding questions	*Point to where Julie started on your map. Where did she stop?* *Which way did Julie go first? What way did she go next?*	**AFTER:** WORKBOOK (VIDEO backup) Look directly in the workbook (zoom in on video if needed/more than one route traced) and determine: Did the child follow the route correctly? 1 if yes, 0 if no. **B6gCQ1:** 0 **Flag for review:**___
Reasoning questions	*Can you use words to describe how the car drove to the store?* **Walk me back through the route. What made you decide to change directions here* [point]*? Why didn't you turn here* [point]*?*	

After specifying the hypothesized learning progression, we designed tasks related to each subcomponent. Interview prompts were developed to address both content and reasoning, to capture students' conceptions of anticipated strategies, and to uncover misconceptions or unanticipated strategies that would serve to support or refute the content and ordering of the learning progressions. Project researchers with deep content knowledge and professional backgrounds in teaching generated the task-based items using their pedagogical expertise and the extant literature from learning progression development (Ketterlin-Geller et al., 2020). By carefully crafting reasoning probes with a combination of positive and negative questions, a variety of questions were posed to begin uncovering how the children were thinking about each of the tasks presented. After reviewing draft protocols individually, we engaged in consensus discussions to refine the interview prompts. We determined the most appropriate materials for the tasks, digitized the protocols, and role-played scripts before pilot testing items with children.

Procedures

Pilot Interviews. To test the initial cognitive interview protocol prior to conducting the full-scale cognitive interviews, we pilot tested all task-based items for all subcomponents twice across three rounds of pilot interviews with a convenience sample of K-2 students. In the first round of pilot interviews, four students in Grades K-2 participated, one each in Grades K-1, and two in Grade 2. In the second round of pilot interviews, six students, two per each Grades K-2, participated, and in the third round of pilot interviews, two students, one in each Grade 1 and 2, participated.

During the pilot interviews, each student engaged one-on-one with an interviewer, who was also a project researcher, trying out items from one of the two targeted learning goals. At the same time, an observer took field notes and captured fidelity data. Each task the student was given focused on a single spatial reasoning subcomponent. This aligns with the intent of the cognitive interviews to better understand how students think about and process a single subcomponent in a progression of knowledge and skills. Between each round of pilot interviews, preliminary data was used to inform revisions to the cognitive interview protocol prior to conducting the full-scale cognitive interviews.

Data Collection

Data from the pilot tests included audio and video recordings, still photos of completed student work products, general field notes, and fidelity of administration data. The field notes and fidelity data were collected in the moment by members of the research team as they watched and listened to the interactions between the interviewer and participant. Specifically, the field notes captured nuances about the context that would not be captured through quantitative data or analysis. For instance, if a child was interviewed after lunch or was anxious about recess, we noted such factors that could indirectly relate to their task performance (see Figure 4 for an example). The observer recorded the fidelity data by interview item using a specific form (see Figure 5 for an example) to ascertain protocol clarity and student comprehension. These data were intended to help us answer research question 2: how does iteratively designing task-based cognitive interview items reduce construct-irrelevant

Figure 4. Sample Excerpt of Field Notes Form for Interviews

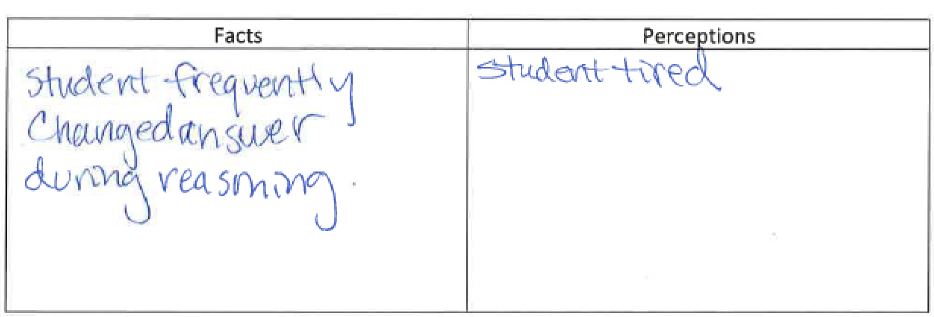

Facts	Perceptions
Student frequently changed answer during reasoning.	Student tired

Figure 5. Sample Fidelity of Administration Form for Interviews

LP: Spatial Reasoning - Between

Skill Code	Did the interviewer reword the question?		Did the interviewer repeat the question?		Did the student seem comfortable with the materials?			How comfortable did the student appear with the task?			
	0 - No	1 - Yes	0 - No	1 - Yes	NA	0 - No	1 - Yes	0 Not comfortable	1	2	3 Very comfortable
SR.B.5.a	0	1	0	1	NA	0	1	0	1	2	3
SR.B.5.b	0	1	0	1	NA	0	1	0	1	2	3
SR.B.5.c	0	1	0	1	NA	0	1	0	1	2	3
~~SR.B.6.a~~	~~0~~	~~1~~	~~0~~	~~1~~	~~NA~~	~~0~~	~~1~~	~~0~~	~~1~~	~~2~~	~~3~~
SR.B.6.b	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.c	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.d	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.e	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.f	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.g	0	1	0	1	NA	0	1	0	1	2	3
SR.B.6.h	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.a	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.b	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.d	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.c	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.e (1)	0	1	0	1	NA	0	1	0	1	2	3
SR.B.7.e (2)	0	1	0	1	NA	0	1	0	1	2	3

variance (CIV)? They included a count of how many times questions were repeated and/or reworded in the interview and a score of students' inferred comfort using a one to four Likert-type scale.

Analysis

We systematically examined the data following each round of pilot interviews to inform our revision discussions. Responses to the content questions were scored for correctness. Audio and video recordings and photographs of students' work were examined together to triangulate evidence and corroborate whether our items produced anticipated behaviors from the students and (Merriam, 1998; Miles & Huberman, 1993). A research team member with extensive qualitative research expertise established the revision procedures and led the reconciliation discussions. Based on the field notes and fidelity of administration data, we discussed how each protocol functioned and perceptions of student responses and engagement. These data were particularly helpful in detecting CIV. For instance, if students required rephrasing of a specific question, that could indicate that the question wording was unclear or used unfamiliar vocabulary and revisions were needed. We examined the videos and discussed our observations; in instances that fidelity data indicated the presence of CIV, we used dialogical intersubjectivity to find consensus amongst the research team on appropriately revising the protocol (Kvale & Brinkmann, 2009). These steps guided our item revisions between pilots and before full-scale cognitive interviews. We revised the task materials, wording, and presentation when participants' responses contradicted the knowledge and skills we had intended to elicit. Examples of the revision process will be illustrated in the results section.

Results

In this section, we describe how the results of the iterative design process were used to revise items for task-based cognitive interviews that assessed children's spatial reasoning. The purpose of the full-scale cognitive interviews is to elicit evidence of students' spatial reasoning without the contamination of CIV. Specifically, we examined students' responses to iterated items to uncover evidence of children's reasoning and the items' construct relevance through improved alignment between the protocol and the subcomponent.

Improved Alignment

We examined children's correctness and evidence of their reasoning and compared those outcomes to associated subcomponents of the construct definition. For instance, one subcomponent reads: "Construct a three-dimensional object or space given at least two images of top, front, or side views" (Pinilla et al., 2020). The task within the first round of the pilot interviews (as seen in Figure 6) required children to create a block structure using 1-inch cubes based on two to three images of a single structure taken from different angles. It was observed that children consistently attempted to build multiple structures to match the multiple pictures. During the revision phase following the first round of pilot interviews, we added a scaffolding prompt to help children "see" that the photos were of the same building. However, during the second round of pilot interviews, children demonstrated an inability to complete the task as anticipated, regardless of how interviewers rephrased the questions. These responses did not align with the skill we intended to elicit, and we found that a material change was needed to enhance the alignment of the task-based item and the targeted subcomponent.

In the revision phase following the second round of pilot interviews, the team reviewed field notes, photos of student work, and the subcomponent itself. We updated the protocol so that children created a farm scene using Duplo® blocks instead of building a single structure. You can see in the fidelity form in Figure 5, that the questions had to be repeated and the student didn't seem completely comfortable with the task.

In the third round of pilot interviews, we presented children with two to three pictures of the scene situated on top of a grid that they used to inform their construction of a single, three-dimensional scene. Interview responses resulting from this revision were better aligned to the subcomponent as children demonstrated an understanding that the photos taken from multiple perspectives all represented a single scene of objects situated in the same places. For instance, one student stated, “"the side helped me with the barn, and then the top one helped me with the cat,”" as they pointed to each photo and the corresponding item in the scene they built. In this example, we see that the design of the scene elicited the students' reasoning around perspective-taking, a skill which typically begins developing around age 5 (Frick et al., 2014) which was the focus of the subcomponent. See Figure 6 for an example item traced through these rounds of iteration.

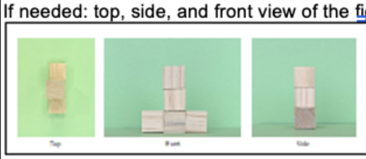
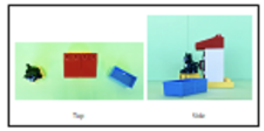
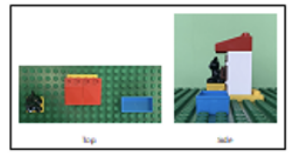
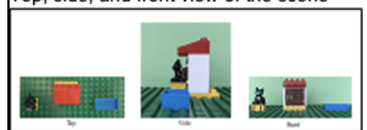
Enhanced Construct Relevance

To maximize construct relevant responses, we examined the fidelity of administration data to ascertain if children understood the questions and were comfortable engaging with the materials and tasks. To illustrate how these data informed revisions related to CIV, we will walk through an example in Figure 7.

The subcomponent reads: "Compose a two-dimensional composite figure or a three-dimensional composite figure in more than one way (e.g., a hexagon can be composed of two trapezoids or six triangles)" (Pinilla et al., 2020). The development of shape composition skills has been thoroughly investigated (Clements et al., 2004; Verdine, 2017), and the protocol for two- and three-dimensional skills intended to align with existing developmental progressions (Sarama & Clements, 2009). The protocol required children to replicate a stimulus figure in multiple ways using provided blocks. In the early item, the stimulus figure was built with three blocks of multiple colors put together in a specific way. Reflection on the data sources, including the fidelity of administration data, field notes, and item scoring, indicated that although students appeared comfortable with the task, they weren't performing as expected. In the pilot interview round 1, participants replicated the exact construction of the stimulus figure based on color, not providing alternative ways to construct the figure. It was hypothesized that the stimulus figure was likely contributing to CIV, as the construction of the figure with different colors was distracting and limited student thinking by potentially engaging their working memory over their spatial reasoning skills (see Just & Carpenter, 1985).

To reduce CIV, the next iteration of the stimulus figure was built using a single color. However, students could still see how it was constructed (separation between the blocks) and continued to replicate only

Figure 6. Early and Iterated Versions of Interview Protocol

	Pilot Interview Task Round 1	Pilot Interview Task Round 2	Pilot Interview Task Round 3
Materials	1" cube blocks (more than required to build the shape) Top and side view of the figure  If needed: top, side, and front view of the figure 	Farm scene from the previous item; toy camera; farm set pieces Top and side view of scene  If needed: top, side, and front view of the scene 	Farm set pieces & Duplo® base Top and side view of scene  Top, side, and front view of the scene 
Interview script	*Here are some pictures of a block tower from different perspectives.* *Use these blocks to create the block tower that you see in these pictures.*	*Here are some pictures of a farm from different perspectives.* *Use these Duplo® blocks to create the farm scene that you see in these pictures.* [Offer front view if needed] If child successfully builds the farm, move to task two. If the child is unsuccessful, end test after task one.	*Here are some pictures of a farm from the top and the side.* *Use these Duplo® blocks to create the farm scene that you see in these pictures.* Before providing third view: *Are you finished building?* *I'm going to show you one more picture. Does looking at this picture make you want to change anything?* [allow child to work if yes].
Scaffolding questions	*How are you starting to think about this question?* *What can you do first to build the tower?*	*What do you need to start using to build first?* *What can you construct next?*	*What can you do first to build the farm?* *What could you add next?*
Reasoning probes	*How did you know to put the blocks in this shape?* **How did you use these pictures to help you build the tower?*	*How did you know to put* [select an object] *facing that direction?*	*How did you use these pictures to help you build the farm?*
Scoring	**DURING:** After the child has completed the interview, take a picture of their tower. **AFTER:** 1 if true; 0 if not true: ____ Total number of blocks is 5 ____ Number of vertical levels 3 ____ Number of blocks in base 3 ____ Number of blocks in mid-level 1 ____ Number of blocks in top level 1 ____Alignment of base and middle Alignment of middle and top	**Scoring not established for this round.**	**DURING:** Did the child make any changes after front view was shown? 1 for yes, 0 for no: _____ **AFTER:** Using student work photo 1 if true; 0 if not true: ____ Cat is on the hay ____ Cat/hay are L of barn ____ Trough R of barn ____ Barn between Cat/hay and trough ____ Barn aligned on base ____ Barn roof aligned with eave over door **Flag for review:** __

Note. Protocol for between objects subcomponent skill: construct a three-dimensional object or space given at least two images of top, front, or side views that configuration. In the final version of the item, a 3D printed stimulus figure with no lines showing construction was provided, removing the construct irrelevant barrier. This final revision led to students building the figure in multiple ways and explaining their reasoning more clearly.

Discussion

The primary purpose of this manuscript was to illustrate the role of an iterative design process in ensuring that data gathered from cognitive interviews is trustworthy and meaningful for their intended uses. We found that engaging in the iterative design process helped us develop a cognitive interview protocol that was better aligned than the initial protocol to specific subcomponents of the learning progression. Moreover, through the process of making incremental improvements, we were able to reduce sources of CIV to improve the data that would ultimately be generated from the full-scale cognitive interviews. Further, we found that collecting multiple data sources, including audio and video recordings, still photos of completed student work products, general field notes, and fidelity of administration data, during pilot interviews was essential for refining the task-based items, as they allowed us to discuss the irregularities in fidelity data and formulate well-reasoned revisions quickly.

While this paper describes the iterative process of developing task-based cognitive interview items, the purpose of the cognitive interviews was to support the empirical recovery of the spatial reasoning learning progression. Because cognitive interviews are intended to help specify the underlying construct, their role in learning progression recovery can be significant. If the data generated from cognitive interviews are not reflective of students' construct relevant understanding, learning progression recovery and, subsequently, future assessment or instruction development efforts may be compromised. When the cognitive interview protocol is tested and refined iteratively to align with the intended constructs and minimize sources of CIV, we can have confidence in the evidence they contribute toward refining our learning progression.

Limitations

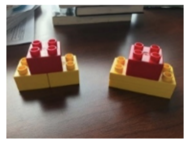
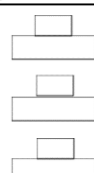
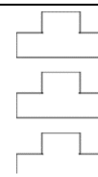
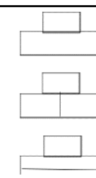
Although the findings from our research are not intended to be generalizable, the process of iteratively developing cognitive interview protocols is applicable to other settings and/or populations. For example, when conducting cognitive interviews for constructs that involve multiple dimensions of students' thinking, such as cross-cutting science principles, pilot testing may help verify the alignment with the construct. Relatedly, our population of young children led us to question whether the tasks would be susceptible to sources of CIV. When working with other populations for which CIV may be present (e.g., students with disabilities), pilot testing may be needed. However, two limitations may impact the direct applicability of this process. First, data gathered during the piloting phases were analyzed using a rapid cycle process in which the researchers and observers participating in the piloting were noticing and observing the participants' interactions with the tasks and determining which aspects of the cognitive interview protocol elicited construct relevant and construct irrelevant processes. Because these researchers and observers were familiar with the tasks and the spatial reasoning learning progression, their observations informed future iterations. However, other observers may have noted different patterns in students' responding behaviors or patterns, thereby limiting the replicability of the findings. Second, the iterative design process described in this manuscript was applied with a small sample size. As such, we were able to synthesize the findings and design the finalized cognitive interview protocols within three cycles of revisions. If the sample were larger, additional observations about the alignment with the intended construct and/or sources of CIV may have emerged, thereby necessitating additional cycles of revisions. As such, it should not be inferred from this research that all protocols can be finalized within two cycles.

Practical and Research Implications

Since the goal of cognitive interviews is to understand response processing (Padilla & Leighton, 2017), our protocol was specifically designed to elicit content and reasoning responses aligned to each subcomponent of our hypothesized learning progression. In doing so, the descriptive evidence gathered allowed us to find interconnections among responses between skills within the learning progression.

Using the iteratively developed protocol, we subsequently conducted cognitive interviews. The data

Figure 7. Early and Iterated Versions of Interview Protocol

	Pilot Interview Task Round 1	Pilot Interview Task Round 2	Pilot Interview Task Round 3
Materials	Stimulus figure that is glued together; blocks to create the stimulus figure (more than required to build the figure) 	Stimulus figure that is glued together; blocks to create the stimulus figure (more than required to build the figure) MUST INCLUDE: • 5 thick 2x2 • 1 thick 2x4 • 2 thin 2x4 	Stimulus figure that is glued together; blocks to create the stimulus figure (more than required to build the figure) MUST INCLUDE: • 5 thick 2x2 • 1 thick 2x4 • 2 thin 2x4 
Interview script	*We're going to build figures with these Duplo® blocks.* *Do the best that you can to build these figures.* *Use these Duplo® blocks* [point to the right-hand side of the mat], *to build this figure.* [point to the left-hand side of the mat]	*We're going to build figures with these Duplo® blocks.* *Do the best that you can to build this figure.* *Use these Duplo® blocks* [point to the blocks], *to build this same figure* [point to the solid figure].	*We're going to build figures with these Duplo® blocks.* *Do the best that you can to build this figure.* *Use these Duplo® blocks* [point to the blocks], *to build this same figure* [point to the solid figure] as many times as you can. If this child only builds one figure, ask them*: Can you build the figure another way?*
Scaffolding questions	If the child does not begin to construct a figure, let child touch figure without picking it up (i.e., allow child to rotate figure on table). If the child is still unable to construct a figure, let the child pick up the figure.	If the child does not begin to construct a figure, let child touch figure without picking it up (i.e., allow child to rotate figure on table). If the child is still unable to construct a figure, let the child pick up the figure.	If the child does not begin to construct a figure, let child touch figure without picking it up (i.e., allow child to rotate figure on table). If the child is still unable to construct a figure, let the child pick up the figure.
Reasoning probes	*How did you know that this block* [point to the top piece on one student-built figure] *should go on top of the other?* *If you had even more blocks, could you create some ways to make the figure again?*	*How did you know that this block* [point to the top piece on one student-built figure] *should go on top of the other?*	*How did you know that this block* [point to the top piece on one student-built figure] *should go on top of the other?*
Scoring	**DURING:** Draw lines to show Duplo® blocks used by child to build the figure: 	**DURING:** Draw lines to show Duplo® Blocks used by child to build the figure: 	**DURING:** Circle the Duplo® block Combinations the child uses. Duplicate combos only count once: 
	Circle the compositions the child used: • 1 thick 2x2, 1 thick 2x4 • 3 thick 2x2s • 2 thin 2x4s, 1 thick 2x2 Did the child create a minimum of 2 compositions to create the figure? 1 for yes, 0 for no.	Take a picture of their structures. **AFTER:** Circle the compositions the child used: • 1 thick 2x2, 1 thick 2x4 • 3 thick 2x2s • 2 thin 2x4s, 1 thick 2x2 Did the child create a minimum of 2 compositions to create the figure? 1 for yes, 0 for no.	Take a picture of each shape for each view. **AFTER:** Circle the compositions the child used: • 1 thick 2x2, 1 thick 2x4 • 3 thick 2x2s • 2 thin 2x4s, 1 thick 2x2 Did the child create a minimum of 2 compositions to create the figure? 1 for yes, 0 for no. **Flag for review:**

Note. Protocol for within objects subcomponent skill: compose a two-dimensional composite figure or a three-dimensional composite figure in more than one way (e.g., a hexagon can be composed of two trapezoids or six triangles) collected in these cognitive interviews would then be used to validate the spatial reasoning learning progression with evidence that the items were well aligned to the subcomponents and that CIV was minimized to the extent possible. For more information about the full-scale cognitive interviews facilitating the empirical recovery of the learning progression, see Pinilla et al., 2020.

Because this study details how we iteratively developed a cognitive interview protocol to capture evidence of children's spatial reasoning, the presented procedures are replicable and could benefit others engaging in similar work to refine cognitive interview protocols based on recursive feedback. This work is significant because spatial reasoning skills are related to flexible thinking, which supports numeric thinking and overall mathematical and academic achievement (Bailey, 2017; Verdine et al., 2017; Wai et al., 2009). By employing high quality research methods to recover the learning progression and subsequently develop items as the foundation for a formative assessment suite, this paper illustrates best practices for how researchers can engage in a rigorous cognitive interview design process.

References

Alonzo, A. C. & Steedle, J. T. (2008). Developing and assessing a force and motion learning progression. *Science Education, 93*, 389-421. https://doi.org/10.1002/sce.20303

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Psychological Association.

Bailey, D. H. (2017). Causal inference and the spatial-math link in early childhood. *Monographs of the Society for Research in Child Development*, *82*(1), 127-136. https://doi.org/10.1111/mono.12288

Bruce, C., Moss, J. & Ross, J. (2012). *Survey of JK to grade 2 teachers in Ontario Canada: Report to the literacy and numeracy secretariat of the ministry of education.* http://www.pme38.com/wp-content/uploads/2014/05/RF-Sinclair-et-al.pdf

Bruce, C. D., Davis, B., Sinclair, N., McGarvey, L., Hallowell, D., Drefs, M., Francis, K., Hawes, Z., Moss, J., Mulligan, J., Okamoto, Y., Whiteley, W., & Woolcott, G. (2017). Understanding gaps in research networks: Using "spatial reasoning" as a window into the importance of networked educational research. *Educational Studies in Mathematics, 95*, 143–161. https://doi.org/10.1007/s10649-016-9743-2

Clements, D. H. (2004). Geometric and spatial thinking in early childhood education. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 267–297). Lawrence Erlbaum Associates.

Clements, D. & Sarama, J. (2011). Early childhood teacher education: the case of geometry. *Journal of Mathematics Teacher Education*, *14*(2), 133–148. https://doi.org/10.1007/s10857-011-9173-0

Corbin, J., & Strauss, A. (2015). Basics of qualitative research: Techniques and procedures for developing grounded theory (4th ed.). Sage.

Creswell, J. W., & Poth, C. (2018). Qualitative inquiry and research design, (4th ed.). Sage.

Davis, B., & Spatial Reasoning Study Group. (Eds.). (2015). *Spatial reasoning in the early years: Principles, assertions, and speculations*. Routledge.

Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education, 47*(2), 123-182. https://doi.org/10.1080/03057267.2011.604476

Ericsson, A., & Simon, H. (1993). *Protocol analysis: Verbal reports as data.* The MIT Press. https://doi.org/10.7551/mitpress/5657.001.0001

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of perspective-taking abilities in 4-to 8-year-olds. *Frontiers in Psychology*, *5*, 386. https://doi.org/10.3389/fpsyg.2014.00386

Gilligan-Lee, K. A., Hawes, Z. C. K., & Mix, K. S. (2022). Spatial thinking as the missing piece in mathematics curricula. *npj Science of Learning,* 7, Article 10. https://doi.org/10.1038/s41539-022-00128-9

Graf, E. A., & van Rijn, P. W. (2015). Learning progressions as a guide for design: Recommendations based on observations from a mathematics assessment. In *Handbook of test development* (pp. 181-205). Routledge.

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *Handbook of visuospatial thinking* (pp. 121–169). Cambridge University Press. https://doi/10.1017/CBO9780511610448.005

Ketterlin-Geller, L. R., Yovanoff, P., Jung, E., Liu, K., & Geller, J. (2013). Construct Definition Using Cognitively Based Evidence: A Framework for Practice. *Educational Assessment, 18*(2), 122-146. https://doi.org/10.1080/10627197.2013.790207

Ketterlin-Geller, L. R., Zannou, Y., Sparks, A., & Perry, L. (2020). Empirical recovery of learning progressions through the lens of educators. *The Journal of Mathematical Behavior, 60*. https://doi.org/10.1016/j.jmathb.2020.100805

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Sage.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 1479-1498. https://doi.org/10.2307/1130467

Moss, J., Hawes, Z., Naqvi, S., & Caswell, B. (2015). Adapting Japanese Lesson Study to enhance the teaching and learning of geometry and spatial reasoning in early years classrooms: a case study. *ZDM*, *47*(3), 377-390. https://doi.org/10.1007/s11858-015-0679-2

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics: An overview*.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. National Academies Press

National Research Council. (2006). *Learning to think spatially*. National Academies Press. https://doi.org/10.17226/11019

National Research Council (NRC). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. C. T. Cross, T. A. Woods, & H. Schweingruber (Eds.), Committee on Early Childhood Mathematics, Center for Education, Division of Behavioral and Social Sciences and Education. National Academies Press.

Padilla, J. L., & Leighton, J. P. (2017). Cognitive interviewing and think aloud methods. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 211-228). Springer, Cham. https://doi.org/10.1007/978-3-319-56129-5_12

Pinilla, R., Perry, L., Hatfield, C., Sparks, A., Ketterlin-Geller, L. (2020). Spatial reasoning: Cognitive interview protocol development (Technical Report No. 20-07) [Technical Report]. Southern Methodist University, Research in Mathematics Education.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352. https://doi.org/10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2017). Discussion and implications: How early spatial predict later spatial and mathematical skills. *Monographs of the Society for Research in Child Development*, *82*(1), 80-109. https://doi.org/10.1111/mono.12285

Zhang, X., Koponen, T., Räsänen, P., Lerkannen, M. K., & Nurmi, J. E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development, 85*(3), 1091-1107. https://doi.org/10.1111/cdev.12173

Citation:
Thomas, E. R., Pinilla, R. K., Ketterlin-Geller, L. R., & Hatfield, C. (2023). Iterative cognitive interview design to uncover children's spatial reasoning. *Practical Assessment, Research, & Evaluation*, 28(12). Available online: https://scholarworks.umass.edu/pare/vol28/iss1/12/

Corresponding Author:
Elizabeth R. Thomas
Stephens College
Email: ethomas [at] stephens.edu