

Spatial Anxiety: A novel questionnaire with subscales for measuring three aspects of spatial anxiety

Ian M. Lyons^{1,1}, Gerardo Ramirez^{1,2,3}, Erin A. Maloney⁴, Danielle N. Rendina⁵,
Susan C. Levine⁶, Sian L. Beilock^{6,7}

¹The first two authors contributed equally to this manuscript and are listed alphabetically.

¹Georgetown University, Department of Psychology

²University of California, Los Angeles, Department of Psychology

³University of California, Los Angeles, Graduate School of Education & Information Studies

⁴University of Ottawa, School of Psychology

⁵University of Wisconsin-Madison, Department of Psychology

⁶University of Chicago, Department of Psychology

⁷Barnard College

Funding

This research was funded by NSF CAREER DRL-0746970 to Beilock, NSF Spatial Intelligence Learning Center (SBE-0541957) to Beilock and Levine, Departmental Start-Up Funds (UCLA, Psychology) to Ramirez, and Departmental Start-Up Funds (Georgetown, Psychology) to Lyons. Send email correspondence to ian.lyons@georgetown.edu.

ABSTRACT

Spatial skills are a strong predictor of achievement and pursuit of employment in STEM fields. However, some individuals experience anxiety arising from situations that require performing spatial tasks in an evaluative context, and as a result, may avoid spatial related mental activities and exposure to spatially relevant experiences. We sought to generate and validate an instrument capable of reliably measuring individual differences in experiences of spatial anxiety. We developed a spatial anxiety data-driven approach, wherein an exploratory factor analysis was conducted within the framework for different types of spatial skills outlined by Uttal et al. (2013). In Study 1, factor analyses revealed that items loaded on three factors that corresponded well with some of the most common spatial abilities that have been discussed in the broader literature: navigation, mental-manipulation and imagery. The three subscales were high in internal reliability and between-scale selectivity. Study 2 then established that external validity was good for the navigation and manipulation subscales: higher anxiety ratings uniquely predicted lower objective performance and lower attitude/ability ratings on established measures within the respective subdomains. External validity was acceptable for the imagery subscale, uniquely predicting lower attitude/ability ratings on an established spatial imagery questionnaire. The overall result is an empirically validated Spatial Anxiety scale for use with adults that also respects the multifaceted nature of spatial processing. This questionnaire has the potential to provide a more comprehensive screening tool for spatial anxiety, and is a step toward identifying potential barriers to STEM education.

Keywords

Spatial Anxiety; Spatial Anxiety Questionnaire; Spatial Processing; Spatial Attitudes; Individual Differences

Spatial Anxiety: A novel questionnaire with subscales for measuring three aspects of spatial anxiety

INTRODUCTION

It is now well recognized that spatial thinking (i.e., the ability to represent and transform, symbolic, non-linguistic information; Gardner, 1983) shapes individuals' capacity to learn and succeed in science, technology, engineering and math (STEM) fields. Spatial ability is a strong predictor of STEM entry and retention (Benbow & Stanley, 1982; Shea et al., 2001; Wai et al., 2009; Uttal & Cohen, 2014). Greater spatial abilities also predict better grades in calculus, physics and chemistry (Sorby, 2009; Kozhevnikov et al., 2002), success in three-dimensional biology problems (Russell-Gebbett, 1985), and bedrock mapping tasks within geology (Hambrick et al., 2011). Importantly there exist individual differences in the capacity for spatial thinking that derives from a number of sources; for instance, hormonal variation (Grimshaw et al., 1995), culture (Berry, 1966; Hoffman et al., 2011), and group differences in environmental exposure to spatial activities (Levine et al., 2012; Terlecki & Newcombe, 2005). Spatial skills are nevertheless malleable and can be enhanced with training and experience (Baenninger & Newcombe, 1989; Brinkmann, 1966; Lord, 1985; Uttal et al., 2013), should individuals seek opportunities to hone their spatial thinking skills.

An important, and often overlooked, factor that relates to spatial ability is spatial anxiety, defined as fear and apprehension towards spatial processing, which can prevent individuals from engaging in experiences and opportunities that might otherwise promote the development of spatial skills. A high degree of spatial anxiety has been found to predict problems with everyday activities such as performance on spatial puzzle tasks (Ramirez et al., 2013), as well as sense of direction (Kremmyda et al., 2016; Lawton, 1994) and math anxiety among adults (Ferguson et al., 2015). Spatial anxiety is also associated with reduced workplace efficacy among educators, as children whose teachers have higher spatial anxiety show reduced gains in spatial skills across the school year (Gunderson et al., 2014).

Studies examining spatial anxiety suggest considerable variation among individuals' experience. Enhancing our ability to assess spatial anxiety can have important ramifications for identifying individuals who are more likely to struggle with important workplace responsibilities (e.g., interpreting

scientific figures, visualizing models, imaging anatomical structure, etc.), as well as everyday spatial activities (e.g., parent-child play around building blocks and puzzles, packing, parallel parking, route planning, etc.; Bronzaft et al., 1976). More importantly, identifying individuals with a high degree of spatial anxiety may enable researchers and policy-makers to make better recommendations for improving spatial training and even in the selection of candidates for career-specific training (e.g. for dentistry and medicine; Hegarty, Keehner, Cohen, Montello, & Lippa, 2007; Hegarty, Keehner, Cohen & Montello, 2009).

The most widely-used measure of spatial anxiety is currently a scale created by Lawton (1994). While this scale represented the dominant conceptualization of spatial processing at the time of its creation, recent advances in our understanding of spatial cognition indicate that this scale is no longer an appropriate measure of general spatial anxiety. Specifically, Lawton's (1994) spatial anxiety scale could be more appropriately described as a measure of environmental navigation anxiety, which thus represents only one dimension of spatial processing. We are now aware of the involvement of spatial skills in a number of activities beyond navigation. For instance, many factor analytic studies implicate various types of spatial abilities (Carroll, 1993; Eliot, 1987; Linn & Petersen, 1985; Lohman, 1988; Thurstone, 1947). Hence, the development of tools that account for the multifaceted nature of spatial skills and are informed by modern typology for spatial thinking is needed.

In this article, we detail the creation of a novel spatial anxiety scale, for use with adults, informed by Uttal et al.'s (2012) four cell classification system. In this well-accepted classification system, there are four categories of spatial processing into which all spatial tasks can be compartmentalized. This framework crosses the intrinsic (i.e., the relation of parts that define an object) vs extrinsic (i.e., the relation among objects in a group) distinction against the static (fixed object) vs dynamic (moving object) distinction. We choose this framework as an initial starting point for development of the current scale because it was developed from a top-down theory-driven analysis of the nature of spatial thinking and grounded with respect to work in STEM disciplines and supported by various lines of research (e.g., Kozhevnikov & Hegarty, 2001; Huttenlocher & Presson, 1973; Kozhevnikov, Hegarty & Mayer 2002; Kozhevnikov, Kosslyn & Shepard, 2005).

Hence, the current study set out to develop a reliable scale that measures individual differences in various sub-types of spatial anxiety (Study 1), and then to provide external validity for the factor structure of this newly developed spatial anxiety scale through ability and self-rating tasks (Study 2). We

took special care to distinguish spatial anxiety from more general trait anxiety, as well as to establish discriminant validity among our various spatial anxiety subscales.

STUDY 1

The goal of Study 1 was to develop a novel Spatial Anxiety Scale incorporating the evidence that spatial processing is not a unitary construct. More specifically, our aim was to identify the appropriate number of subscales to include, and then to identify items most representative of each subscale.

Methods

Participants

Participants were 485 adults recruited via Amazon's Mechanical Turk. Of these, 449 generated a complete dataset; hence, subsequent analyses for Study 1 (with the exception of initial item triage – see below) preceded with $N=449$ (227 female; age: range=18.1 to 67.5yrs, $M=33.59$, $SD=11.34$).

Procedure

All procedures and materials were reviewed and approved by the University of Chicago Institutional Review Board (IRB). The study consisted of a main questionnaire comprising the 80 candidate items (see Table 1 below) presented in randomized order. Several filler questionnaires and basic demographics information were also collected to prevent participants from inferring the purpose of the study. The various questionnaires were presented in randomized order, with the exception that the demographics survey was always collected last. All participants gave informed consent at the beginning of the survey. The study took approximately 20-25 minutes to complete. Participants were compensated \$4 for their time.

Stimuli and Materials

Initial Item Generation and Triage

Initially, 130 items were generated to address the four categories of spatial skills put forth by Uttal et al. (2013): intrinsic-static (e.g., detailed object imagery), intrinsic-dynamic (e.g., mental rotation or mental manipulation), extrinsic-static (e.g., comparing scales on a map), and extrinsic-dynamic (e.g., navigation).

A major goal in creating a spatial anxiety survey is to discriminate reliably between high and low spatially anxious individuals. Hence, an important preliminary requirement is eliminating candidate items that show floor/ceiling effects and/or little variability in responses. To this end, we conducted a brief pilot/triage survey via Mechanical Turk ($N=64$, 31 female, mean age=35.58yrs). The 130 initial candidate

items were scored on a 0 (not at all anxious) to 4 (very anxious) scale. To pass triage and thus be considered for the main study, a given item was required to generate a mean response of greater than 1.0 (no floor effect), less than 3.0 (no ceiling effect), in addition to a standard deviation of at least 1.0 (sufficient variability). These prerequisites eliminated 50 items, and ensured that the remaining 80 items are likely to capture at least some meaningful variance with respect to spatial anxiety. The remainder of the study proceeded with these items.

Items

The 80 items that passed triage are given in Appendix A. Also shown is their ostensible a priori categorization into subscales with respect to Uttal et al.'s (2013) classification. Note that we refer henceforth to the four categories using slightly more intuitive terms, which are abbreviated accordingly: Imagery (I) for intrinsic-static, Mental Manipulation (M) for intrinsic-dynamic, Scalar Comparison (S) for extrinsic-static, and Navigation (N) for extrinsic-dynamic. Where a given item is ambiguous with respect to category, multiple categories may be indicated (e.g., 'IM'). These categorizations are entirely a priori, and were thus verified or rejected according to the factor analysis that follows (see Results).

To avoid bias, the categories from above (and in Appendix A) were not shown to participants, and all 80 items were presented in a random order (randomized across participants, regardless of category). Participants were given the following instructions: "The items in the questionnaire below refer to situations and experiences that may cause tension, apprehension, or anxiety. For each item, mark the response that describes *how much you would be made to feel anxious by it*. Work quickly, but be sure to think about each item." Response options: 'not at all', 'a little', 'a fair amount', 'much', 'very much'. Items were scored 0 (not at all) to 4 (very much).

Results

As noted in the Introduction, spatial processing – and hence spatial skill or ability – is hardly a unitary construct. Hence, the aims of this study were twofold: (1) to identify the number and nature of subscales appropriate to the broader goal of measuring spatial anxiety, and (2) to identify the items that best comprise each subscale. To achieve the first goal, we merged a data-driven approach by employing an exploratory factor analysis with a more theoretically driven approach using item-labels that were informed by the framework for different types of spatial skills outlined by Uttal et al. (2013) (see Appendix A).

Factor Analysis

We first entered all 80 items into an exploratory factor analysis using maximum likelihood extraction. We opted for this extraction method over orthogonal extraction (principal component analysis) because we deemed it reasonable to allow for the possibility that various aspects of spatial processing, and consequently spatial anxiety, are related to one another (i.e., not fully orthogonal). For the same reason, to generate rotated solutions, we thus used the Direct Oblimin ($\delta=0$) rotation method.

Table 1 – Factors

Factor	Eigenvalue	% Variance
1	30.03	37.54
2	4.97	6.21
3	3.38	4.23
4	2.31	2.89
5	1.76	2.20
6	1.48	1.85
7	1.27	1.58
8	1.18	1.48
9	1.09	1.36
10	1.02	1.28
11	1.01	1.27

Extraction yielded 11 factors with eigenvalues greater than 1 (Table 1). However, 11 different subscales would prove impractical, so we instead opted for the more conservative initial cut-off of eigenvalues of at least 2 (above the line in Table 1). Note that this also limited us to factors capturing at least 2.5% of the total variance. Moreover, the 4 factors corresponds nicely to the number of spatial skill-types proposed by Uttal et al. (2013), which was also the basis for how we generated the items in the first place. To get a clearer sense of what each factor represents, we next examined the loadings of individual items for the rotated 4-factor solution. The rotated factor loadings are shown in Table 2a. Note that loadings were taken from the pattern matrix; thus they represent *unique* factor loadings.

Factors 2 and 4 appeared to correspond clearly to a priori categories. Specifically, Factor 2 loaded highly on Navigation (N or SN) items almost exclusively. Some Scalar Comparison (S) items also loaded highly on this factor, though these tended to be thematically related to navigation (e.g., S07: “Memorizing routes and landmarks on a map for an upcoming exam”). Factor 4 loaded most highly on Imagery (I or IS) items.

Factor 1 appeared to comprise a combination of Mental-Manipulation (M) items, Scalar-Comparison (S) items, and items that were thought to be a combination thereof (SM). Indeed, half of the top 8 loadings belonged to combination (SM) items. One perspective is thus that Factor 1 simply merged the two Mental-Manipulation (M) and Scalar-Comparison (S) categories. On the other hand, two-thirds (14 of 21) M items showed loadings on Factor 1 of .5 or greater, while less than half (5 of 12) S items did so. In addition, the S and SM items that loaded highest on Factor 1 (bold items in Table 2) dealt with dynamic and often multi-dimensional mental imagery (e.g., SM01: “Asked to imagine the 3-dimensional structure of a complex molecule using only a 2-dimensional picture for reference”, SM09: “Asked to imagine the motion of a mechanical system given a static picture of the system”). For this reason, we concluded this factor and the items that loaded highest on it as comprising a component best characterized as (anxiety about) Spatial Mental-Manipulation (or the intrinsic-dynamic category from Uttal et al., 2013).

Factor 3 appears to have comprised items largely about anxiety (or comfort, given the negative factor loadings) about performing spatially related tasks in front of a classroom. Indeed, all four items highlighted in Table 2 for Factor 3 mention doing some activity ‘in class’ or ‘in front of one’s class’. Hence, this Factor was deemed as one of no interest for present purposes (and these items were of course omitted from the final version of the scale).

Given that Factor 3 failed to correspond to any of the four a priori categories, one possibility is that a 5th factor may be warranted – and in particular, that said 5th factor might correspond to the missing category: Scalar-Comparison (S). To this end, we also examined rotated factor loadings from a 5-factor solution. However, this factor showed an absolute loading greater than 0.5 (-.532) on only one item (M06: “Asked to imagine how the orbit of a comet changes over time”). Moreover, the loadings of the other items on the first four factors remained largely unchanged.

Taking the above results together, we believe that the most robust interpretation of the factor analysis is that we identified three key components of spatial anxiety which correspond most closely to the three a priori categories Spatial Mental-Manipulation (Factor 1), Spatial Navigation (Factor 2), and Spatial Imagery (Factor 4). Therefore, the Spatial Anxiety Scale described in the following section comprises three subscales corresponding to these three factors/categories.

Table 2 – Rotated Factor Loadings

Item	Factor1	Factor2	Factor3	Factor4	Item	Factor1	Factor2	Factor3	Factor4
I01				0.604	N01				
I02				0.519	N02		0.768		
I03				0.514	N03		0.809		
I04					N04		0.738		
I05					N05		0.805		
I06					N06		0.823		
I07				0.723	N07	0.613			
I08					N08		0.771		
I09				0.643	N09		0.529		
I10				0.787	N10				
I11				0.537	N11		0.520		
I12					N12		0.534		
I13					N13		0.683		
I14					S01	0.714			
IM01					S02	0.685			
IS01					S03			-0.543	
IS02			-0.626		S04			-0.711	
IS03					S05				
IS04				0.508	S06	0.555			
M01					S07		0.597		
M02			-0.534		S08		0.538		
M03					S09	0.720			
M04	0.721				S10				
M05	0.700				S11	0.534			
M06	0.712				S12		0.586		
M07					SM01	0.788			
M08	0.724				SM02	0.599			
M09	0.771				SM03	0.577			
M10	0.633				SM04	0.762			
M11	0.677				SM05	0.532			
M12	0.673				SM06	0.543			
M13	0.603				SM07	0.596			
M14					SM08				
M15					SM09	0.736			
M16	0.574				SM10				
M17	0.522				SM11	0.783			
M18	0.595				SM12				
M19					SN01		0.573		
M20	0.602				SN02		0.586		
M21	0.551				SN03		0.573		

Note. Table 2 shows rotated factor loadings for the 4-factor solution for all 80 items. Loadings whose absolute value are less than .5 are suppressed. The top 8 loadings for each factor are in **bold**. See Table 1 for abbreviations. See Appendix A for item text.

*Spatial Anxiety Scale***Table 3 – Spatial Anxiety Scale**

Subscale	Item
M	Asked to imagine the 3-dimensional structure of a complex molecule using only a 2-dimensional picture for reference
M	Asked to determine how a series of pulleys will interact given only a 2-dimensional diagram
M	Asked to imagine and mentally rotate a 3-dimensional figure
M	Asked to imagine a 3-dimensional structure of the human brain from a 2-dimensional image
M	Asked to imagine the motion of a mechanical system given a static picture of the system
M	Imagining on a test what a 3-dimensional landscape model would look like from a different point of view
M	Asked to imagine the 3-dimensional shape created by rotating a complex 2-dimensional plane on an exam
M	Using a 3-dimensional model of an airport to complete a homework assignment
N	Finding your way to an appointment in an area of a city or town with which you are not familiar
N	Finding your way back to your hotel after becoming lost in a new city
N	Asked to follow directions to a location across town without the use of a map
N	Finding your way back to a familiar area after realizing you have made a wrong turn and become lost while driving
N	Trying to get somewhere you have never been to before in the middle of an unfamiliar city
N	Trying a new route that you think will be a shortcut without the benefit of a map
N	Asked to do the navigational planning for a long car trip
N	Memorizing routes and landmarks on a map for an upcoming exam
I	Asked to recall the shade and pattern of a person's tie you met for the first time the previous evening
I	Asked to give a detailed description of a person's face whom you've only met once
I	Asked to recall the exact details of a relative's face whom you have not seen in several years
I	Asked to recreate your favorite artist's signature from memory
I	Describing in detail the cover of a book to a bookseller because you've forgotten both the title and author of the book
I	Tested on your ability to create a drawing or painting that reproduces the details of a photograph as precisely as possible
I	Asked to imagine and describe the appearance of a radio announcer or someone you've never actually seen
I	Given a test in which you were allowed to look at and memorize a picture for a few minutes, and then given a new, similar picture and asked to point out any differences between the two pictures

Note. Table 3 gives the complete final Spatial Anxiety Scale broken into its three subscales: Mental Manipulation (M), Navigation (N), and Imagery (I). Instructions: "The items in the questionnaire below refer to situations and experiences that may cause tension, apprehension, or anxiety. For each item, mark the response that describes *how much you would be made to feel anxious by it*. Work quickly, but be sure to think about each item." Response options: 'not at all', 'a little', 'a fair amount', 'much', 'very much'. Scoring: 0 (not at all) to 4 (very much); sum scores across the 8 items for each subscale.

To generate the final Spatial Anxiety Scale, we used the three factors identified in the previous section. For the specific items, we selected the 8 highest-loading items for each factor (bold items in Table 2). We chose the number 8 for several reasons. First, all three factors yielded at least 8 items with loadings of at least .5 or greater. Second, this number of items struck a pragmatic balance between respecting the specific subscales (and hence different types of spatial processing), and keeping the overall scale to a reasonable length. Specifically, 8 items is generally considered acceptable for establishing internal reliability (see also the section on reliability below), and the three subscales together thus yield 24 items in total, which is typical for comparable anxiety measures in other domains (Alexander & Martray, 1989; Hopko, Mahadevan, Bare & Hunt, 2003; Suinn, Taylor & Edwards, 1988). The final scale is given in Table 3.

For norming purposes, Table 4 gives means and standard-deviations for each of the three subscales. Scores were computed by summing self-ratings (scored 0 to 4) across all 8 items in each subscale for each participant. Scores thus range from 0 to 32. Table 4 shows comparable means and standard-deviations across the subscales with means slightly below the middle of the scales' range. All three subscales were moderately positively skewed to a comparable degree.

Table 4 – Subscale Summary Statistics

Subscale	Mean	SD	Skew
M	11.80	7.62	0.45
N	12.64	7.49	0.48
I	10.73	6.43	0.45

Reliability, Interrelatedness, and Selectivity

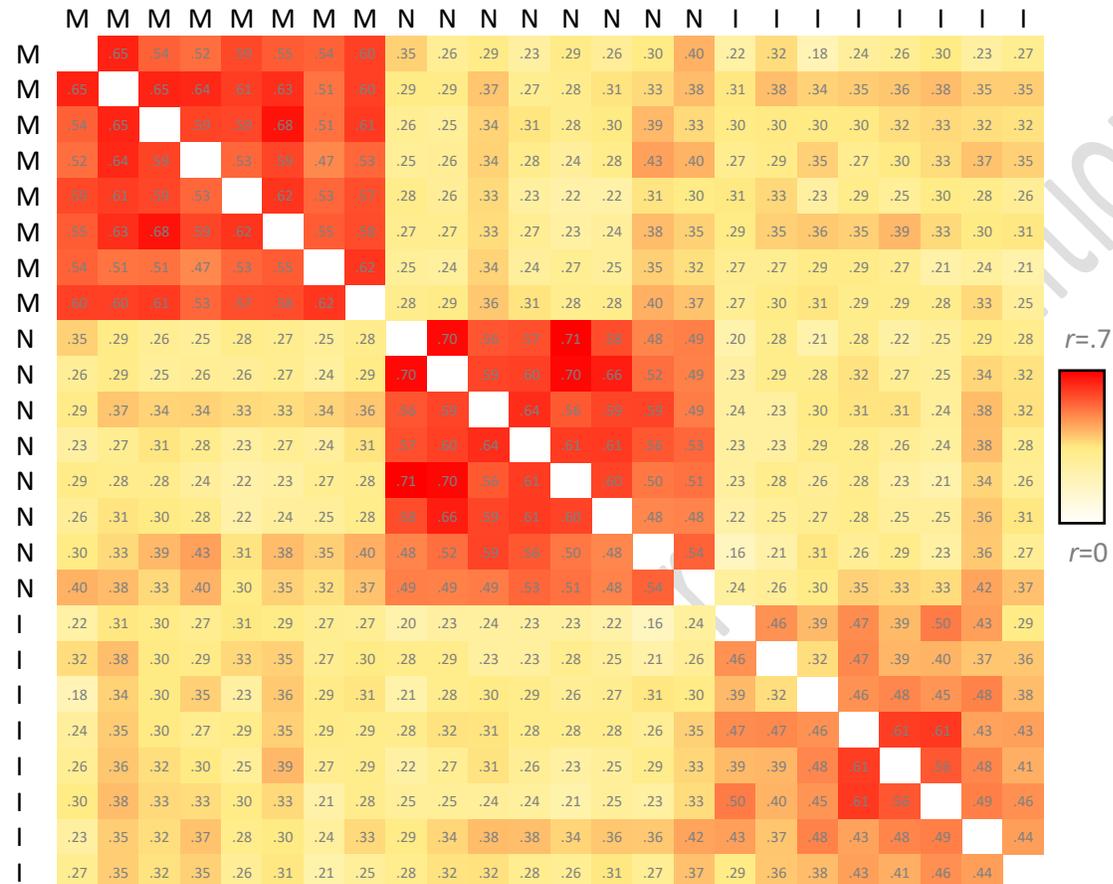
Using Cronbach's α , internal reliability was good to excellent for all three subscales (from Table 2). M: $\alpha=.917$, N: $\alpha=.914$, I: $\alpha=.862$.

The subscales were related to one another, which is perhaps unsurprising as each assesses some aspect of spatial anxiety ($M \leftrightarrow N: r=.476$, $M \leftrightarrow I: r=.526$, $N \leftrightarrow I: r=.486$, all $ps \leq 9E-27$). Note that the correlations are also not exceptionally high, indicating low likelihood of collinearity, which may prove useful when comparing the subscales in a multiple regression equation, for instance.

Finally, Figure 1 gives the item-wise correlation matrix for all 24 items in the final scale. Items from a given subscale were more related to other items from the same subscale than to items from the other subscales (all $ps \leq 2E-11$). In other words, specific items tended to show good selectivity or 'preference'

for items within their own respective subscale. In sum, the subscales were reliable, moderately interrelated, and selective.

Figure 1 – Inter-Item Correlation Matrix



Caption. Correlations among individual items from the three subscales (from Table 3). Mean r s for same-scale correlations: $M \leftrightarrow M = .579$, $N \leftrightarrow N = .569$, $I \leftrightarrow I = .443$. Mean r s for different-scale correlations: $M \leftrightarrow N = .300$, $M \leftrightarrow I = .300$, $N \leftrightarrow I = .278$.

Discussion

The goal of Study 1 was to develop a Spatial Anxiety Scale that respected the notion that spatial processing is not a unitary construct. Hence, our final scale (Table 3) includes three subscales that assess anxiety about Spatial Mental-Manipulation (M), Spatial Navigation (N) and Spatial Imagery (I). We arrived at these subscales by combining a data-driven exploratory factor analysis with a more theoretically driven view of what the main spatial skills are (taken from Uttal et al., 2013). In this way, our three subscales closely correspond to three of the Uttal et al. categories of spatial skills (M \approx intrinsic-dynamic, N \approx extrinsic-dynamic, I \approx intrinsic-static). We initially generated items intending to

capture Uttal et al.'s fourth category (extrinsic-static, corresponding to Scalar-Comparison, S, here); however, factor analyses failed to reveal a separate factor onto which these items reliably loaded. Hence, in order to respect the data-driven prong of our approach, we refrained from including a fourth subscale for this category.

The result was three spatial anxiety subscales that were reliable and selective. The subscales were also moderately related to one another which potentially reflects their common theme of anxiety about spatial processing more generally, but the moderate correlations indicate the three subscales may predict unique variance in individual's spatial performance. With that said, Study 1 did not assess the external validity of either the scale or its subscales. For this, we now turn to Study 2.

STUDY 2

The goal of Study 2 was to assess the external validity of the Spatial Anxiety Scale (and in particular its three component subscales) developed in Study 1: Spatial Mental-Manipulation (M), Spatial Navigation (N), and Spatial Imagery (I). To do so, we examined the unique relation between each subscale (controlling for the other subscales and a measure of general anxiety) and attitudes and ability in the relevant spatial sub-domain using measures previously established in the literature. To establish a more complete picture, we assessed both self-rated attitudes/abilities and actual abilities.

Methods

Participants

Participants were 251 students at the University of California, Los Angeles (UCLA). 18 participants were removed from further analysis either for failing to respond correctly to catch survey items or for below-chance performance on one of the cognitive tasks. The final N for this study was thus 233 (164 female; age: range=18.0 to 34.3yrs, $M=21.12$, $SD=2.11$).

Procedure

All procedures and materials were reviewed and approved by the UCLA Institutional Review Board (IRB). Participants completed a survey battery and three cognitive tasks. The survey battery consisted of several questionnaires presented in randomized order (with the exception that a basic demographics survey was always presented last). The order of the three tasks and survey battery were counterbalanced across participants. Detailed descriptions of the surveys and tasks are given below. Participants were given 1.5 research participation credits upon completion of the approximately 90 minute study. Up to three participants were tested per session in their own individual work station.

Participants were told that the goal of the study was to examine cognitive performance. To de-emphasize social comparison, we also told participants that everyone would be completing a different set of tasks so they should ignore the fact that some students might finish before they do. All participants were also given headphones to block out extraneous noise.

Surveys

The surveys of primary interest were the Spatial Anxiety Scale developed in Study 1 and self-rated attitudes/ability scales for the three spatial sub-domains: Mental-Manipulation (M), Navigation (N), and Imagery (I). To remove variance due to general anxiety, we measured trait-anxiety. Several filler questionnaires of no interest were included to mask the nature of the study. Catch trials (e.g., “Select ‘not at all’”) were included at random intervals to ensure participants were indeed considering each item. If participants missed more than one catch item, they were excluded from further analysis. Descriptive statistics for the surveys can be found in Table 5. For previously established surveys, data were scored according to published norms.

Spatial Anxiety Scale (M, N, I)

This scale was nearly identical to that developed in Study 1 (Table 3). The only difference was that, due to a computation error, two of the items in the Mental-Manipulation (M) subscale differed. Specifically, items M04 and S09 were replaced with M06 and S01 (see Table A-1). However, given very similar factor loadings (.721 and .720; .714 and .712, respectively; see Table 2), the fact that the two versions of the subscale were correlated at $r=.97$, and that other results from Study 1 were nearly identical if one used M06 and S01 instead of M04 and S09, we believe that results from Study 2 should be highly indicative of the external reliability of the final version of the M-subscale given in Table 3. As in Study 1, item-order was randomized across participants regardless of subscale to prevent thematic grouping. Scores for each subscale ranged from 0 to 32, with a higher score indicating higher anxiety.

Self-Rated Attitude and Ability Scales (OSIQ-S, SBS-D, OSIQ-O)

To assess self-rated Navigation attitudes and ability, we used the Santa Barbara Sense of Direction Scale (SBS-D; Hegarty et al., 2002). This scale consists of 15 items pertaining to navigational abilities [examples: “My ‘sense of direction’ is very good”, “I very easily get lost in a new city” (reverse-coded)]. In keeping with the scale’s original design, participants respond on a 1-7 scale (strongly disagree to strongly agree). Responses are then averaged across items to give a participant’s final score (range: 1-7), with a higher score corresponding to higher self-rated attitude/ability. In the validation study, the test–retest reliability of the SBS-D was .91 (Hegarty et al., 2002).

To assess self-rated Mental-Manipulation and Imagery attitudes/abilities, we used the Object-Spatial Imagery Questionnaire (OSIQ; Blajenkova et al., 2006). The OSIQ, in fact, consists of two separate parts: 'object' and 'spatial' factors (15 items of each). Examples of 'object' items are, "My images are very vivid and photographic", "When I imagine the face of a friend, I have a perfectly clear and bright image". Examples of 'spatial' items are, "I can easily imagine and mentally rotate 3-dimensional geometric figures", "I have excellent abilities in technical graphics". Hence, for present purposes, we treated the 'object' component as self-rated Imagery attitude/ability (OSIQ-O), and we treated the 'spatial' component as self-rated Mental-Manipulation attitude/ability (OSIQ-S). In keeping with the scale's original design, participants responded on a 1-5 scale (strongly disagree to strongly agree). Responses were then summed across items to give a participant's final score (range: 15-75), with a higher score corresponding to higher self-rated attitude/ability. Cronbach's alpha for the spatial scale in the validation study was .83. Cronbach's alpha for the Object scale in the validation study was .79. Both scales have an excellent one-week test-retest reliability, ($r=.81$ for the spatial scale and $r=.95$ for the object scale; Blajenkova et al., 2006).

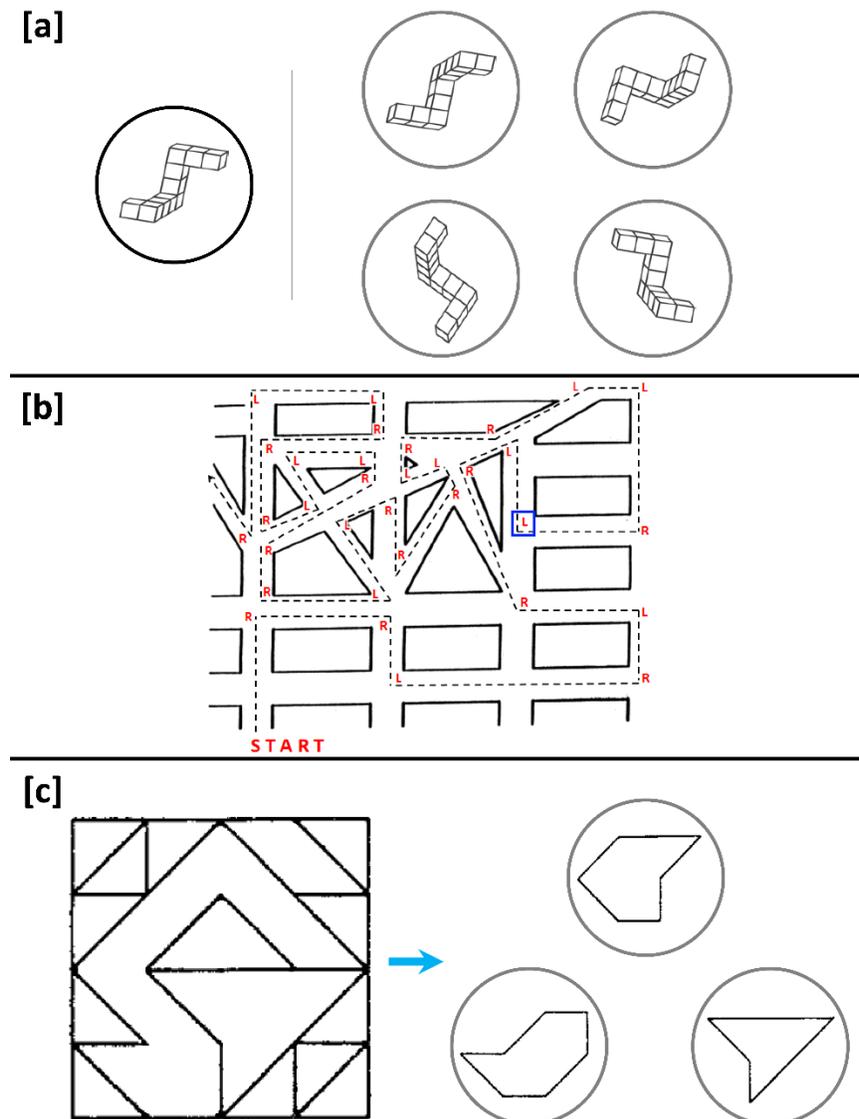
General Trait Anxiety (STAI)

To control for general anxiety, we used the 'trait' component of the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970). The trait scale consists of 20 items and in the scale's instructions participants are encouraged to indicate how they *generally* feel [examples: "I feel that difficulties are piling up so that I cannot overcome them", "I feel satisfied with myself" (reverse-coded)]. In keeping with the scale's original design, participants respond on a 1-4 scale (almost never to almost always). Responses are then summed across items to give a participant's final score (range: 20-80); hence, a higher score corresponds to higher general anxiety. Note that this scale was intended to serve as a control measure.

Tasks

Descriptive statistics for all tasks can be found in Table 5. Task Examples are shown in Figure 2.

Figure 2



Caption. Figure 2 shows examples of the three behavioral tasks.

[a] Mental-Manipulation (MRT): participants determine which of the two figures in grey circles can be rotated to match the standard (black circle).

[b] Navigation (MapNav): The task proceeds along the route marked with the dashed line by highlighting (blue square) each turn in succession. Participants determine if the highlighted marker indicates the correct direction the route takes, from the perspective of a person following that route as if it were a street-plan.

[c] Imagery (EmbFig): Participants are shown the complex figure on the left for several seconds, after which it disappears and is replaced by the three options in grey circles. Participants indicate which of the three options was embedded within the prior complex shape.

Mental-Manipulation Ability (MRT)

To assess objective Mental-Manipulation ability, we used a standard Mental Rotation task (MRT; Weisberg, et al., 2014). In this version of the task, on the left-hand side of the screen, participants saw a line drawing of an abstract three-dimensional figure comprised of concatenated cubes. Participants were also shown 4 probe figures containing similar figures on the right. Two of the probe figures were the same as the figure on the left, just rotated in space; two were foils. Participants’ task was to determine which 2 of the 4 probe figures were the same as the first (left-most) figure, just having been

rotated. The task began with instructions and three practice trials that were untimed. Participants received feedback if they choose the incorrect probe figures. After the practice trials were over, participants were presented with two blocks of trials. Each block presented participants with 10 trials and gave them 3 minutes to solve as many trials as possible. Participants were encouraged to work as quickly as possible without sacrificing accuracy. To control for guessing, hit-rates (H) and false-alarm-rates (FA) were computed across all trials for each participant. H and FA were then used to compute d' (or 'sensitivity') estimates via the formula $d' = Z(H) - Z(FA)$, where $Z(x)$ corresponds to the inverse of the cumulative (Gaussian) distribution function (Stanislaw & Todorov, 1999). A higher value of d' indicates better performance. An example trial is shown in Figure 2a.

Navigation Ability (MapNav)

A revised computerized version of the Money Road Test was used to validate our Navigation anxiety subscale (Ferguson et al., 2015; Money et al., 1965). In this version of the task, participants are presented with an on-screen image of a street map with a walking route marked by a dashed line. The walking route meanders across the map and makes various left and right turns. Each turn is labeled with either an "R" or "L" to indicate a right or left turn. However, not all "R" and "L" labels correctly correspond with the turn that was taken. Participants are instructed to imagine that they are walking along the path and to click the "Y" button if the label corresponds with the actual direction of the turn taken, and the "N" button if the label does not correspond with the actual turn taken. The task began with instructions and a set of practice problems to familiarize participants with the response format. Once participants were ready for the main trials, they were instructed to respond as quickly and accurately as possible. The test contained 33 turns with 10 turns labeled incorrectly.

Accuracy (in the form of error-rates (ERs), as the proportion incorrect) and response times (RTs, in milliseconds) were recorded on each trial. To correct for potential speed-accuracy trade-offs, and to reduce the number of statistical comparisons that needed to be run (thereby reducing the risk of false-positives), overall performance was computed by combining ERs and RTs via the formula $P = RT(1 + K \cdot ER)$, where K is the number of response options (in this case 2) (Lyons et al., 2014). This formula linearly weights RTs as a function of ERs on a scale from actual RT (in the case of 0% errors) to $2x$ RT (in the case of chance performance); hence, a higher score (P) indicates worse performance. An example trial is shown in Figure 2b. Further, test re-test reliability for this instrument is acceptable ($r = .72$; Ferguson et al., 2015).

Imagery Ability (Pictures).

To assess Spatial-Imagery ability, we adapted an Embedded Figures task from Ekstrom et al. (1976). Participants were shown a complex two-dimensional line drawing for six seconds and told to commit it to memory as best as possible. After the initial drawing disappeared (followed by a brief visual mask), three simpler line figures were shown. Participants' task was to identify which of the three simpler figures was part of (i.e., 'embedded in') the line drawing from a moment before. Participants had 10 seconds to respond. Note that there was a fixed set of 5 simpler line figures, and the 3 candidates on a given trial were always drawn from this set; moreover, participants were given time to familiarize themselves with the set during instructions. Participants completed a total of 30 trials (inter-trial-interval = 1000msec). Accuracy (in the form of error-rates (ERs), as proportion incorrect) and response times (RTs, in milliseconds) were recorded on each trial. Performance (P) was computed in the same manner as the Navigation task above, with the exception that the number of response options (K) was 3 in this case. A higher score (P) indicated worse performance. An example trial is shown in Figure 2c.

Results

Descriptive statistics and the correlation matrix for all variables are given in Tables 5 and 6, respectively. The main goal of Study 2 was to establish external validity for each of the three Spatial Anxiety subscales. To do so, we used previously established measures to assess self-rated attitudes/ability and objective ability thought to correspond to the particular type of spatial processing addressed by each anxiety subscale. We thus had 6 spatial assessments – a self-report and objective score that corresponded to each of the three subscales. In the results that follow, we regressed each of the 6 spatial assessments on the three spatial anxiety subscales, as well as general trait anxiety as a control measure (i.e., STAI). In this way, we were able to determine the unique relation between a given anxiety subscale and performance in the corresponding subdomain task (controlling for the contributions of the other two anxiety subscales and general anxiety).

Table 5 – Descriptive Statistics

	Measure	Mean	SD	Skew
Anxiety	M	9.90	6.88	.62
	N	15.21	6.78	.32
	I	10.26	5.50	.47
Self-Ratings	OSIQ-S	41.09	7.94	.19
	SBSD	3.88	1.15	.08
	OSIQ-O	50.68	7.81	-.30
Ability	MRT ⁺ (<i>d'</i>)	1.85	1.35	.30
	MapNav ⁻ (<i>P</i>)	2934.15	2047.37	2.83
	EmbFig ⁻ (<i>P</i>)	2619.80	1108.08	1.58
	STAI	47.70	11.32	.39

Table 6 – Correlation Matrix

	M	N	I	OSIQ-S	SBSD	OSIQ-O	MRT	Map Nav	Emb Fig	STAI
M		1E-07	1E-24	3E-06	.027	.604	6E-08	8E-04	.008	8E-05
N	.327		6E-07	9E-07	4E-24	.390	3E-04	.004	.135	7E-08
I	.588	.309		.898	.848	.019	.017	.837	.782	4E-06
OSIQ-S	-.291	-.304	.008		3E-12	.010	2E-04	.002	.292	9E-05
SBSD	-.140	-.581	-.012	.422		.060	7E-05	6E-06	.040	7E-05
OSIQ-O	-.033	-.054	-.148	-.163	.119		.877	.629	.049	.624
MRT	-.334	-.226	-.151	.233	.249	.010		1E-09	.002	.997
MapNav	.211	.183	.013	-.198	-.282	-.031	-.372		.025	.502
EmbFig	.168	.095	.018	-.067	-.130	-.124	-.192	.142		.935
STAI	.246	.333	.285	-.244	-.248	-.031	.000	.043	.005	

Note. Below the diagonal are (zero-order) *r*-values; above the diagonal are *p*-values.

Verification – Factor Analysis and Reliability

Here we sought to replicate the factor loading structure identified in Study 1. Specifically, we tested whether items from the three different subscales loaded onto different factors, and whether items from the same subscale loaded most strongly on the same factor. The factor analysis was conducted using the same parameters as in Study 1, with the exception that only the 24 items from the final scale (Table 3) were included, and the rotated solution was limited to 3 factors (corresponding to the three subscales). Results were highly consistent with the notion that (1) items from different subscales load onto separate factors, and (2) items from the same subscale load onto the same factor. Consistent with this, internal reliability was good for all three subscales: M: $\alpha=.877$, N: $\alpha=.864$, I: $\alpha=.810$. In sum, both the underlying factor structure and internal reliabilities from Study 2 replicated the overall pattern of results from Study 1, and thus provide further support for the notion that the three subscales should be treated as measuring anxiety about separate aspects of spatial processing.

Table 7 – Rotated Factor Loadings

Item	Factor1	Factor2	Factor3
M	.731		
M	.780		
M	.651		
M	.563		
M	.785		
M	.442		
M	.637		
M	.628		
N		.716	
N		.678	
N		.646	
N		.625	
N		.673	
N		.618	
N		.678	
N		.632	
I			.663
I			.753
I			.586
I			.475
I			.411
I			.429
I			.618
I			.411

Note. Table 7 shows rotated factor loadings for the 3-factor solution for the 24 items included in the final version of the scale separated into the three subscales. Loadings whose absolute value are less than .4 are suppressed. Items are listed in the same order as Table 3.

Spatial Mental-Manipulation Anxiety (M)

We hypothesized that higher M-anxiety would uniquely predict lower M-attitude/ability self-ratings; because higher scores indicated higher self-ratings, we expected a *negative* relation. We also hypothesized that higher M-anxiety would predict lower M-performance; because a higher score on this task indicated better performance, we expected a *negative* relation. Multiple regression results (Table 8a-b, with the most relevant results highlighted in blue) were consistent with both hypotheses.

Table 8 – Regression Results (M)

(a) DV: OSIQ-S						(b) DV: MRT					
Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>	Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>
M	-.441	.084	-5.28	3E-07	-.330	M	-.062	.015	-4.14	5E-05	-.265
N	-.238	.076	-3.15	.002	-.204	N	-.036	.014	-2.69	.008	-.175
I	.495	.105	4.73	4E-06	.299	I	.015	.019	0.79	.431	.052
STAI	-.128	.044	-2.87	.005	-.187	STAI	.016	.008	2.03	.044	.133

Note: Overall model fits: (a) adjusted $R^2=.207$, $p=1E-11$; (b) adjusted $R^2=.116$, $p=2E-06$. r_p : partial correlation.

Spatial Navigation Anxiety (N)

We hypothesized that higher-N anxiety would uniquely predict lower N- attitude/ability self-ratings; because higher scores indicated higher self-ratings, we expected a *negative* relation. We also hypothesized that higher N-anxiety would predict lower N-performance; because a higher score on this task indicated worse performance, we expected a *positive* relation. Multiple regression results (Table 9a-b, with the most relevant results highlighted in orange) were consistent with both hypotheses.

Table 9 – Regression Results (N)

(a) DV: SBSO						(b) DV: MRT					
Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>	Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>
M	-.009	.011	-0.82	.414	-.054	M	87.19	23.31	3.74	2E-04	.240
N	-.098	.010	-9.85	3E-19	-.546	N	43.50	21.07	2.06	.040	.135
I	.047	.014	3.38	9E-04	.180	I	-68.74	29.25	-2.35	.020	-.154
STAI	-.011	.006	-1.80	.073	-.118	STAI	-5.14	12.40	-0.41	.679	-.027

Note: Overall model fits: (a) adjusted $R^2=.343$, $p=9E-21$; (b) adjusted $R^2 = .071$, $p=3E-04$. r_p : partial correlation.

Spatial Imagery Anxiety (I)

We hypothesized that higher I-anxiety would uniquely predict lower I-ability self-ratings; because higher scores indicated higher self-ratings, we expected a *negative* relation. We also hypothesized that higher I-anxiety would predict lower I-performance; because a higher score on this task indicated worse performance, we expected a *positive* relation. Multiple regression results (Table 10a-b, with the most relevant results highlighted in yellow) were consistent with only the former hypothesis: I-anxiety uniquely predicted I-attitude/ability ratings, but not objective I-performance at the traditional alpha level of .05.

Table 10 – Regression Results (I)

(a) DV: SBSO						(b) DV: MRT					
Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>	Predictor	<i>b</i>	<i>se</i>	<i>t</i>	<i>p</i>	<i>r_p</i>
M	.097	.092	1.05	.293	.070	M	39.54	12.85	3.08	.002	.200
N	-.048	.083	-0.58	.561	-.039	N	14.13	11.62	1.22	.225	.080
I	-.258	.115	-2.24	.026	-.147	I	-27.46	16.13	-1.70	.090	-.112
STAI	.033	.049	0.67	.506	.044	STAI	-5.16	6.84	-0.75	.452	-.050

Note: Overall model fits: (a) adjusted $R^2 = .008$, $p = .215$ ($p = .042$ if only I-anxiety is included as a predictor); (b) adjusted $R^2 = .035$, $p = .016$. r_p : partial correlation.

Gender Effects

We tested for gender differences on each of the three spatial anxiety subscales. Women showed significantly higher anxiety ratings for M-anxiety [$t(231) = 2.56$, $p = .011$, $d = .328$; Women: 10.64 ($se = .55$), Men: 8.14 (.74)] and N-anxiety [$t(231) = 2.15$, $p = .033$, $d = .273$; Women: 15.83 (.55), Men: 13.75 (.71)], but not I-anxiety [$t(231) = 0.18$, $p = .860$, $d = .018$; Women: 10.30 (.44), Men: 10.16 (.63)].

Percentiles

For future norming purposes, scores corresponding to percentiles (in quintiles) are given for each subscale in Table 11.

Table 11 – Percentiles

Percentile	M	N	I
20	3.0	9.0	5.0
40	7.6	12.6	9.0
60	11.0	16.4	11.0
80	15.0	21.2	15.0

Discussion

The main goal of Study 2 was to establish external validity for each of the three Spatial Anxiety subscales. All three subscales showed the predicted (unique) negative relation with established measures of self-rated ability/attitudes in their respective subdomain of spatial processing. Two of the subscales (M- and N-anxiety) also showed the predicted relation with objective performance in the relevant spatial subdomain, wherein higher anxiety corresponded to lower performance. These results thus suggest good external validity for the M- and N-anxiety subscales. External validity was less definitive for I-anxiety, with the predicted relation obtaining significance in the case of self-rated attitude/ability but only marginal significance in the case of objective ability. However, one may note that effect-sizes (partial- r s) were relatively similar for the unique relations between I-anxiety and imagery ability/attitudes ($-.147$) and between I-anxiety and objective imagery ability ($-.112$). Additionally, the lower relation between I-anxiety and objective performance may have been due to problems in task selection; that is, the Embedded Figures task may not be the best measure of spatial imagery. We take this and other issues up further in the General Discussion below.

GENERAL DISCUSSION

Spatial skills are an important component of STEM success. Yet, some individuals may be reluctant to engage in spatial related mental activities, in part, because they are made anxious by such experiences. We sought to develop and validate a tool to measure individual differences in spatial anxiety. To respect the well-established notion that spatial processing is not monolithic (i.e., it very likely comprises multiple disparate sub-skills), we developed spatial anxiety subscales in correspondence with a prominent theory-driven framework of spatial abilities (Uttal et al., 2013). This theory-driven approach was supplemented by a more data-driven approach, wherein we let the data determine precisely which factors comprised the resultant subscales and the specific items that comprised each subscale. Specifically, the factor analyses from Study 1 revealed that items loaded on three factors that corresponded well with some of the most common spatial abilities that have been discussed in the broader literature (i.e., navigation, manipulation and imagery), including three of the four domains outlined by Uttal et al. (2013). Internal reliability and between-scale selectivity was high; moreover, external validity was good for two of the subscales (M and N) and moderate for the third (I). The result is

an empirically validated Spatial Anxiety scale that also respects the variegated nature of spatial processing (Table 3). We discuss several points of consideration and potential limitations below.

One point of consideration is that factor analyses in Study 1 led to retention of just 3 of the 4 categories of spatial processing proposed by Uttal et al. (2013). Our initial set of items was generated based on the 2x2 conceptual matrix for understanding the breakdown of different types of spatial skills by Uttal et al., (2013). In that framework, spatial skills are classified as a factorial combination of static/dynamic and intrinsic/extrinsic factors. While we expected our items to fall into these four a priori defined categories, our data-driven approach led to the 'loss' of the 'extrinsic-static' category. This was driven by the fact that these extrinsic-static ('S') items largely loaded on the manipulation and navigation factors (Table 2). Note that this result remained unchanged even if a 5th candidate factor were added to the rotated model solution. Thus, at least with respect to spatial anxiety, it appears that the extrinsic-static category is largely indistinguishable from (anxiety about) manipulation and navigation. It is for this reason that we omitted this category from our final scale (though it is perhaps interesting to note that several items initially labeled as 'S' items did make it into the final M and N subscales). Here it is important to point out that we do not see this as a confirmation or rejection of the four-category framework proposed by Uttal et al. (doing so was not the aim of this paper); instead, we merely used that framework as an initial jumping-off point. Future work might be aimed specifically at developing an anxiety scale that more closely matches the Uttal et al. framework.

With respect to self-rated attitude/ability ratings, external validity (using rating scales previously established elsewhere in the literature) was excellent. One would expect that higher anxiety about a given domain should be related to lower ability/attitude ratings, which is precisely what we found. Each anxiety subscale (M, N, I) was a significant unique negative predictor of anxiety/ability ratings in its respective category (Tables 8a, 9a, 10a). With respect to objective ability scores, external validity was good, albeit perhaps not as strong as the ability/attitude ratings. For mental-manipulation, M-anxiety was indeed a highly significant unique predictor of lower MRT (M-ability) performance (Table 8b). For navigation, N-anxiety was a statistically significant unique predictor of lower performance on the Map-Navigation task (Table 9b). It is worth noting that M-anxiety was in fact also a significant predictor of poor performance on this task. This may be due in part to the fact that we used a computer-based in-lab task to assess navigation ability, which may have lent to greater prevalence of mental-manipulation strategies being employed by some participants. A more active task where a person is asked to actually navigate a real (or virtual) space in future studies might show a stronger relation with N-anxiety. That

said, we should emphasize that, despite this potential concern, N-anxiety was nevertheless a significant unique predictor of poor performance on the Map-Navigation task, over and above the contribution of M-anxiety. Finally, it is important to point out that I-anxiety was only marginally significantly ($p=.09$) related to performance on the imagery ability (Embedded-Figures) task (Table 10b). We discuss potential reasons for this in the paragraph below. In sum, all three subscales showed acceptable to good validity with respect to anxiety/ability ratings; however, with respect to objective ability, while M-anxiety showed good external validity and N-anxiety showed acceptable validity (with the important caveat that task-selection may have influenced the results), I-anxiety fell just short. These results may be useful to consider when employing the various anxiety subscales in future studies.

Our spatial imagery anxiety (I-anxiety) subscale uniquely predicted lower imagery ability/attitude self-ratings, but it did so only marginally for our measure of objective spatial imagery ability (Embedded-Figures task). One point worth noting is that the effect-sizes for the unique relations between I-anxiety and imagery ability/attitudes ($-.147$) and between I-anxiety and objective imagery ability ($-.112$) were quite similar, falling just on either side of the arbitrary significance threshold of $.05$. With a slightly larger sample-size, both effects may have been significant at the traditional threshold. That said, the overall relatively small partial-correlations seen for the I-anxiety subscale may be the result of our I-anxiety scale being a suboptimal measure of anxiety about spatial imagery processing. However, we do not believe this to be the case given (1) the clear imagery-related nature of the items (see Table 3), (2) that these items all loaded on a factor separate from the M and N factors (Table 2); and (3) the fact that I-anxiety ratings *did* uniquely predict lower spatial imagery ability/attitude self-ratings using an established measure from elsewhere in the literature (the 'I' portion of the OSIQ; Blajenkova et al., 2006). Another possibility is that the task we selected – the Embedded-Figures task – is not an ideal measure of spatial imagery ability. We, admittedly, had difficulty identifying a well-established task that provides a relatively pure measure of mental imagery ability. The Embedded-Figures task, while measuring imagery ability, also draws heavily on more domain-specific short-term-memory processes, which could have reduced our ability to assess imagery performance specifically. Furthermore, there may be a difference between the vividness of one's mental images and their preference for representing and processing colorful pictorial images of individual objects. While they may be related, these are two distinct constructs. For example, in Blajenkova et al., (2006), the object subscale of the OSIQ and the Vividness of Visual Imagery Questionnaire, which measures how vivid one's mental imagery is (VVIQ; Marks, 1973) were only correlated at $r=.48$. Thus, performance on the Embedded-Figures task may be more related to the quality of one's imagery rather than their propensity to use such imagery. This is an

empirical question that future research could address. Furthermore, as is evidenced by the strong relation between performance on the Embedded Figures Task and M-anxiety, it may involve a strong mental manipulation component. As such, though we believe the I-anxiety subscale to be reliable and valid with respect to self-reported spatial imagery attitudes, we take a more reserved position with respect to objective spatial imagery ability. More broadly, we suggest the literature would be well-served by the development of a task that provides a more distilled measure of spatial imagery abilities.

In the current study the zero-order correlations for the I-anxiety subscale does not correlate with the OSIQ-S, SBS-D, and MapNav (see Table 6). In the multiple regressions the I-anxiety subscale is a strong and significant predictor of all these variables but in the negative direction. While this may, at first, appear problematic, it is actually in line with previous research. Indeed, considerable cognitive and neuroscience research (e.g., Farah, Hammond, Levine, & Calvanio, 1988; Kosslyn, 1994; Kosslyn & Koenig, 1992; Levine, Warach, & Farah, 1985) suggests that mental imagery is not a unitary construct, and instead argues that there are two distinct object and spatial imagery subsystems that encode and process visual information in different ways. For instance, the OSIQ spatial scale has been found to be significantly correlated with the Paper Folding Test and with the Vandenberg-Kuse Mental Rotation Test, but not the Degraded Pictures Test. Further, the OSIQ object scale was significantly correlated with the Degraded Pictures Test and was not significantly correlated with either the Paper Folding Test or the Vandenberg-Kuse Mental Rotation Test. Perhaps most important, the spatial and object sub-scales of the OSIQ have been found to be either uncorrelated (Blajenkova et al., 2006 Study, 2a) or negatively correlated with one another (Blajenkova et al., 2006, Study 2b).

The moderate correlations among the three subscales (.3 to .6 range), along with the multiple regression results (Tables 8-10), suggest that the subscales can – and perhaps should – be treated separately. Using the scales separately may be particularly important for researchers interested in understanding how anxiety may diminish spatial ability and why some individuals do not respond to spatial training. Indeed, in a recent and comprehensive review of the influence of training on spatial thinking, Uttal et al. (2013) concluded that “spatial skills are highly malleable, and that training in spatial thinking is effective, durable, and transferable” (p. 365). This finding is particularly encouraging given the importance of spatial ability for success in STEM fields. This is also welcome news for researchers interested in the remediation of spatial anxiety. Specifically, it will be important for spatial anxiety remediation techniques to target the relevant sub-domain of spatial processing.

We should also make clear that the work reported here is correlational in nature; thus, the causal direction of the spatial anxiety and decreased spatial ability relation is not yet clear. Poor spatial skills may pre-dispose one to develop spatial anxiety. Although we find that spatial anxiety predicts performance above and beyond self-reported ability ratings but this result in of itself is not enough to establish a causal direction. Follow up research should examine whether early performance around spatial relevant tasks contributes to growth in spatial anxiety or whether spatial anxiety derails the student performance outcomes for spatial tasks. Recent research in the comparable domain of math anxiety (e.g., Foley, Herts, Borgonovi, Guerriero, Levine, & Beilock, 2017), leads us to theorize that there is likely a bi-directional relation between spatial anxiety and performance on spatial tasks with some work suggesting the ability to anxiety relation is likely stronger than the anxiety to ability relation (Gunderson, Park, Maloney, Beilock, & Levine, 2017; Ma & Xu, 2004; Ramirez, et al., 2017; Ramirez, Shaw & Maloney, 2017).

Regardless, the central focus of the current paper was to develop a tool that measures the various facets of spatial anxiety. We hope that this tool will prove useful for future research that aims to unpack the causal relation(s) between spatial anxiety, spatial ability, and attitudes about spatial situations in a manner that also respects the variegated nature of spatial processing.

There exists a number of limitations that should be acknowledge. For instance, a rule of thumb in studies attempting to validate scales is to collect 10 participants per item. Despite our large sample size we did not to follow this rule of thumb. We also did not measure state anxiety which could have provided a clearer picture of how spatial anxiety manifests itself in the moment. An addition limitation is that these results are limited to adults in North America. As such, an interesting future direction is to examine spatial anxiety as a function of various demographic factors, including age and geographical location. Our inclusion of the percentile norms for these data can serve to facilitate these future studies.

Conclusion

To conclude, spatial ability is an important domain of general cognitive factors predicting STEM entry and retention. Researchers have extensively studied how hormonal variation and environmental exposure contribute towards individual differences in spatial ability. However, the existing literature lacks an empirically reliable and validated scale for measuring spatial anxiety that also respects the widely-held view that spatial processing can and should be treated as comprising multiple sub-domains. By combining theory- and data-driven techniques, we developed a set of three spatial anxiety subscales to address this gap. Moreover, we showed that the majority of the subscales possess good and selective

reliability. We believe this scale may be of considerable value to researchers and education stakeholders interested in addressing the affective factors predictive of spatial ability.

in press – J Numerical Cognition

REFERENCES

- Alexander, L., & Martray, C. R. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development*, 22(3), 143-150.
- Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex roles*, 20(5), 327-344.
- Berry, J.W. (1966). Temne and Eskimo Perceptual Skills. *International Journal of Psychology*, 1(3), 207-229.
- Blajenkova O, Kozhevnikov M, Motes M. A. (2006). Object-Spatial Imagery: A New Self-Report Imagery Questionnaire. *Applied Cognition Psychology*, 20:239-63.
- Brinkmann, E. H. (1966). Programed instruction as a technique for improving spatial visualization. *Journal of Applied Psychology*, 50(2), 179-184.
- Benbow, C., & Stanley, J. (1982). Intellectually talented boys and girls: Educational profiles. *Gifted Child Quarterly*, 26, 82-88.
- Bronzaft, A. L., Dobrow, S. B., & O'Hanlon, T. J. (1976). Spatial orientation in a subway system. *Environment and Behavior*, 8, 575-594.
- Carroll, J. B. (1993). Human cognitive abilities: A survey of factor analytic studies. New York: Cambridge University Press Cambridge.
- Eliot, J. (1987). Models of psychological space: Psychometric, developmental and experimental approaches. New York, NY: Springer-Verlag. Eliot, J., & Fralley, J. S. (1976). Sex differences in spatial abilities. *Young Children*, 31(6), 487-498.
- Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service: Princeton, New Jersey.
- Foley, A. E., Herts, J. B., Borgonovi, Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link: A global phenomenon. *Current Directions in Psychological Science*, 26, 52-58.
- Hegarty, M., Keehner, M., Cohen, C. A., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In G. L. Allen (Ed.), *Applied spatial cognition: From research to cognitive technology*, chapter 11 (pp. 285-315). Mahwah, NJ: Lawrence Erlbaum.
- Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D. R. (2009). How spatial abilities enhance, and are enhanced by, dental education. *Learning and Individual Differences*, 19(1), 61-70.
- Hegarty M, Richardson A.E., Montello D.R., Lovelace, K. & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425-48.
- Hopko, D. R., Mahadevan, R., Bare, R. L., and Hunt, M. K. (2003). The abbreviated math anxiety scale (AMAS) construction, validity, and reliability. *Assessment*, 10:178-182.
- Ferguson, A. M., Maloney, E. A., Fugelsang, J., & Risko, E. F. (2015). On the relation between math and spatial ability: The case of math anxiety. *Learning and Individual Differences*, 39, 1-12. Chicago
- Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences* (Tenth-anniversary ed.). New York: Basic Books.

- Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2013). Teachers' spatial anxiety relates to 1st – and 2nd – graders' spatial learning. *Mind, Brain and Education*, 7(3), 196-199.
- Gunderson, E., Park, D., Maloney, E.A., Beilock, S.L. & Levine, S.C. (2017). *Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school*. Manuscript submitted for publication.
- Hambrick, D., & Meinz, E. (2011). Limits on the predictive power of domain-specific knowledge and experience for complex cognition. *Current Directions in Psychological Science*, 20(5), 275-279.
- Hoffman, M., Gneezy, U., & List, J. A. (2011). Nurture affects gender differences in spatial abilities. *Proceedings of the National Academy of Sciences*, 108(36), 14786-14788.
- Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology*, 4(2), 277-299. Chicago
- Kremmyda, O., Hübner, K., Flanagan, V. L., Hamilton, D. A., Linn, J., Strupp, M., ... & Brandt, T. (2016). Beyond dizziness: virtual navigation, spatial anxiety and hippocampal volume in bilateral vestibulopathy. *Frontiers in human neuroscience*, 10.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5), 745-756.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20(1), 47-77.
- Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & cognition*, 33(4), 710-726.
- Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, 30(11-12), 765-779.
- Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Developmental psychology*, 48(2), 530-542.
- Linn, M. C. & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 56(6), 1479-1498.
- Lohman, D. F. (1988). Spatial abilities as traits, processes and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 181–248). Hillsdale, NJ: Erlbaum
- Lord, T. R. (1985). Enhancing the visuo-spatial aptitude of students. *Journal of Research in Science Teaching*, 22(5), 395-405.
- Lyons I.M., Price G.R., Vaessen, A., Blomert, L. & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, 17(5):714-26.
- Ma, X., & Xu, J. (2004). The causal ordering of mathematics anxiety and mathematics achievement: A longitudinal panel analysis. *Journal of Adolescence*, 27, 2, 165-79.
- Money, J., Alexander, D., & Walker, H. T. (1965). *A standardized road-map test of direction sense: Manual*. Johns Hopkins Press.
- Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2012). Spatial anxiety relates to spatial abilities as a function of working memory in children. *The Quarterly Journal of Experimental Psychology*, 65(3), 474-487.
- Ramirez, G., Fries, Gunderson, E., Schaeffer, M.W., Maloney, E. A., Beilock S. L. & Levine,

- S.C. (2017). *Reading Anxiety – An early affective impediment to children’s success in reading*. Manuscript submitted for publication.
- Ramirez, G., Shah, S., & Maloney, E. A. (2017). *Math Anxiety: Past research, promising interventions and a new interpretation framework*. Manuscript submitted for publication.
- Russell-Gebbett, J. (1985). Skills and strategies—pupils' approaches to three-dimensional problems in biology. *Journal of Biological Education*, 19(4), 293-298.
- Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93(3), 604-614.
- Sorby, S. A. (2009). Educational Research in Developing 3-D Spatial Skills for Engineering Students. *International Journal of Science Education*, 31(3), 459-480.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Stanislaw H., & Todorov N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1):137-49.
- Suinn, R. M., Taylor, S., & Edwards, R. W. (1988). Suinn Mathematics Anxiety Rating Scale for Elementary School Students (MARS-E): Psychometric and Normative Data. *Educational and Psychological Measurement*, 48, 979-986.
- Terlecki, M., Newcombe, N., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology*, 22, 996-1013.
- Thurstone, L. L. (1947). *Multiple-factor analysis*. University of Chicago Press.
- Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how? In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 57, pp. 147– 181). San Diego, CA: Academic Press.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2012). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 138, 352–402.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817-835.

SUPPLEMENTARY MATERIALS

Full Datasets for both studies can be found at [\[doi to be supplied\]](#).

in press – J Numerical Cognition

APPENDIX A

Table A-1

Category	Item Text
I (01)	Asked to recreate your favorite artist's signature from memory
I (02)	Tested on your ability to create a drawing or painting that reproduces the details of a photograph as precisely as possible
I (03)	Asked to imagine and describe the appearance of a radio announcer or someone you've never actually seen
I (04)	Asked to distinguish and identify a specific bird in a flock of birds of the same species
I (05)	Playing a competitive game where one is asked to recall a visual detail about a scene that others are unlikely to have noticed
I (06)	Asked to draw a detailed picture of a key landmark in your neighborhood
I (07)	Asked to give a detailed description of a person's face whom you've only met once
I (08)	Tested on your ability to recall the layout of a standard keyboard as quickly as possible, without looking at an actual keyboard
I (09)	Asked to recall the exact details of a relative's face whom you have not seen in several years
I (10)	Asked to recall the shade and pattern of a person's tie you met for the first time the previous evening
I (11)	Describing in detail the cover of a book to a bookseller because you've forgotten both the title and author of the book
I (12)	Asked to draw the front of an imaginary building
I (13)	Asked to scan a complex visual scene for a specific item
I (14)	Trying to recall where you left your cell phone
IM (01)	Attending an architectural design course
IS (01)	Recalling from memory where the electrical sockets are in the adjacent room to avoid cutting any wires when drilling into the wall
IS (02)	Required to recreate a diagram of a circuit board from memory in front of a classroom
IS (03)	Asked to redraw a map from memory
IS (04)	Given a test in which you were allowed to look at and memorize a picture for a few minutes, and then given a new, similar picture and asked to point out any differences between the two pictures
M (01)	Visualizing your opponent's next chess move
M (02)	Asked to rotate and draw a complex molecule from a novel angle in front of your class
M (03)	Asked to replicate the steps someone took to complete a Rubik's cube
M (04)	Asked to imagine the 3-dimensional shape created by rotating a complex 2-dimensional plane
M (05)	Imagining whether the evening shadows cast by a large statue would cover a nearby crosswalk if the statue was rotated 90 degrees
M (06)	Asked to imagine how the orbit of a comet changes over time
M (07)	Tested on your ability to follow instructions for creating an origami design
M (08)	Imagining on a test what a 3-dimensional landscape model would look like from a different point of view
M (09)	Asked to imagine and mentally rotate a 3-dimensional figure
M (10)	Taking a test where you have to imagine what the sliced face of an apple would look like if the apple was sliced on different planes

Category	Item Text
M (11)	Asked to predict where a laser will hit after being reflected off of several mirrors
M (12)	While looking at two differently sized windmills, predict when a specific blade from each will be pointing directly up at exactly the same time
M (13)	Recalculate the trajectory of a cannon to hit a small, distant target while playing a competitive video game
M (14)	Asked to complete a slider puzzle (moving tile game) with a strict time limit
M (15)	Asked to write the alphabet with the letters upside-down as quickly and accurately as possible
M (16)	Tested on a task that requires you to judge whether two shapes, rotated at different angles, are the same
M (17)	Tested on your ability to visualize what the creases in a folded napkin would look like when unfolded
M (18)	Tested on your ability to visualize what a piece of paper would look like when folded along several drawn creases
M (19)	Predict which direction an arrow painted on a spinning wheel will be pointing when the wheel comes to rest
M (20)	Asked to imagine how the shadow of an irregularly shaped tree changes over the course of a sunny day
M (21)	Asked to predict which direction a gymnast will be facing after they complete a complex dismount involving several twists and flips
N (01)	Asked to describe where you are to someone over the phone in terms of cardinal directions (N, S, E, W)
N (02)	Trying to get somewhere you have never been to before in the middle of an unfamiliar city
N (03)	Finding your way back to your hotel after becoming lost in a new city
N (04)	Trying a new route that you think will be a shortcut without the benefit of a map
N (05)	Asked to follow directions to a location across town without the use of a map
N (06)	Finding your way to an appointment in an area of a city or town with which you are not familiar
N (07)	Tested on your ability to visualize what an object would look like from a different angle if you imagined moving yourself around the object
N (08)	Finding your way back to a familiar area after realizing you have made a wrong turn and become lost while driving
N (09)	Trying to figure out which direction you'd be facing when you came out of a subway station in an unfamiliar city
N (10)	Trying to navigate your way through a large hedge maze
N (11)	Trying to exit a large, unfamiliar building with no distinctive features or landmarks inside
N (12)	Asked to memorize routes and landmarks during a walk that you will be tested on later
N (13)	Asked to do the navigational planning for a long car trip
S (01)	Asked to imagine how gravity interacts with a passing light beam
S (02)	Shown a graph depicting an object's position over time (position graph), and then shown a number of graphs depicting an object's velocity over time (velocity graph), and asked to choose the velocity graph that best describes the object described in the position graph
S (03)	Asked to interpret a 3-dimensional graph and explain it to a classroom of your peers
S (04)	Required to interpret a diagram of a circuit board in front of a classroom
S (05)	Trying to complete a jigsaw puzzle in a short period of time to receive a prize in front of an audience

Category	Item Text
S (06)	Asked to find a location on a topographical map using information from the contour lines
S (07)	Memorizing routes and landmarks on a map for an upcoming exam
S (08)	After studying a map for a few minutes, tested on your ability to recall the routes between various items on the map
S (09)	Using a 3-dimensional model of an airport to complete an assignment
S (10)	Tested on your ability to estimate the shortest route by air between two cities on a globe
S (11)	Asked to describe the landscape on a topographical map
S (12)	Tested on your ability to find the shortest route between a series of cities on a road map
SM (01)	Asked to imagine the 3-dimensional structure of a complex molecule using only a 2-dimensional picture for reference
SM (02)	Asked to describe the interaction between two geological landforms
SM (03)	Asked to visualize what two atoms bonding would look like
SM (04)	Asked to imagine the 3-dimensional structure of the human brain from a 2-dimensional image
SM (05)	Tested on your ability to solve a complicated 3-dimensional puzzle
SM (06)	Attending an astronomy class where you will have to imagine multiple complex orbits on the final exam
SM (07)	Asked to read and explain a complex 2-dimensional graph on a test
SM (08)	Describe in accurate detail how the water moves through the pipes in your house to a plumber
SM (09)	Asked to imagine the motion of a mechanical system given a static picture of the system
SM (10)	Asked to design a room based on specific instructions about its layout
SM (11)	Asked to determine how a series of pulleys will interact given only a 2-dimensional diagram
SM (12)	Describing how the rooms of a large museum fit together to someone trying to draw an accurate representation of the floor-plan
SN (01)	Deciding whether your car can fit into an available parallel parking spot while people watch and cars are waiting to pass
SN (02)	Determining whether there is enough space for you to merge onto an adjacent lane on a busy highway
SN (03)	Giving someone directions to a hospital over the phone in reference to a map only they can see

Note. Complete list of items used for the factor analysis in Study 1. Abbreviations: I=Imagery, M=Mental-Manipulation, N=Navigation, S=Scalar-Comparison. Numbers in parentheses are arbitrary indices for referencing individual items in the factor analysis results (Study 1).