SMU | ANNETTE CALDWELL SIMMONS SCHOOL OF EDUCATION & HUMAN DEVELOPMENT

RESEARCH IN MATHEMATICS EDUCATION

Spatial Reasoning: Home Environment Phase 2 Distribution and Analysis

Technical Report 21-05

Spatial Reasoning: Home Environment Distribution and Analysis

Anthony Sparks • Sarah Wellberg • Josh Geller • Sydney Holder • Cassandra Hatfield • Leanne Ketterlin-Geller

Southern Methodist University

Spring 2021

Published by
Southern Methodist University
Department of Education Policy & Leadership
Simmons School of Education & Human Development
PO Box 750114
Dallas, TX 75275-0114
Contact information: rme@smu.edu

This research was supported by The National Science Foundation (Grant #1721100). Opinions expressed herein do not necessarily reflect those of The National Science Foundation or individuals within.

Abstract

This report describes the distribution and analysis efforts for the Spatial Reasoning Home Environment Survey. The survey is intended for parents of students in grades K-2 and measures at-home spatial reasoning activities. We distributed the survey virtually through Qualtrics and retained 201 observations. We conducted exploratory factor analyses and item response theory analyses. The analyses revealed a two-factor instrument. Implications for research and next steps are discussed.

Table of Contents

Spatial Reasoning: Home Environment Phase 2 Distribution and Analysis

Introduction

The purpose of this technical report is to describe the distribution and analyses of the redevelopment of the Spatial Reasoning Home Environment Questionnaire. The first iteration of the survey is described in the Spatial Reasoning Home Environment Questionnaire Phase 1 technical report (Sparks et al., 2020). The redevelopment of the home environment questionnaire is described in the Spatial Reasoning Home Environment Questionnaire Phase 2 Development technical report (Wellberg et al., 2020).

Distribution

To meet the goals of the analysis, we targeted distribution of the survey to 200 parents. To ensure national distribution of the survey, it was distributed through two different groups of contacts. Prior to receiving the actual survey, a short eligibility questionnaire was sent to both groups to identify if potential participants met a minimum criterion and to screen out computer bots. Eligible participants were then selected at random and sent a link to the Spatial Reasoning Home Use Questionnaire. Once a respondent finished the survey, they were offered a $10 gift card and, if they agreed to submit their name and email address, were immediately sent the gift card via an automated rewards service. Further description of the distribution details is provided below.

Contact Groups

Educators subscribing to RME's newsletter listserv comprised the first contacts group to which the Eligibility Questionnaire was sent. These educators include teachers, administrators, coaches, and researchers, many of whom are former attendees of RME's annual research-to-practice conference or have worked with RME on past research projects. A flier was sent to all members of the listserv (nearly 3,000 people) including a brief description of the survey, its importance, and mention of the $10 gift card incentive. Parents of children in kindergarten through second grade were encouraged to follow the link to the questionnaire and/or forward the link to parents they may know.

Twitter was also used for distribution of the survey. RME has over 1200 Twitter followers for which an electronic flier was posted similar to the one sent through the newsletter listserv. Parents of children in kindergarten through grade two were encouraged to follow the link to the questionnaire, retweet, or share the link with other parents.

Eligibility Questionnaire

The reasons for designing an eligibility questionnaire as an entry point for potential participants were three-fold. First, since the survey offered a gift card incentive, and that incentive was

automatically sent to those successfully completing the survey, researchers wanted to be sure that participants were authentic. Additionally, based on prior survey distribution experience the researchers had evidence to suggest that computer generated "bots" could gain access to the survey and game the system for financial gain. Finally, the survey was being distributed through Qualtrics, which has the ability to track partial and complete responses and to send automated reminders to participants who have not yet begun or had only partially completed the survey. By using an eligibility questionnaire, researchers were able to gather email addresses and utilize the tracking functionality within Qualtrics.

The eligibility questionnaire, built using Qualtrics online survey software, contained seven questions including a "reCAPTCHA" verification against robots, an acknowledgement of informed consent, two questions used to determine whether parents had a child in kindergarten through second grade, and two spaces in which to enter and to verify an email address. (See Appendix A for the full eligibility survey). Ineligibility was based on these criteria:

- Both emails were not an exact match
- Selecting "no" to the question "Do you have one or more children who are in the grade of kindergarten through second grade during the current 2020-2021 school year?"
- Indicating their child was born outside of the years range 2012-2015. This date of birth year range was used as a second check for confirming parents had a child in grades kindergarten through two.

The Eligibility Questionnaire amassed over 3,000 respondents in the twelve days it was open. Qualtrics has a feature that describes the source of a respondent's access to the survey as either "social," referring to a social media link (Twitter, in this case), or "anonymous" which refers to access via a link sent in an email. Approximately 85% of respondents were from Twitter and 15% from the newsletter listserv.

Upon reviewing the robust data Qualtrics offers for each response, researchers created the final list of eligible participants by removing ineligible participants based on the above criteria, removing duplicate emails (most from Twitter), duplicate IP addresses, along with, duplicate longitude and latitude entries. After these removals, 752 respondents remained as eligible to take the actual Home Environment Survey.

Distribution of the Home Environment Survey

Of the final eligible respondents, a random number was generated for each respondent and 250 respondents were randomly selected from which to send the home environment survey link. Researchers retained the 85% social, 15% anonymous proportion for the selection. Although our goal was to get 200 respondents, we oversampled by starting with 250 because we didn't expect to receive responses from everyone.

An email was sent via the Qualtrics software using Research in Mathematics Education (RME) as the sender inviting participants to follow the survey link, take the survey, and submit their name and email address to receive the gift card. Specific language was used to reduce frustration

if gift card distribution fell short: "…only the first 200 participants will receive the gift card," and "If you don't receive an email with a link to the gift card, then all gift cards were claimed."

Researchers designed the Qualtrics link to be valid for 5 days, after 3 days, researchers sent a reminder email to those who had not yet completed the survey. After the reminder was sent, and a few more respondents completed the survey but the target goal of 200 respondents had not yet been met. Researchers selected another random sample of an additional 70 people from the Eligibility Questionnaire and pushed the survey out again. At closing of the survey, there were 273 responses.

Incentive

Researchers offered $10.00 gift cards for those who completed the survey. The gift cards were distributed through Rewards Genius, an online incentive coordination platform that liaises with Qualtrics and automatically sends gift cards to those you specify as completing specific questions within Qualtrics. A total of 250 gift cards were purchased and distributed, and there were 273 responses to the survey.

Data Cleaning

We downloaded a raw data file from the survey data collection software, Qualtrics®. Our sample size started at 273 observations. We first removed individuals who were either RME staff or representatives of Qualtrics®, and we removed responses we already identified as duplicates before data collection had completed. Lastly, we removed partially complete responses from the dataset. These steps brought the number of observations down to 265.

We next checked multiple variables for duplicates. The first variable was IP Addresses of the respondents. We found five instances where an IP address repeated twice. Removing these ten observations brought the sample size down to 255. We next checked for duplicates on the email address in which we sent respondents. We found no duplicates on the sent email. At the end of the survey, respondents were prompted to enter an email to qualify for the $10 incentive. We analyzed the email addresses that respondents entered and found one instance of a duplicate. These two email addresses were removed, and the sample size dropped to 253.

Next, we analyzed similar response patterns to further identify duplicate responses. After identifying similar response patterns, we analyzed their location and responses to other questions on the survey. We found two instances of duplicate response. One response occurred four times, while another response occurred three times. We decided to keep the first response in each and discard the remaining responses. Furthermore, we identified two additional duplicates when comparing the emails to the Rewards Genius® emails. These steps brought the sample size to 238. We also checked the IP addresses of the responses to ensure that the responses all came from inside the United States. No responses were recorded outside of the US. Lastly, we identified unusual responses based on the average adult words read per minute. Based on the number of words in the survey, the quickest an average adult could have read the survey is within three minutes. Therefore, we decided to remove observations that took less than three minutes to respond to the survey. This brought the number of observations down to 201.

Analysis

Two research associates and an undergraduate student worker collaborated to conduct the analyses. The data were first cleaned as described above. We began by describing the demographics and the academic characteristics of the parents and of the students of the parents that were surveyed. Then, we analyzed the response distribution of four-point scale items in which the parents were asked whether their child could perform various activities. These items had the response options “Not Yet”, “Yes, with a lot of help”, “Yes, with a little bit of help”, and “Yes, without help”. Last, we analyzed how frequently the students performed tasks using the five-point scale items with response options of “Never”, “1-2 times per year”, “1-2 times per month”, “1-2 times per week”, and “Almost daily”. The results are described in the next section.

To better understand the structure of the instrument, we conducted a series of exploratory factor analyses. We accounted for the polytomous nature of the data by using a polychoric correlation matrix. We then used theoretically based approaches to better understand the appropriate number of factors and item loadings.

To better understand the items on the instrument, we conducted a series of item response theory (IRT) analyses. (IRT) allows us to gain insight into additional features of the items (e.g., discrimination, difficulty). We used the factor structure from the results of the exploratory factor analyses to run multidimensional IRT analyses beginning with a graded-response model followed by a partial credit model. In addition, we ran item and person fit statistics to assess the model fit. Items were removed as necessary.

Results

In this section, we describe the results of the analyses.

Descriptive Statistics and Recoding Decisions

Table 1 describes the demographics and academic characteristics of the K-2nd students of the parents that were surveyed.

Table 1

Student Characteristics of Parent Surveyed

Characteristic		Count (%)
Grade		
	K	43 (21%)
	1	107 (53%)
	2	51 (25%)
Gender		
	Male	121 (60%)
	Female	79 (39%)
	Gender non-binary	0 (0%)

	Prefer not to answer	1 (1%)
Race		
	Asian	2 (1%)
	American Indian/Alaskan Native	10 (5%)
	Black/African American	35 (17%)
	Native Hawaiian/Other Pacific Islander	3 (2%)
	White	142 (71%)
	Two or More Races	6 (3%)
	Prefer not to answer	3 (2%)
Ethnicity		
	Hispanic	63 (31%)
	Non-Hispanic	132 (66%)
	Prefer not to answer	6 (3%)
Home Languages		
	English	198 (98%)
	Spanish	3 (2%)
School Type		
	Public	132 (66%)
	Private – Non-religious	24 (12%)
	Private – Religious	38 (19%)
	Charter	6 (3%)
	Other: Homeschool	1 (1%)
Bilingual enrollment		
	Yes	160 (80%)
	No	41 (20%)
Pre-K		
	Yes	179 (89%)
	No	21 (10%)
	Prefer not to answer	1 (1%)
After-school program		
	Yes	151 (75%)
	No	50 (25%)
504/IEP		
	Yes	110 (55%)
	No	86 (43%)
	Prefer not to answer	5 (3%)
Internet Access		
	Home	140 (70%)
	School	42 (21%)
	Community hotspot	10 (5%)
	Cellular device	9 (5%)
	Other	0 (0%)

Table 2 describes the gender and level of education of the K-2nd parents that were surveyed.

Table 2

Parent Characteristics

Characteristic		Count (%)
Gender		
	Male	102 (51%)
	Female	99 (49%)
	Gender non-binary	0 (0%)
	Prefer not to answer	0 (0%)
Level of Education		
	Some high school	0 (0%)
	High school diploma	10 (5%)
	Vocational certification	18 (9%)
	Associates degree	76 (38%)
	Undergraduate degree	79 (39%)
	Graduate degree	18 (9%)

Table 3 describes the response distribution of four-point scale items. The parents were asked if their child engaged in various activities and were given the response options “Not Yet”, “Yes, with a lot of help”, “Yes, with a little bit of help”, and “Yes, without help”.

Table 3

Home Environment Survey Response Distribution for Four-Point Scale Items

Does your child do any of the following activities on their own?				
	Not Yet	Yes, with a lot of help	Yes, with a little bit of help	Yes, without help
Papercraft	4 (2%)	77 (38%)	78 (39%)	42 (21%)
Draw maps	15 (8%)	51 (25%)	81 (40%)	54 (27%)
Draw plans for building of spaces	16 (8%)	84 (42%)	73 (36%)	28 (14%)
Draw pictures from a bird’s-eye view	28 (14%)	55 (27%)	93 (46%)	25 (12%)
If prompted, does your child do any of the following?				
	Not Yet	Yes, with a lot of help	Yes, with a little bit of help	Yes, without help
Recognize that shapes have the same name	2 (1%)	78 (39%)	76 (38%)	45 (22%)

even when they are facing different ways our are different sizes				
Identify that two or more objects are the same shape even if they have different sizes or orientations	8 (4%)	60 (30%)	95 (47%)	38 (19%)
Recognize a photo of an object or a location taken from a different point of view	13 (7%)	70 (35%)	78 (39%)	40 (20%)
Find his/her location on a map with a grid	26 (13%)	84 (42%)	72 (36%)	19 (10%)
Notice the shape or an object's flat face after it has been cut into parts or sliced in half	11 (6%)	66 (33%)	86 (43%)	38 (19%)
Associate or draw real-world objects as shapes	13 (7%)	58 (29%)	95 (47%)	35 (17%)

Table 4 describes the response distribution of five-point scale items. The parents were asked about how frequent their child does certain activities and given choices "Never", "1-2 times per year", "1-2 times per month", "1-2 times per week", and "Almost daily".

Table 4

Home Environment Survey Response Distribution for Five-Point Scale Items

About how often does your child use a computer, video game, phone, or tablet application to do the following activities?					
	Never	1-2 times per year	1-2 times per month	1-2 times per week	Almost daily
Build things	1 (1%)	16 (11%)	49 (35%)	62 (44%)	13 (9%)
Organize or arrange shapes to match or fit a space	1 (1%)	27 (16%)	65 (39%)	50 (30%)	24 (14%)
Move a digital avatar through space	1 (1%)	11 (9%)	44 (36%)	42 (35%)	23 (19%)
Navigate or move through virtual spaces using a map	1 (1%)	20 (16%)	37 (30%)	53 (43%)	13 (11%)
About how often does your child play with the following item/toys?					
	Never	1-2 times per year	1-2 times per month	1-2 times per week	Almost daily

Jigsaw Puzzles	1 (1%)	25 (12%)	70 (35%)	71 (35%)	34 (17%)
Blocks	2 (1%)	31 (15%)	41 (20%)	79 (39%)	48 (24%)
Board games in which they move a player through a route with other players	5 (3%)	19 (10%)	73 (36%)	84 (42%)	20 (10%)
How often have you noticed your child doing the following?					
	Never	1-2 times per year	1-2 times per month	1-2 times per week	Almost daily
Describe the features of a figure	4 (2%)	22 (11%)	82 (41%)	70 (35%)	23 (11%)
Using landmarks or specific places to describe locations	5 (3%)	40 (20%)	65 (32%)	58 (29%)	33 (16%)
Describing an object's position relative to other objects	4 (2%)	17 (9%)	74 (37%)	76 (38%)	30 (15%)
Using hand motions while they are describing an object's position	1 (1%)	29 (14%)	55 (27%)	82 (41%)	34 (17%)
About how often do you (or someone in your household) do the following with your child?					
	Never	1-2 times per year	1-2 times per month	1-2 times per week	Almost daily
Build things with your child by following a set of written, illustrated, or oral instructions	1 (1%)	22 (11%)	64 (32%)	79 (39%)	35 (17%)
Use hand motions or other movements when describing an object's position	3 (2%)	34 (17%)	60 (30%)	72 (36%)	32 (16%)
Ask your child to place or retrieve an object using positional language	7 (4%)	23 (11%)	58 (29%)	74 (37%)	39 (19%)
Ask your child why a shape has a certain name	7 (4%)	19 (10%)	71 (35%)	77 (38%)	27 (13%)
Ask your child to identify an object that represents a certain shape	8 (4%)	28 (14%)	60 (30%)	73 (36%)	32 (16%)

Given the very small percentages of parent respondents who reported that their child never engaged in the activities described in these items, we decided to collapse the "Never" and the "1-2 times per year" response categories. Furthermore, since respondents were only directed to answer the Digital Experience items if they reported that their child had access to the associated technology, there was quite a bit of missingness for those items. As a result of this missingness, we only had 68 complete observations out of the 201 respondents. For this reason, we decided to recode any missing responses to these four items as the collapsed "Never/1-2 times per year" category. We felt justified in doing this because it is likely that if a child does not have access to certain technology, they never engage in the activity.

Exploratory Factor Analysis (EFA)

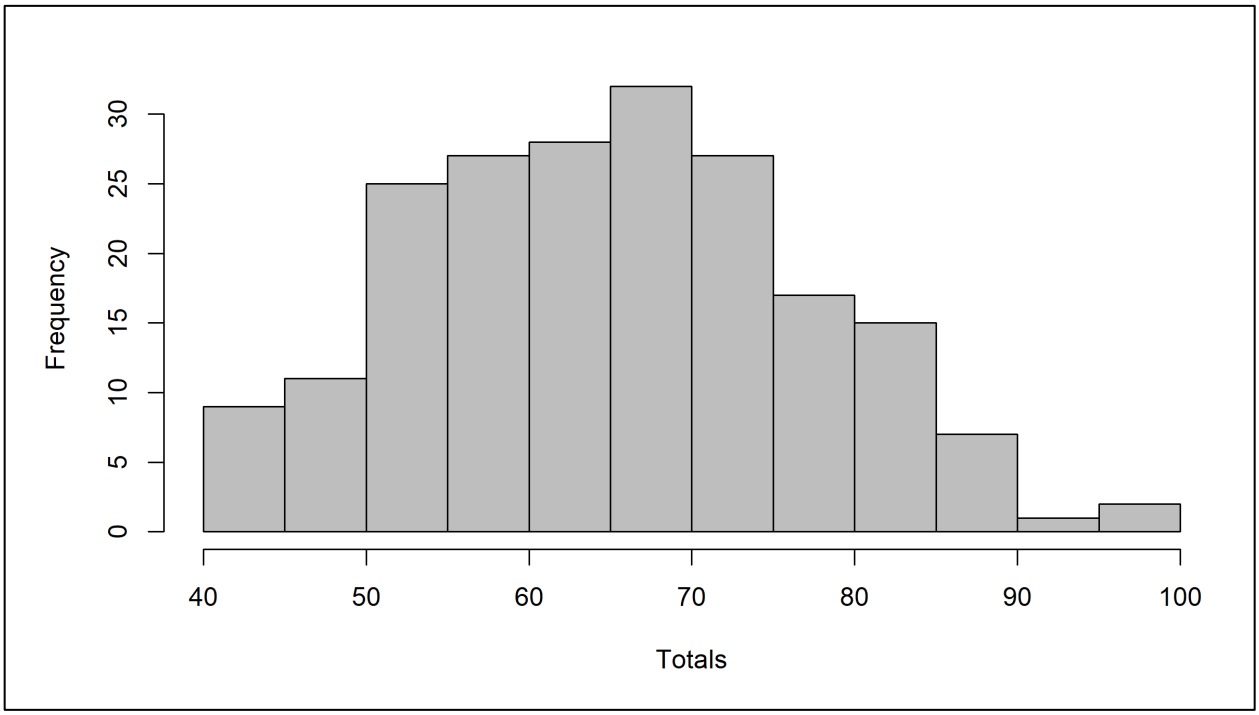
In this section, we describe the exploratory factor analysis.

Descriptive Analyses

After making our recoding decisions, we attempted to determine the survey's factor structure using an exploratory factor analysis (EFA). To obtain valid information from an EFA, certain assumptions about the data must be met (Tabachnick & Fidell, 2001). First, our sample size of 201 respondents and 26 items is sufficient to estimate loadings for the two factors that we expect (within- and between- shape reasoning) and could support six factors if each item block represented its own factor. In looking at the distribution of total scores (where independence items are coded 1-4 from "Not yet" to "Yes, without help", and frequency items are coded 1-4 from "Never/1-2times per year" to "Almost daily") in Figure 1, we can see that the totals are roughly normally distributed. Though there is a slight uptick in the interval from 95-100, there are no clear outliers among the respondents.

Figure 1

Survey Total Distribution



Another consideration is multicollinearity. We used the polychoric correlations between the items to investigate this rather than the traditional Pearson correlations because the items are all on Likert response scales, which are ordinal, not interval scales. We can see in Table 5 that almost all pairs of items have moderately positive correlations with none larger than 0.66. There are five item pairs for which we see negative correlations, but these are small. In fact, three of these are small enough that they round to 0.00. Furthermore, examination of scatterplots revealed that the items tended to have slight positive linear relationship with one another, except for the five item pairs with negative correlations in Table 5. These item pairs had very slight negative associations, which were not statistically distinguishable from 0.

Table 5

Polychoric Correlation Matrix

		Build/Draw				Identify						Digital				Play			Language				Interaction				
		paper	maps	plans	bird	shape2d	shape3d	photo	grid	slice	realshape	build	org	avatar	map	puzzle	blocks	board	figfts	landmark	relpos	hands	build	hands	poslang	shapename	idshape
Build/Draw	paper	1.00	0.26	0.26	0.06	0.51	0.28	0.20	0.11	0.37	0.23	0.24	0.33	0.31	0.15	0.34	0.24	0.30	0.29	0.27	0.20	0.17	0.41	0.30	0.22	0.28	0.36
	maps	0.26	1.00	0.14	0.11	0.42	0.30	0.21	0.11	0.32	0.28	-0.15	0.10	0.20	0.08	0.21	0.28	0.09	0.17	0.40	0.29	0.19	0.29	0.33	0.20	0.24	0.08
	plans	0.26	0.14	1.00	0.38	0.32	0.15	0.47	0.27	0.30	0.32	0.08	-0.09	0.24	0.15	0.00	0.04	0.11	0.17	0.17	0.25	0.13	0.09	0.05	0.17	0.12	0.04
	bird	0.06	0.11	0.38	1.00	0.18	0.12	0.16	0.44	0.20	0.10	0.14	0.10	0.13	0.15	0.03	0.00	0.12	0.11	0.02	0.17	0.16	0.21	0.14	0.09	0.05	0.10
Identify	shape2d	0.51	0.42	0.32	0.18	1.00	0.28	0.49	0.09	0.57	0.59	0.26	0.27	0.16	0.21	0.30	0.35	0.32	0.36	0.47	0.46	0.35	0.39	0.50	0.31	0.26	0.32
	shape3d	0.28	0.30	0.15	0.12	0.28	1.00	0.20	0.14	0.44	0.35	0.07	0.11	0.18	0.07	0.04	0.25	0.24	0.04	0.41	0.24	0.37	0.25	0.43	0.29	0.15	0.17
	photo	0.20	0.21	0.47	0.16	0.49	0.20	1.00	0.23	0.28	0.47	0.14	0.18	0.26	0.21	0.05	0.18	0.26	0.23	0.25	0.29	0.22	0.20	0.26	0.28	0.25	0.08
	grid	0.11	0.11	0.27	0.44	0.09	0.14	0.23	1.00	0.12	0.09	0.03	0.05	0.05	0.13	0.15	0.08	0.13	0.22	0.12	0.28	0.07	0.22	0.06	0.16	0.27	0.16
	slice	0.37	0.32	0.30	0.20	0.57	0.44	0.28	0.12	1.00	0.42	0.31	0.26	0.13	0.23	0.19	0.24	0.29	0.23	0.33	0.40	0.37	0.30	0.40	0.32	0.23	0.31
	realshape	0.23	0.28	0.32	0.10	0.59	0.35	0.47	0.09	0.42	1.00	0.13	0.18	0.08	0.13	0.09	0.21	0.32	0.17	0.42	0.28	0.38	0.20	0.30	0.53	0.18	0.22
Digital	build	0.24	-0.15	0.08	0.14	0.26	0.07	0.14	0.03	0.31	0.13	1.00	0.18	0.38	0.56	0.16	0.14	0.22	0.26	0.16	0.23	0.22	0.22	0.31	0.29	0.04	0.33
	org	0.33	0.10	-0.09	0.10	0.27	0.11	0.18	0.05	0.26	0.18	0.18	1.00	0.12	0.00	0.58	0.43	0.50	0.37	0.25	0.34	0.34	0.62	0.36	0.27	0.42	0.57
	avatar	0.31	0.20	0.24	0.13	0.16	0.18	0.26	0.05	0.13	0.08	0.38	0.12	1.00	0.47	0.29	0.26	0.17	0.19	0.31	0.24	0.30	0.20	0.34	0.22	0.28	0.24
	map	0.15	0.08	0.15	0.15	0.21	0.07	0.21	0.13	0.23	0.13	0.56	0.00	0.47	1.00	0.01	0.15	0.14	0.20	0.22	0.10	0.13	0.19	0.23	0.20	0.08	0.12
Play	puzzle	0.34	0.21	0.00	0.03	0.30	0.04	0.05	0.15	0.19	0.09	0.16	0.58	0.29	0.01	1.00	0.51	0.50	0.52	0.35	0.40	0.27	0.62	0.31	0.24	0.38	0.47
	blocks	0.24	0.28	0.04	0.00	0.35	0.25	0.18	0.08	0.24	0.21	0.14	0.43	0.26	0.15	0.51	1.00	0.30	0.38	0.49	0.36	0.36	0.50	0.53	0.22	0.54	0.35
	board	0.30	0.09	0.11	0.12	0.32	0.24	0.26	0.13	0.29	0.32	0.22	0.50	0.17	0.14	0.50	0.30	1.00	0.35	0.36	0.37	0.38	0.53	0.31	0.33	0.29	0.42
Language	figfts	0.29	0.17	0.17	0.11	0.36	0.04	0.23	0.22	0.23	0.17	0.26	0.37	0.19	0.20	0.52	0.38	0.35	1.00	0.25	0.41	0.31	0.58	0.26	0.23	0.40	0.45
	landmark	0.27	0.40	0.17	0.02	0.47	0.41	0.25	0.12	0.33	0.42	0.16	0.25	0.31	0.22	0.35	0.49	0.36	0.25	1.00	0.37	0.46	0.34	0.66	0.31	0.28	0.33
	relpos	0.20	0.29	0.25	0.17	0.46	0.24	0.29	0.28	0.40	0.28	0.23	0.34	0.24	0.10	0.40	0.36	0.37	0.41	0.37	1.00	0.46	0.26	0.43	0.44	0.46	0.45
	hands	0.17	0.19	0.13	0.16	0.35	0.37	0.22	0.07	0.37	0.38	0.22	0.34	0.30	0.13	0.27	0.36	0.38	0.31	0.46	0.46	1.00	0.30	0.49	0.45	0.41	0.37
Interaction	build	0.41	0.29	0.09	0.21	0.39	0.25	0.20	0.22	0.30	0.20	0.22	0.62	0.20	0.19	0.62	0.50	0.53	0.58	0.34	0.26	0.30	1.00	0.25	0.29	0.36	0.54
	hands	0.30	0.33	0.05	0.14	0.50	0.43	0.26	0.06	0.40	0.30	0.31	0.36	0.34	0.23	0.31	0.53	0.31	0.26	0.66	0.43	0.49	0.25	1.00	0.25	0.41	0.34
	poslang	0.22	0.20	0.17	0.09	0.31	0.29	0.28	0.16	0.32	0.53	0.29	0.27	0.22	0.20	0.24	0.22	0.33	0.23	0.31	0.44	0.45	0.29	0.25	1.00	0.32	0.31
	shapename	0.28	0.24	0.12	0.05	0.26	0.15	0.25	0.27	0.23	0.18	0.04	0.42	0.28	0.08	0.38	0.54	0.29	0.40	0.28	0.46	0.41	0.36	0.41	0.32	1.00	0.40
	idshape	0.36	0.08	0.04	0.10	0.32	0.17	0.08	0.16	0.31	0.22	0.33	0.57	0.24	0.12	0.47	0.35	0.42	0.45	0.33	0.45	0.37	0.54	0.34	0.31	0.40	1.00

Round 1

With the EFA assumptions adequately met, we began exploring the factor structure by creating a scree plot with parallel analysis. The parallel analysis suggested the use of seven factors, which can be seen in Figure 2. For the first round of EFA, we used the seven suggested factors with a principal axis factor extraction and polychoric correlations. Since we expected the items to all relate to spatial reasoning with likely two correlated factors having to do with within- and between-shape reasoning, we used a promax rotation in all our analyses.

Figure 2

Scree Plot

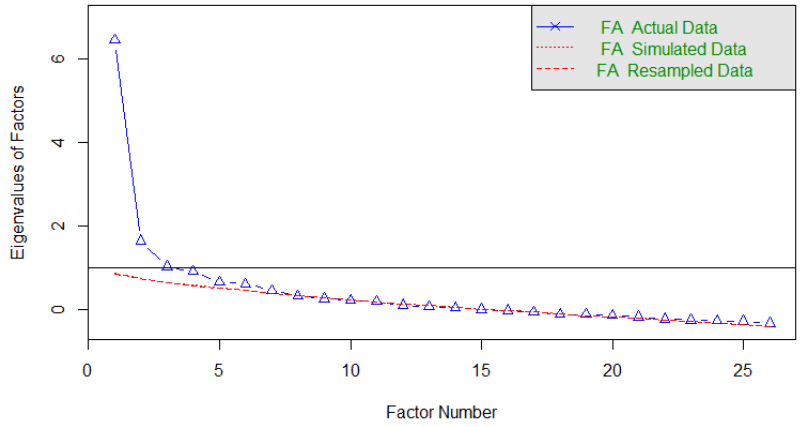


Table 6

Initial Factor Structure (Promax Rotation)

Category	Item	1	2	3	4	5	6	7
Build/Draw								
	Papercraft							
	Maps		0.518					
	Plans			0.609				
	Bird-Eye View						0.712	
Identify								
	2D Shapes			0.613				
	3D Shapes		0.695					
	Photo			0.68				
	Grid						0.61	
	Slice							
	Real Shapes			0.691		0.452		
Digital								
	Build				0.873			

	Organize	0.813				
	Avatar			0.534		
	Maps			0.704		
Play						
	Puzzles	0.822				
	Blocks					
	Board Games	0.53				
Language						
	Features of Figures	0.599				
	Landmarks		0.639			
	Relative Position					
	Hand Motions				0.496	
Interaction						
	Build	0.907				
	Hand Motions		0.77			
	Positional Language				0.527	
	Shape Names					0.534
	Shape ID	0.63				

Round 2

The results of the seven-factor EFA with promax rotation (Table 6) indicated that only three of those factors had loadings larger than 0.4 for four or more items, so we ran a new analysis using three factors (Table 7). This round resulted in two main factors and a third that loaded very heavily on three of the four items from the digital item block. We removed the three items which had loadings less than 0.3 on all three factors ("Build Papercraft", "Draw Bird-Eye View", "Location Grid") and the item that had loadings higher than 0.3 on two factors ("Relative Position"). Once these items were removed, we ran a two-factor EFA because only three items loaded on the third factor. H2 represents the communalities of the items, or the ratio of unique variance to the shared variance. Red cells represent items with weak loadings or items with low communalities (h2) that are flagged for removal. Gray cells represent items that did not load at a 0.30 threshold.

Table 7

Round 2 EFA Loadings (3 Factors)

Category	Item	1	2	3	h2
Build/Draw					
	Papercraft				0.27
	Maps		0.60		0.29
	Plans		0.50		0.28
	Bird-Eye View				0.11
Identify					

	2D Shapes		0.70		0.57
	3D Shapes		0.62		0.30
	Photo		0.57		0.34
	Grid				0.08
	Slice		0.57		0.41
	Real Shapes		0.81		0.49
Digital					
	Build			0.84	0.61
	Organize	0.85			0.58
	Avatar			0.45	0.30
	Maps			0.75	0.49
Play					
	Puzzles	0.91			0.64
	Blocks	0.60			0.44
	Board Games	0.53			0.38
Language					
	Features of Figures	0.59			0.40
	Landmarks		0.55		0.45
	Relative Position	0.34	0.37		0.41
	Hand Motions		0.40		0.38
Interaction					
	Build	0.77			0.58
	Hand Motions		0.45		0.44
	Positional Language		0.38		0.30
	Shape Names	0.53			0.37
	Shape ID	0.70			0.50

Round 3 & 4

This third round of EFA resulted in factor loadings less than 0.3 on two of the digital items and a communality of only 0.12 on the other digital item that had loaded on the third factor in the previous round (Table 8). We decided to remove these three items for the final round of analysis. This resulted in the final factor loadings seen in Table 9. We can see evidence of simple structure such that there are two main factors that explain the variance observed in these items, and each factor is loaded on by disjoint sets of items.

Table 8

Round 3 EFA Loadings (2 Factors)

Category	Item	1	2	h2
Build/Draw				
	Maps		0.42	0.19
	Plans		0.56	0.21

Identify				
	2D Shapes		0.70	0.55
	3D Shapes		0.55	0.27
	Photo		0.63	0.33
	Slice		0.61	0.41
	Real Shapes		0.77	0.49
Digital				
	Build			0.14
	Organize	0.86		0.58
	Avatar			0.18
	Maps		0.36	0.12
Play				
	Puzzles	0.92		0.63
	Blocks	0.59		0.43
	Board Games	0.55		0.38
Language				
	Features of Figures	0.62		0.38
	Landmarks		0.52	0.45
	Hand Motions		0.41	0.39
Interaction				
	Build	0.81		0.61
	Hand Motions		0.47	0.45
	Positional Language		0.45	0.30
	Shape Names	0.52		0.34
	Shape ID	0.71		0.47

Table 9

Final Factor Loadings

Category	Items	1	2	h2
Build/Draw				
	Maps		0.44	0.22
	Plans		0.53	0.20
Identify				
	2D Shapes		0.70	0.57
	3D Shapes		0.56	0.29
	Photo		0.61	0.32
	Slice		0.60	0.41
	Real Shapes		0.80	0.54
Digital				
	Organize	0.84		0.58

Play				
	Puzzles	0.90		0.62
	Blocks	0.59		0.44
	Board Games	0.55		0.38
Language				
	Features of Figures	0.62		0.38
	Landmarks		0.51	0.45
	Hand Motions		0.40	0.39
Interaction				
	Build	0.81		0.61
	Hand Motions		0.45	0.44
	Positional Language		0.43	0.29
	Shape Names	0.52		0.34
	Shape ID	0.70		0.46

We can see in Table 10 that the correlations between the two factors are moderately high at 0.55. The proportions of variance explained by each factor, however, are low, at 0.23 and 0.19 for the first and second factor, respectively (Table 11). This indicates that there may be some other sources of variance that are contributing to the differences in observed responses to the items in this survey.

Table 10

Factor Correlations

	Factor 1	Factor 2
Factor 1	1	0.55
Factor 2	0.55	1

Table 11

Variance Explained

	Factor 1	Factor 2
SS Loadings	4.33	3.60
Proportion Variance	0.23	0.19
Cumulative Variance	0.23	0.42

This survey was created to measure children's experiences with spatial reasoning in their home environment. The items were designed to align with sub-constructs that have to do with within-shape and between-shape spatial reasoning. The items that ultimately loaded on the first factor tend to be those that inquire about within-shape tasks, which have to do with knowledge of shape properties and the transformation and de/composition of shapes (Sarama & Clements, 2009b). The items that loaded on the second factor represent the between-shape subconstruct. These ask about spatial orientation and positional relationships between objects (Sarama & Clements,

2009a). It does, therefore, appear as though the two spatial reasoning sub-constructs are distinct enough to load clearly on two separate factors.

The overall reliability estimate for the 19 items that remained at the end of this process is 0.87, which is fairly high. Furthermore, the reliability estimates for the eight items representing within-shape reasoning and the 11 items representing between-shape reasoning are 0.85 and 0.80, respectively. This indicates a fairly high level of internal consistency within each factor.

Additional Analyses

Because analyses supported the removal of three of the four Digital Experiences items, we decided to try redoing this process beginning with those four items removed. This resulted in more items being filtered from the survey, though those that were left followed the same loading patterns. Because the digital organization item remained in the original analyses, we also tried running the EFA with only that item in at the start and the other three digital items removed. We ended up with results that were similar to those we had obtained with all four of the digital items removed. To keep as many items as possible, we decided to stick with the original analysis described above. After a meeting with the team to discuss the qualitative nature of the factors, we settled upon Factor 1 as a spatial visualization factor, while Factor 2 as a spatial orientation factor.

The exploratory factor analysis estimates loadings across items even when their loadings are not significant. To better understand the fit of the model without the cross loadings, we conducted a confirmatory factor analysis with the simple structure identified from the exploratory factor analysis. Some fit indices are diminished due to the number of parameters estimated (Hu & Bentler, 1999). The confirmatory factor analysis allowed us to understand the model fit without the additional parameter estimates from the exploratory factor analysis. Table 12 describes the fit indices of the two models. We note the significant improvement of the model fit when cross loadings are constrained to zero. Literature indicates better model fit when the TLI is closer to 1.0, the RMSEA is closer to 0.05, and when the upper bound of the confidence interval is less than 0.10 (Hu & Bentler, 1999).

Table 12

Fit Indices for Exploratory and Confirmatory Factor Analysis

Fit Index	Exploratory Factory Analysis	Confirmatory Factor Analysis
TLI	0.744	0.969
RMSEA	0.107	0.067
90% Confidence Interval	(0.097, 0.119)	(0.055, 0.079)

Item Response Theory (IRT)

The results from the exploratory factor analyses informed the item response theory (IRT) analyses. The purpose of the IRT analyses is to better understand features of individual items on the survey. The exploratory factor analyses uncovered a multidimensional instrument with two

distinct factors. Items from the exploratory factor analysis did not cross load across factors, leading to an instrument with simple structure. Therefore, we conducted multidimensional IRT analyses.

To model the items on the survey, we conducted a series of graded response models, which are used to model polytomous-scored items. These models provide estimates of both item discrimination and item difficulty, which is represented by the thresholds at which adjacent response levels are equally likely to be selected by a respondent with a given ability (theta) level (Samejima, 1969). Due to the four-point scale, the graded response model estimates three threshold parameters. We also modeled the partial credit model, which only estimates the difficulty parameter, and tested it against the graded response model (e.g., two-parameter) using a chi-squared test. Significant differences between the two models indicate the need to model the additional parameter.

After modeling the items with the graded response model, we then analyzed item and person fit statistics. Item fit statistics provide an indication of how well the proposed model fits with the item responses. In this analysis, we conduct the signed chi-squared test.

$$S - X^2 = \sum_{k}^{n-1} N_k \frac{(O_k - E_k)^2}{E_k(1 - E_k)}$$

Person fit refers to "the alignment between an examinee's response pattern and the IRT model selected for modeling the response data" (Desjardins & Bulut, 2018, p. 136). We report the Z_h statistics.

$$Z_h = \sum [\log L \,|\theta_j - \sum E(\log L \,|\theta_j)] / \sqrt{(\sum V(\log L \,|\theta_j))}$$

The statistic follows a standard normal distribution with an expected value of zero when the response pattern aligns with the IRT model. Large quantities of negative values (< -2) provide indication of model misfit.

When we detected significant item misfit, we dropped the item with the lowest p-value less than 0.05. This process is a variation of a methods known as backwards elimination (James et al., 2013).

Table 13 describes this process with the 19 remaining items on the survey. We describe the iterative steps below.

- In the first iteration with the full survey, we identified ID 3D Shapes with the lowest p-value. After removing ID 3D Shape item, the Digital Organize item was then identified as having the lowest p-value.

- Lastly, the Interaction – Shape Names item was identified with the lowest p-value once the previous two items had been removed. This led to a survey with 16 items. However, we noticed that one factor had six items, which was not evidence of a strong factor.

Table 13

Iterations of Item Fit with P-Values

Category	Item	Round 1	Round 2	Round 3	Round 4
Build/Draw					
	Maps	0.006	0.184	0.303	0.262
	Plans	0.954	0.572	0.070	0.522
Identify					
	2D Shapes	0.948	0.887	0.816	0.357
	3D Shapes	0.005	-	-	-
	Photo	0.010	0.504	0.820	0.358
	Slice	0.642	0.364	0.330	0.125
	Real Shapes	0.458	0.116	0.067	0.383
Digital					
	Organize	0.006	0.039	-	-
Play					
	Puzzles	0.047	0.245	0.695	0.633
	Blocks	0.197	0.116	0.220	0.470
	Board Games	0.012	0.073	0.356	0.090
Language					
	Features of Figures	0.458	0.183	0.603	0.463
	Landmarks	0.108	0.053	0.109	0.307
	Hand Motions	0.875	0.167	0.516	0.441
Interaction					
	Build	0.982	0.625	0.224	0.200
	Hand Motions	0.778	0.199	0.391	0.067
	Positional Language	0.156	0.368	0.363	0.167
	Shape Names	0.202	0.061	0.008	-
	Shape ID	0.355	0.059	0.132	0.198

After consulting with the project team, we decided to approach the analyses in three different ways. The first approach was to remove the three lowest p-value items simultaneously. The next two approaches were to remove the next-to-lowest p-value items one at a time and continue with the backward selection approach. Table 14 describes the results of the first approach. After the removal of the first three items, additional items were identified for removal.

Table 14

Item Statistics with Removing Three Lowest Simultaneously

Category	Item	Round 1	Round 2	Round 3
Build/Draw				
	Maps	0.006	-	-
	Plans	0.954	0.628	0.403
Identify				

	2D Shapes	0.948	0.927	0.981
	3D Shapes	0.005	-	-
	Photo	0.010	0.718	0.889
	Slice	0.642	0.786	0.035
	Real Shapes	0.458	0.599	0.113
Digital				
	Organize	0.006	-	-
Play				
	Puzzles	0.047	0.710	0.246
	Blocks	0.197	0.879	0.228
	Board Games	0.012	0.315	0.135
Language				
	Features of Figures	0.458	0.473	0.491
	Landmarks	0.108	0.134	0.050
	Hand Motions	0.875	0.477	0.821
Interaction				
	Build	0.982	0.215	0.309
	Hand Motions	0.778	0.347	0.248
	Positional Language	0.156	0.018	-
	Shape Names	0.202	0.142	0.461
	Shape ID	0.355	0.469	0.529

Table 15 describe the results of removing Draw Maps item first. This analysis also found additional items for removal.

Table 15

Approach When Removing Draw Maps

Category	Item	Round 1	Round 2
Build/Draw			
	Maps	0.006	-
	Plans	0.954	0.130
Identify			
	2D Shapes	0.948	0.771
	3D Shapes	0.005	0.044
	Photo	0.010	0.304
	Slice	0.642	0.784
	Real Shapes	0.458	0.311
Digital			
	Organize	0.006	0.025
Play			
	Puzzles	0.047	0.290
	Blocks	0.197	0.166
	Board Games	0.012	0.141
Language			

	Features of Figures	0.458	0.470
	Landmarks	0.108	0.009
	Hand Motions	0.875	0.837
Interaction			
	Build	0.982	0.444
	Hand Motions	0.778	0.121
	Positional Language	0.156	0.685
	Shape Names	0.202	0.251
	Shape ID	0.355	0.125

The last iteration included the removal of the Digital Organize item first. Table 16 describes these results. After removing the item, we found no additional items were identified for misfit. We moved forward with this approach.

Table 16

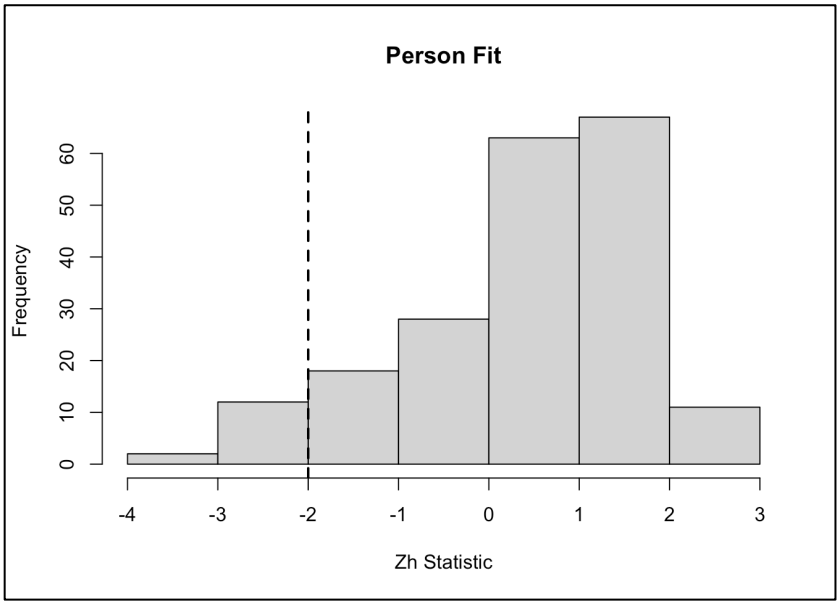
Approach When Removing Digital Organize

Category	Item	Round 1	Round 2
Build/Draw			
	Maps	0.006	0.501
	Plans	0.954	0.822
Identify			
	2D Shapes	0.948	0.945
	3D Shapes	0.005	0.074
	Photo	0.010	0.094
	Slice	0.642	0.764
	Real Shapes	0.458	0.407
Digital			
	Organize	0.006	-
Play			
	Puzzles	0.047	0.512
	Blocks	0.197	0.410
	Board Games	0.012	0.247
Language			
	Features of Figures	0.458	0.682
	Landmarks	0.108	0.604
	Hand Motions	0.875	0.297
Interaction			
	Build	0.982	0.937
	Hand Motions	0.778	0.540
	Positional Language	0.156	0.575
	Shape Names	0.202	0.063
	Shape ID	0.355	0.243

Another indication of model fit is an analysis of person fit. We use the Zh statistic, which is a standardized statistic for measuring person fit of a model. Large quantities of extreme Zh values might indicate model misfit. Figure 3 is the distribution of the Zh statistic for the current model. Due to the small number of participants with extreme Zh values, we do not have evidence of model misfit.

Figure 3

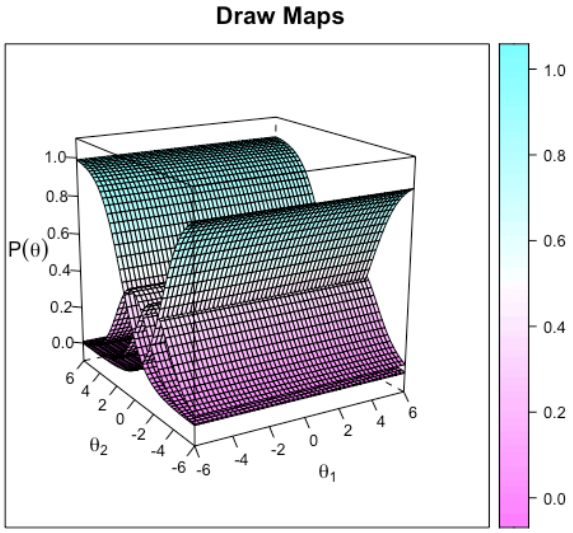
Distribution of Person Fit Statistic



Next, we plot the item characteristic curves of each item. Figure 4 is the ICC for the first item. The remaining items can be found in Appendix B.

Figure 4

ICC for Draw Maps



We include the test information surface in addition to the standard error surface in Figures 5 and 6, respectively. The test information surface indicates higher levels of information from the instrument at the intersection of both ability scales. Conversely, the standard error surface indicates less error at this intersection.

Figure 5

Test Information Surface

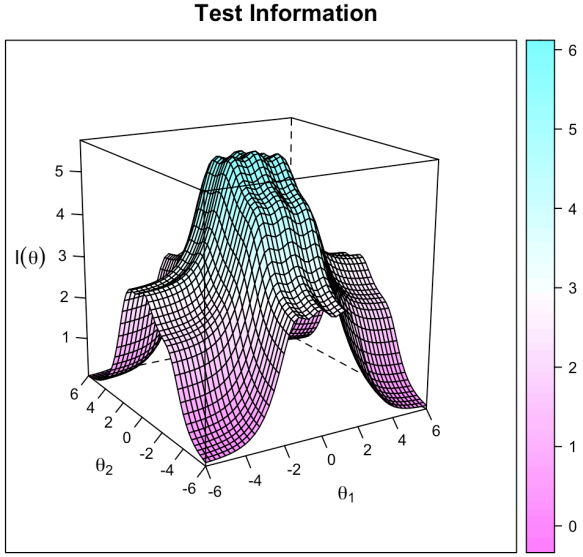
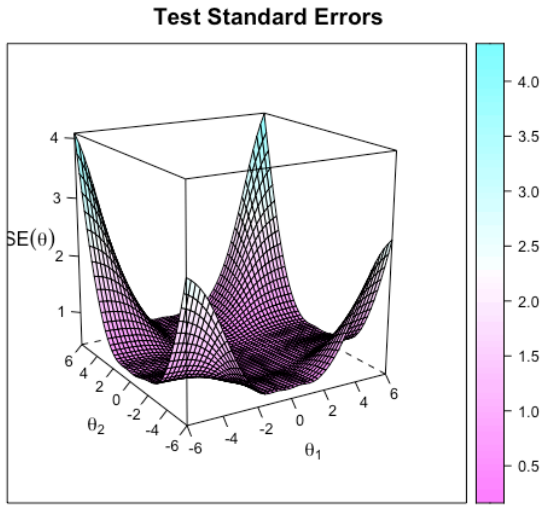


Figure 6

Standard Error Surface



Lastly, we report the item difficulties (thresholds) and discriminations in Table 17. We note that the item discriminations range from 0.67 to 2.46 and item thresholds range from -6.21to 2.83. We also report the 95% confidence intervals for each threshold, which are based on the standard error. We do not note any overlap in the 95% confidence intervals, which provides evidence in support of the item scales, as overlapping intervals may indicate a need to collapse adjacent response categories.

Table 17

Item Difficulties and Discriminations

Category	Item	Discrimination	Threshold 1	Threshold 2	Threshold 3
Build/Draw					
	Maps	0.946	-2.286 (-3.415, 2.256)	-0.887 (-1.236, -0.538)	1.141 (0.771, 1.510)
	Plans	0.666	-2.640 (-3.180, 2.100)	-0.056 (-0.359, 0.248)	1.967 (1.532, 2.402)
Identify					
	2D Shapes	2.086	-6.209 (-7.886, 4.532)	-0.803 (-1.285, -0.322)	2.011 (1.403, 2.618)
	3D Shapes	1.099	-3.601 (-4.371, -2.832)	-0.855 (-1.217, -0.494)	1.761 (1.312, 2.209)
	Photo	1.123	-3.146 (-3.788, -2.503)	-0.509 (-0.860, -0.158)	1.698 (1.259, 2.136)
	Slice	1.456	-3.593 (-4.337, -2.849)	-0.732 (-1.128, -0.335)	1.958 (1.452, 2.463)
	Real Shapes	1.626	-3.497 (-4.245, -2.749)	-0.933 (-1.359, -0.507)	2.165 (1.601, 2.729)

Play					
	Puzzles	2.000	-3.032 (-3.729, -2.335)	-0.259 (-0.711, 0.193)	2.517 (1.877, 3.157)
	Blocks	1.365	-2.134 (-2.634, -1.634)	-0.751 (-1.140, -0.363)	1.538 (1.093, 1.982)
	Board Games	1.331	-2.543 (-3.098, -1.988)	-0.121 (-0.486, 0.245)	2.820 (2.206, 3.433)
Language					
	Features of Figures	1.600	-2.696 (-3.301, -2.090)	0.169 (-0.230, 0.569)	2.822 (2.192, 3.453)
	Landmarks	1.627	-1.759 (-2.249, -1.270)	0.252 (-0.148, 0.652)	2.314 (1.750, 2.877)
	Hand Motions	1.354	-2.232 (-2.740, -1.724)	-0.445 (-0.818, -0.073)	2.079 (1.575, 2.584)
Interaction					
	Build	2.461	-3.651 (-4.569, -2.733)	-0.548 (-1.073, -0.022)	2.830 (2.053, 3.606)
	Hand Motions	1.537	-2.058 (-2.566, -1.550)	-0.167 (-0.556, 0.223)	2.304 (1.744, 2.863)
	Positional Language	1.141	-2.116 (-2.587, -1.644)	-0.389 (-0.737, -0.041)	1.719 (1.277, 2.161)
	Shape Names	1.219	-2.377 (-2.896, -1.857)	-0.073 (-0.428, 0.282)	2.374 (1.849, 2.899)
	Shape ID	1.524	-2.097 (-2.611, -1.583)	-0.130 (-0.517, 0.257)	2.299 (1.753, 2.845)

Conclusions

This technical report describes the distribution and analysis of the updated spatial reasoning home environment survey. We distributed the survey to parents of students in grades K-2. After data cleaning, we retained data from 201 parents. Exploratory factor analyses indicated a two-factor instrument in which 19 of the original 26 items were retained. Further IRT analyses resulted in the retention of 18 of these 19 items. Further work to confirm this factor structure and to contribute additional evidence to the internal structure can include a confirmatory factor analysis with a larger sample size.

References

Desjardins, C. D., & Bulut, O. (2018). *Handbook of educational measurement and psychometrics in R.* CRC Press.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: A multidisciplinary journal*, *6*(1), 1-55.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* New York: Springer.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement,* No. 17.

Sarama, J., & Clements, D. (2009a). Geometry and Spatial Thinking. In *Early childhood mathematics education research: Learning trajectories for young children* (pp. 247–269).

Sarama, J., & Clements, D. (2009b). Shape. In *Early childhood mathematics education research: Learning trajectories for young children* (pp. 199–246).

Sparks, A., Geller, J., Perry, L., & Ketterlin-Geller, L. (2020). *Spatial reasoning: Home environment questionnaire phase 1* (Tech. Rep. No. 20-09). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (Vol. 4). Pearson.

Wellberg, S., Geller, J., McMurrer, J., Sparks, A., & Ketterlin-Geller, L. (2020). *Spatial reasoning: Home environment questionnaire phase 2 development* (Tech. Rep. No. 20-30). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Appendix A – Eligibility Survey

SR Home Use Eligibility Survey

Start of Block: CAPTCHA Block

Q10

Thank you for your interest in Research in Mathematics Education's Spatial Reasoning Home Use Survey.

The questions in this short questionnaire are intended to identify if you are eligible to take the survey. The survey is designed for parents of students between kindergarten and second grade. If you are eligible, you will receive a followup email within 1-2 business days with the link to complete the survey. Upon completion of that survey, you will receive a **$10 Amazon gift card.**

If you have any questions, you may contact us at rme@smu.edu

*Si desea continuar en español, seleccione "Español" arriba.

Q12 Verification

End of Block: CAPTCHA Block

Start of Block: Block 2

Q14 Before you participate in this eligibility survey, you need to read this form. When you read a form like this to learn about a research study, it is called "informed consent." When you give your consent for something, it is the same thing as giving your permission. Your personal participation in this study is voluntary.

If you agree to take part and then change your mind, you can withdraw for any reason. There are no penalties if you withdraw, decline to participate, or skip any parts of the survey. If you agree to participate in this eligibility survey, you will answer questions to confirm your eligibility to participate in a research study we are conducting to learn more about the developmental appropriateness of spatial reasoning skills in K-2 mathematics.
You will be asked to provide your email address so that we can send you a personalized link to complete the online research study survey, if you meet the eligibility requirements.
Risks associated with this survey are minimal. You have a full right to privacy. This means that only the researchers who are part of this study will see the information about you from this survey. No results from this survey will be shared outside of the study research team. All data

from this survey will be kept safe from access by people who should not see it, through password-protection. If you have questions about this study or the procedures, please email lkgeller@smu.edu.
If you have concerns or questions about the study or have a research-related injury, you may contact:

Leanne Ketterlin Geller, Ph.D.
Texas Instruments Chair in Education
Professor, Education Policy & Leadership
Director, Research in Mathematics Education
Simmons School of Education and Human Development
Phone: 214/768-4947
Email: lkgeller@smu.edu

If you have questions about your rights as a participant or feel you have been placed at risk, you may contact:
Professor Christopher Dolder, IRB Chair researchcomplaince@smu.edu 214-768-2033
Would you like to participate in this research study?
By clicking "I agree" below, you agree to participate in this research study.

- ○ I agree (1)
- ○ I disagree (2)

Skip To: End of Survey If Before you participate in this eligibility survey, you need to read this form. When you read a fo... = I disagree

End of Block: Block 2

Start of Block: Default Question Block

Q5 Do you have one or more children who are in the grade of kindergarten through second grade during the current 2020-2021 school year?

- ○ Yes (1)
- ○ No (2)

Skip To: End of Survey If Do you have one or more children who are in the grade of kindergarten through second grade during... = No

*

Q16 In what year was one of your children who is in kindergarten through second grade during the current 2020-2021 school year born?

YYYY

_____

Q7 A valid email address is required to receive a link to the survey. Please enter a valid email address below:

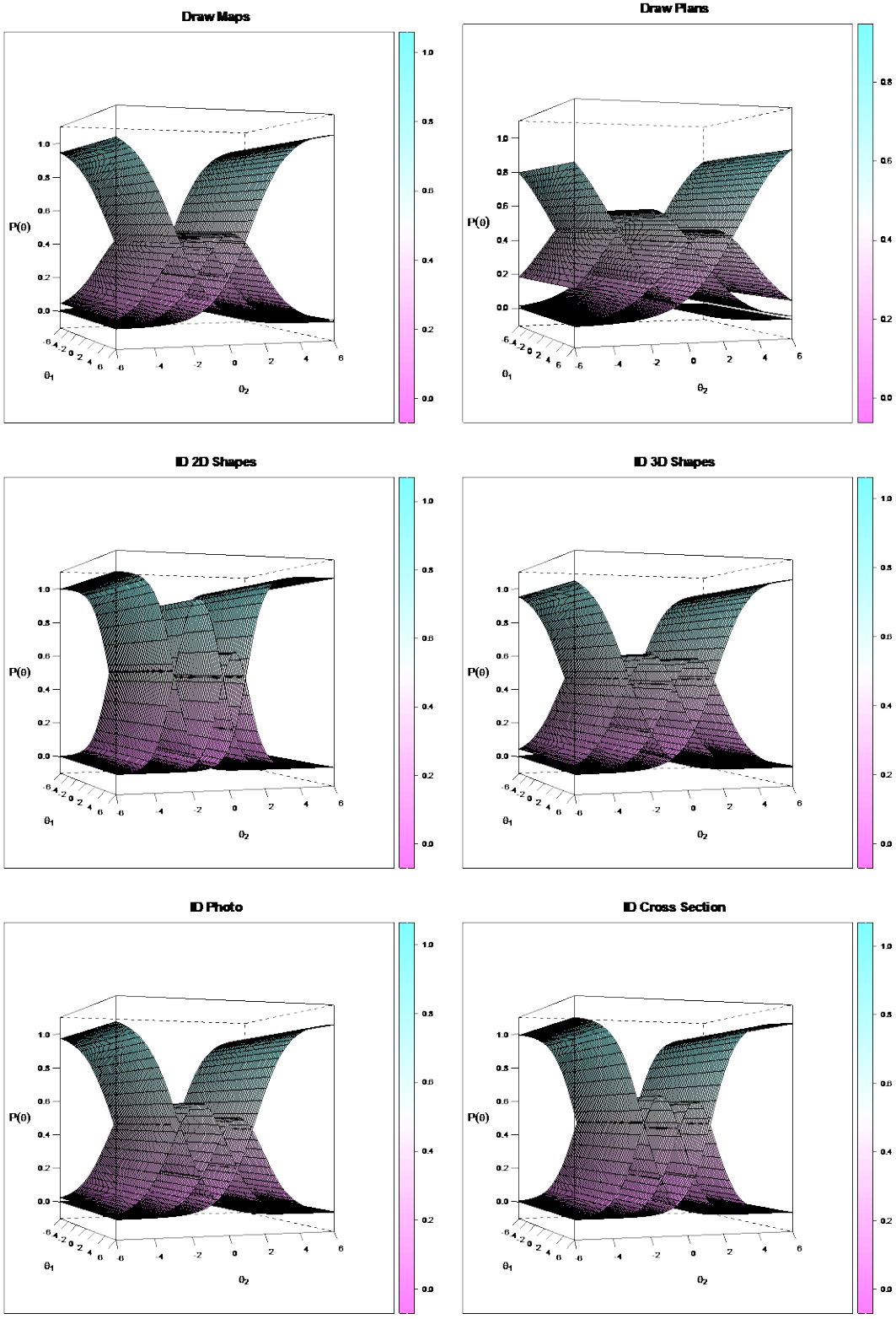
_____

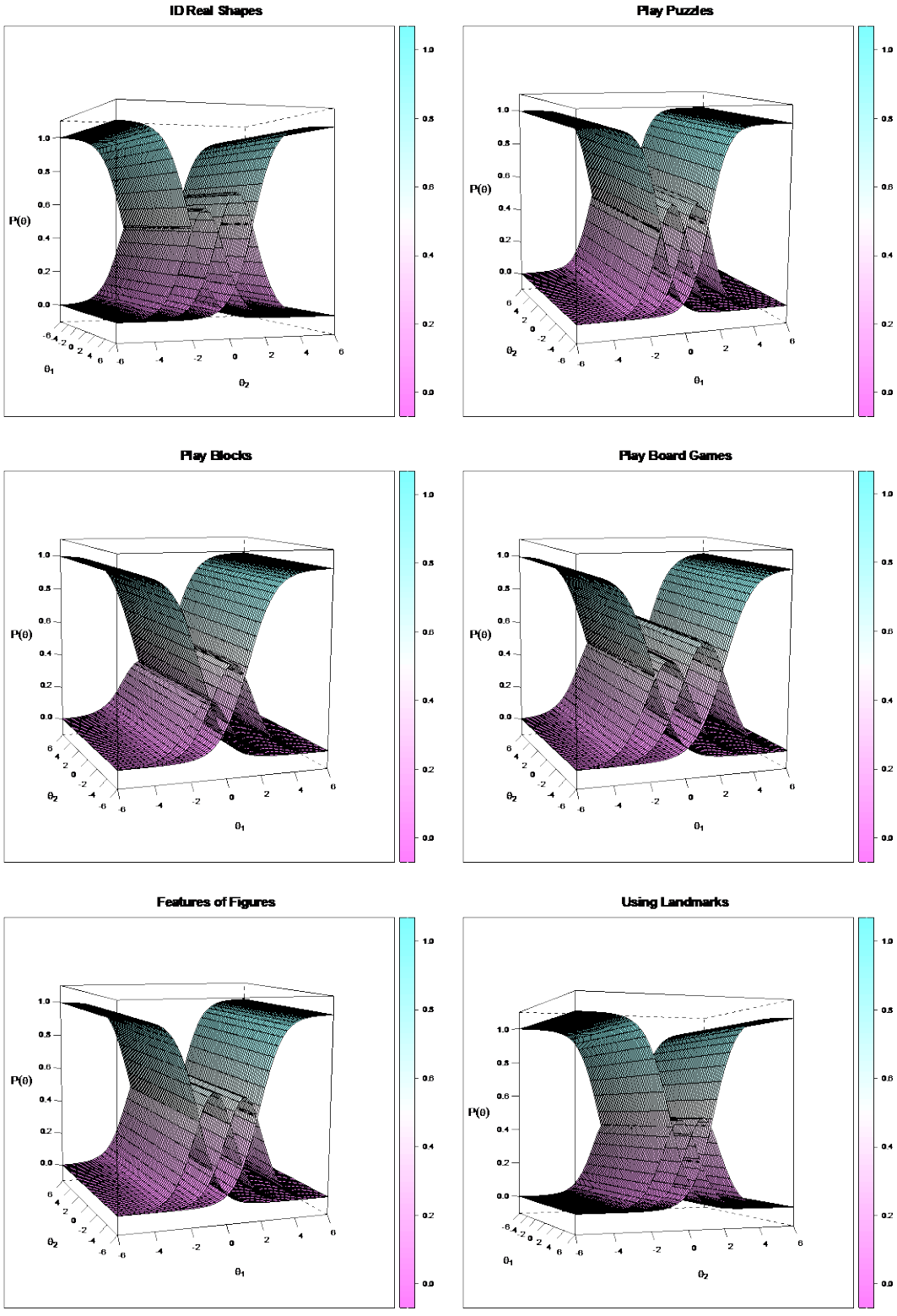
Q9 Confirm email address below:

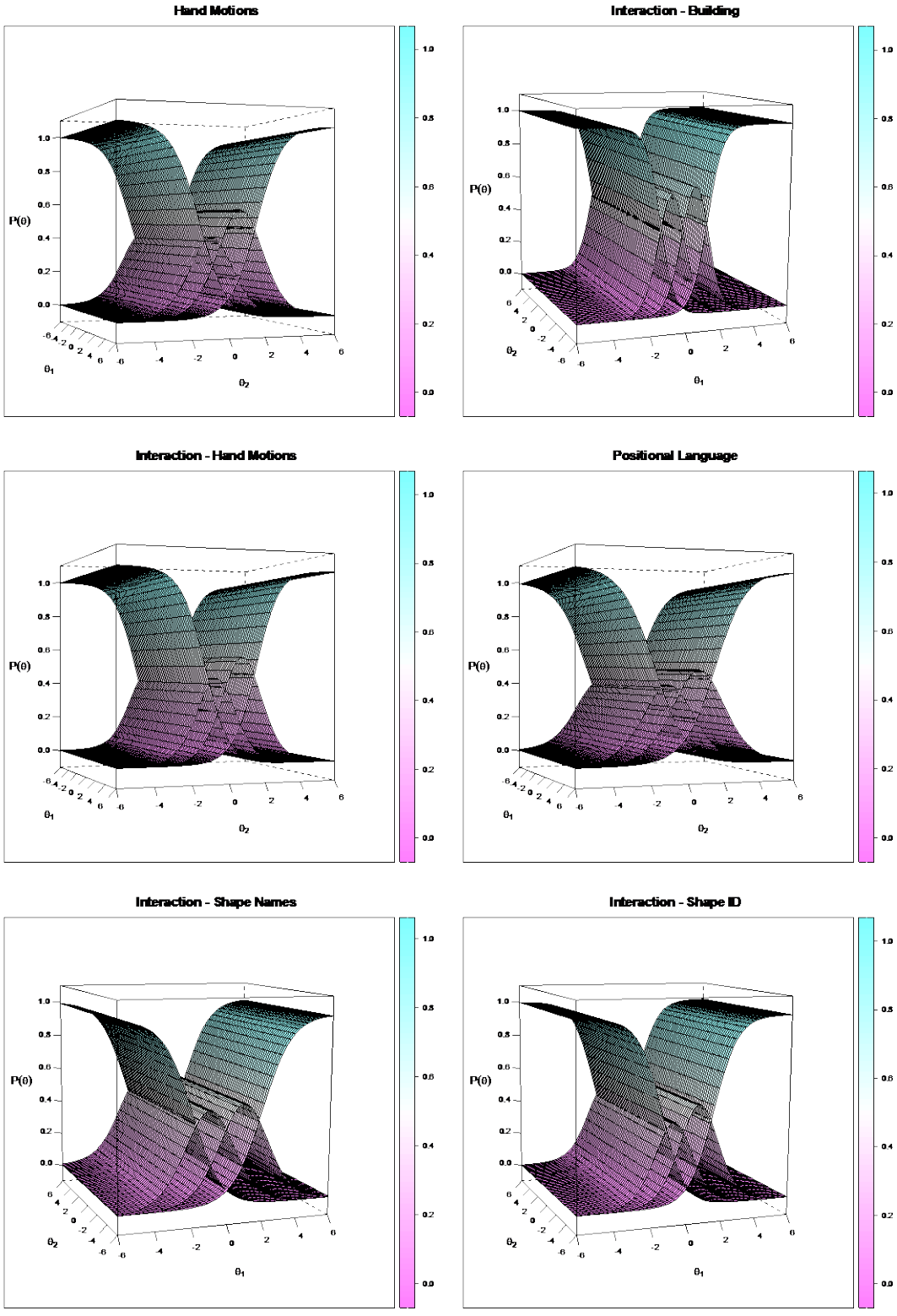
_____

End of Block: Default Question Block

Appendix B – Item Characteristic Curves




ID Real Shapes
Play Puzzles
Play Blocks
Play Board Games
Features of Figures
Using Landmarks
P(θ)
θ1
θ2


Hand Motions
Interaction - Building
Interaction - Hand Motions
Positional Language
Interaction - Shape Names
Interaction - Shape ID
P(θ)
θ1
θ2