the british psychological society
promoting excellence in psychology

British Journal of Educational Psychology (2020)

www.wileyonlinelibrary.com

Spatially gifted, academically inconvenienced: Spatially talented students experience less academic engagement and more behavioural issues than other talented students

Joni M. Lakin[1]* and Jonathan Wai[2]
[1]Auburn University, Auburn, Alabama, USA
[2]University of Arkansas, Fayetteville, Arkansas, USA

Background. Spatially talented students have a capacity for success that is too often overlooked by educational services. Because these students may lack appropriate challenge, theorists suggest these students experience greater academic struggles than other gifted students, including behavioural problems and lack of academic engagement.

Aims. The goal of this research was to explore empirical evidence for the claim that spatially talented students would experience more academic struggles than other gifted students. We sought to understand the size of the 'spatially talented' population and their patterns of behavioural and academic struggles in high school. We also looked at long-term outcomes, including degree completion.

Samples. This article explores characteristics of spatial talent in three US nationally representative data sets: Project Talent (1960), High School and Beyond (1980), and the National Longitudinal Study of Youth (1997). Combined, these data provide a 60-year longitudinal study of student outcomes.

Methods. This study utilized factor analysis, analysis of variance (ANOVA), and regression methods to explore the research questions for each data set.

Results and Conclusions. From our analyses, we estimate that 4–6% (at least 2 million) of the 56.6 million students in the US K-12 system are spatially talented students that are not identified by common gifted and talented screening processes. These students had greater academic challenges, including reading difficulties, poor study habits, and behavioural troubles. We also found that spatially talented students were less likely to complete college degrees compared to other talented students. Our findings support the need for greater services to these talented students.

Introduction

It is well established that there is a population of 'spatially gifted' students, or individuals who have much greater than average 'ability to generate, retain, retrieve, and transform well-structured visual images' (p. 97, Lohman, 1996). A large body of literature has linked early spatial gifts with later achievement and contributions to society in areas that rely on

*Correspondence should be addressed to Joni Lakin, Department of Educational Foundations, Leadership, and Technology, 4036 Haley Center, Auburn University, Auburn, AL 36849, USA (email: jml0035@auburn.edu).

DOI:10.1111/bjep.12343

visual–spatial talents, such as architecture, visual arts, and many traditional science, technology, engineering, and mathematics (STEM) fields including engineering and chemistry (Kell, & Lubinski, 2013; Lubinski, 2010; Shea, Lubinski, & Benbow, 2001; Wai, Lubinski, & Benbow, 2009).

Although the societal value of spatially talented individuals is clear, traditional curricula continues to focus heavily on verbal reasoning and mathematical communication. Those students who excel based on their spatial skills do so *despite* a lack of traditional talent development opportunities found in most schools or curricula.[1] Researchers have therefore speculated that these students will experience greater academic struggles and less academic success than students with talents in verbal or quantitative domains as they progress through K-12 education (Andersen, 2014; Lohman, 1994). Wai and Worrell (2016) argued that many spatially gifted students likely 'hold out' for college to have their talents valued and developed and are therefore often bored or academically disengaged in the K-12 setting. These students may even be more likely to leave high school without a degree (Gohm, Humphreys, & Yao, 1998). Although there is theoretical consensus on the relationship of spatial talents to academic struggles, this issue has not been adequately explored empirically.

There is also the possibility that this spatial talent pool includes students who come from less educationally enriched and less socio-economically advantaged backgrounds (Lohman, 1994; Wai & Worrell, 2016). Because socio-economic status is less correlated with spatial reasoning in the population than is math and verbal reasoning, this means that deliberately selecting for spatially talented students using reliable and valid measures of spatial reasoning would improve the background and demographic diversity of those students who are considered 'talented' (Wai & Allen, 2019; Webb, Lubinski, & Benbow, 2007).

Spatial ability

There are several types of spatial reasoning abilities, and many assessments that measure these different aspects of spatial ability. Through a literature review, we found two types of spatial thinking that were repeatedly confirmed through psychometric research to reflect distinct skills (Hegarty & Waller, 2005; Mix *et al*., 2017). The first important set of skills is Visualization, which is the ability to encode spatial information and reason about those representations (often under speeded conditions). A classic task in this area is the paper folding test where examinees imagine a piece of paper that is folded and then cut and have to determine what the unfolded paper will look like. The other key set of skills is Rotation, which is the ability to mentally rotate 2- and 3-dimensional objects in space. A commonly used assessment of this skill is the Purdue Spatial Visualization Test-Rotations (revised by Yoon, 2011) where examinees determine how one abstract figure has been rotated and then imagine how another abstract figure could be rotated in the same way. Other researchers have defined additional factors such as orientation and mechanical reasoning (Hegarty & Waller, 2005; Linn & Petersen, 1985; Lohman, 1996; Newcombe & Shipley, 2015). Others have organized the specific spatial ability components differently, such as the extrinsic/static continua defined by Uttal *et al*. (2013). However, these additional dimensions often blend visualization and rotation skills (or achievement in the

[1] *A few spatially oriented competitions do exist, including robotics programmes gaining popularity across the globe (Coxon, 2012) and some aspects of science and math competitions (Bicknell, 2008).*

case of mechanical reasoning) and do not seem independently important to academic outcomes related to STEM (Harris, Hirsh-Pasek, & Newcombe, 2013).

Previous work has considered the relationship between spatial reasoning abilities and general intelligence (*g*). Factor analytic approaches suggest that complex spatial tasks may be little more than measures of *g* (Lohman, 1996). However, there is a specific factor, measured by both complex and simple spatial tasks that form reliable specific variance (Carroll, 1993).

The argument for a spatial dimension of reasoning and problem-solving is further bolstered by the experiences of individuals. Lohman (1994) makes the case that personal experience plays a large role in whether one believes that some ability or capacity is worthy of investigation or is of importance, noting how even scientists pursue their questions through ways, such as personal ability strengths, in which they are most familiar. In the same way, for those of us who do not share spatial strengths, it may be hard to understand how there might be a population of students who do have that strength. Shepard (1990) describes how one who sees the world through spatial imagery might describe it:

> My efforts toward the faithful externalization of particular, spontaneous visual images began in earnest following my involuntary experience of an extraordinarily vivid and geometrically regular visual image just before awakening one morning in 1970. With eyes still closed on that morning, I suddenly saw before me an immense, luminously shimmering, golden array of diamond-shaped panels separated by burnished beveled strips. . . . my memory of the image remained so vivid and my feeling of awe at its vast scale, its pristine regularity, and its preternaturally luminous and shimmering quality remained so keen that I immediately set about making a pencil sketch of it . . . (p. 35)

The factor analytic and applied research (Carroll, 1993; Lohman, 1996) clearly suggests that some students with spatial talents will also have verbal and mathematical talents. However, there will also be spatially talented students who have significantly lower verbal and mathematical abilities. This is important because the former group would still be identified in typical talent searches or K-12 and college admissions testing. These assessments reflect the types of abilities that school systems typically value and the skills known to relate to positive work and life outcomes (Kuncel, Hezlett, & Ones, 2004). However, students with *only* spatial talents strong enough to be identified are likely missed in screening procedures throughout school and in those national searches. This population of spatially talented students with relatively lower math and verbal talents composes a potentially neglected and vulnerable pool of students who may have great potential but may not have the opportunity to develop it fully because they are never identified.

Lohman's (1994) central argument in his article aptly titled 'Spatially Gifted, Verbally Inconvenienced' is that the K-12 curriculum is heavily reliant on verbal communication and, to a lesser degree, mathematical representations. Lohman (1994, p. 262) wrote that 'I do believe that achievement tests – with their heavy emphasis on specific language abilities and mathematical calculation abilities – do miss children who can be among our most creative thinkers'. In particular, many high spatial students may be less verbally fluent, the ability to easily retrieve words on the basis of sound patterns, to be able to speak fluently on cue. In other words, serving spatially talented students will require considerable curricular innovation, which might be one factor in the lack of services for this population.

Previous work indicating spatial talent could be associated with academic struggles

A limited area of research has explored the association that spatial ability may have with underachievement, personality traits, personal preferences (e.g., Barton, Cattell, & Silverman, 1974; Webb *et al.*, 2007), and other traits relevant to educational outcomes. The most prominent work in this area has been primarily anecdotal and theoretical (e.g., Andersen, 2014; Lohman, 1994; West, 1991). Lohman (1996) suggested several reasons why empirical evidence was lacking at the time he was writing. One was that the academic outcomes available to many researchers at the time were too aligned to verbal outcomes of schooling. Thus, work from longitudinal databases that offer a broader range of outcome measures has been vital to providing empirical support.

One of the few empirical studies of the issue was a study by Gohm *et al*. (1998; see also Humphreys, Lubinski, & Yao, 1993) where they used the 1960 Project Talent data to identify spatial and mathematically talented individuals in the top 1% of ability in one of the two domains (excluding those with high scores in both). They found differences in interests, where spatially talented students expressed more interest in hands-on activities like art, model building, and electrical working. Relevant to this study, they also analysed item-level responses to questions about study habits and academic motivation. They found spatially talented students reported weaker study habits, more reasons for not going to college, and lower grades in a variety of academic domains. Previous work did not examine latent constructs representing academic challenges and focused on a narrow definition of high ability. Thus, this work limited their focus to students with just one area of exceptional strength. Previous work also did not consider behavioural problems or degree completion. However, this work did provide the important foundation and evidence for the idea that students with spatial strengths might be bored in school and/or not being served sufficiently and having challenges to develop their talents. And, given that the focus of selection and curricula is on math and verbal types of symbol systems, spatial skills have also not been traditionally valued.

Current study

The goal of this study was to broadly examine the empirical evidence and hypotheses that spatially gifted students will experience less engagement and academic success than students who are gifted in mathematics or verbal domains. Like Gohm *et al*. (1998), we used Project Talent data, but we analysed the data using latent variables and considered long-term degree outcomes. To assess whether similar findings were found in other, independent samples, we also analysed the data from High School and Beyond, first surveyed in 1980, as well as the National Longitudinal Study of Youth, first surveyed in 1997. Using three data sets allows us to utilize Lykken's (1968) approach of *constructive replication* as preserving focal constructs in each database but varying construct-irrelevant design features. We reason that though each individual longitudinal data set has its limitations and is set in a specific period of time, findings that replicate across three separate databases that span the last 60 years may be considered more robust. See Figure 1. Our guiding research questions were as follows:

1. How much does including spatial talents expand the pool of identified, talented students in comparison to the pool identified solely with verbal and quantitative measures?

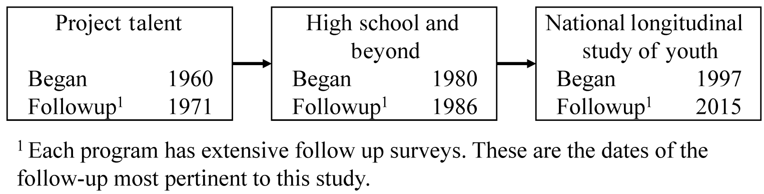

Figure 1. Longitudinal data study design comprising three cohort replications.

2. Do spatially talented students, who do not have verbal or quantitative strengths, have higher reported difficulty with school compared to non-identified or verbal/quantitative identified students?
3. Are spatially talented students, who do not have verbal or quantitative strengths, have greater reported behavioural issues compared to non-identified or verbal/quantitative identified students?
4. Are spatially talented students less likely than other talented students to complete high school or college?

Methods

We identified three suitable data sets collected by or available through the National Center for Educational Statistics (NCES) in the United States by looking for data sets that had all of these key measures: (1) measures of achievement or ability in verbal and quantitative domains, (2) at least one measure of visualization or rotation skills, (3) measures of concurrent academic issues such as disengagement or behavioural problems, and (4) measures of degree completion (available through follow-up surveys). Table 1 shows the representation of constructs in each survey programme. Table 2 provides the demographic information and sample size for each programme.<!–[BJEP_12343_

Project Talent, 1960 (PT60)

Project Talent is a longitudinal study in the United States that began in 1960 with four independent cohorts of 9th- through 12th-grade students across the United States with follow-up surveys over the ensuing years and continuing today (Flanagan *et al.*, 1962). This remarkably comprehensive study claims to have assessed 5% of the US high school population at the time (around 440,000 students from over 1,300 schools; American Institutes for Research, n.d.). Our focus is on the academic variables measured in the 1960 baseline data. PT60 has, by far, the most complete ability assessment of the three different data sets for this study, and included comprehensive measures of reading and math achievement as well as a reading comprehension test. For spatial reasoning, it includes both a visualization in 2D (24 items, figure rotation), a 3D visualization test (16 items, box folding), and a mechanical reasoning task (20 items). They also included a measure of abstract reasoning with 15 figural matrices items. See Figure 2. The student surveys also asked questions about struggles with reading (e.g., 'I read and reread without comprehension') and overall difficulty in school (e.g., 'lack of interest increases distraction'). In follow-up surveys, they also gathered information about later degree completion.

Table 1. Variables used from each database

Variables	PT 1960	HSB 1980	NLSY 1997
Race	Self-report	Self-report	Self-report
Ethnicity	*Not collected*	Self-report	Self-report
Sex	Self-report	Self-report	Self-report
SES	Socio-economic Environment Index	Self-report family income	*Not collected*
Rurality	Location based	*Not collected*	Location based
Tests, subtests (# items) [reliability estimate]			
English Achievement	Spelling (16), Capitalization (33), Punctuation (27), English Usage (25), Expression (12) [r^a = .87]	(20) [r^b = .80]	*Not collected*
Reading Comprehension	(48) [r^c = .84]	*Not collected*	ASVAB Word Knowledge (35), Paragraph Comp. (15)
Mathematics Achievement	Basic through high school math (54) [r^a = .83]	(25) [r^b = .84]	ASVAB Arithmetic Reasoning (30) and Math Knowledge (25)
Abstract Reasoning	Figure matrices (15) [r^a = .65]	*Not collected*	
2D Spatial Reasoning	2D rotation (24) [r^c = .80]	*Not collected*	ASVAB Assembling Objects (25)
3D Spatial Reasoning	3D visualization, box folding (16) [r^c = .66]	(16) [r^b = .70]	
Mechanical Reasoning	(20) [r^a = .69]	*Not collected*	ASVAB Mechanical Comprehension (25)
Personality measures	Tidiness (11) [r^d = .88], Maturity scales (24) [r^d = .93]	*Not collected*	Organized (1), Conscientious (1), Dependable (1), Thorough (1)
Academic challenges[f]	Engagement (21), Reading problems (4)[f] See Table 2	Engagement (3), Interference (4)[f] See Table 3	Breaks School Rules (1), Thorough or Careless (1), Hard Worker (1), Disorganized or Organized (1), Amount of Work (1)
Youth report of problems[f]	*Not collected*	Discipline problems (4)[f] See Table 3	Trouble Paying Attention-Boys (1) Behavioral/ Emotional Scale (4) [Boys r^e = .51, Girls r^e = .53]

Continued

Table 1. (*Continued*)

Variables	PT 1960	HSB 1980	NLSY 1997
Parental report of problems	*Not collected*	*Not collected*	Poor school work-girls (1) Behavioural/Emotional Scale (4) [Boys r^e = .65, Girls r^e = .57]
Long-term degree	11 years post- graduation self-report	6-year follow-up self-report	18-year follow-up self-report
Long-term income	11 years post-graduation self-report	6-year follow-up self-report	18-year follow-up self-report

Note. [a]KR-20/21 estimate from Flanagan and American Institutes for Research in the Behavioral Sciences (1972), averaged across 9th- to 12th-grade samples. Flanagan and American Institutes for Research in the Behavioral Sciences (1972) reported KR-21 and split-half reliability methods for different test scales based on data availability; [b]KR-20 from Talboy (2011); [c]Split-half estimate from Flanagan and American Institutes for Research in the Behavioral Sciences (1972), averaged across 9th- to 12th-grade samples; [d]Cronbach's alpha estimate from Pozzebon *et al*. (2013); [e]Cronbach's alpha estimate from Child Trends, Inc., the Center for Human Resource Research (1999).; [f]For PT60 and HSB80, these items were subjected to factor analysis; therefore, reliability of original scales is not relevant.

Table 2. Demographic characteristics of the samples

	PT 1960, *N* = 103,893		HSB 1980, *N* = 28,240		NLSY 1997, *N* = 8,984	
	Frequency	Per cent	Frequency	Per cent	Frequency	Per cent
Low SES[a]	33,727	25%	4,653	16%	–	
Middle	32,253	46%	8,789	31%	–	
High	31,216	15%	10,630	38%	–	
Missing	–		4,168	15%		
Rural	12,201	17%	–		1,484	17%
Non-rural	71,863	67%	–		5,227	59%
Urban (over 1.5 M)	3,779	15%	–		–	
Missing	16,050	15%	–		2,223	25%
Male	52,566	51%	12,907	46%	4,598	51%
Female	51,327	49%	14,086	50%	4,386	49%
Missing	–		1,247	4%	–	
White	34,489	33%	21,328	76%	4,665	52%
African American	1,883	2%	3,965	14%	2,335	26%
Asian	276	<1%	377	1%	160	2%
Hispanic	–		3,177	11%	1,901	21%
Native American	–		255	1%	–	
Other	227	<1%	1,793	6%	83	1%
Missing	67,018	65%	522	2%	–	

Note. [a]Socio-economic status composite variable from each study reclassified into three levels.

High School and Beyond, 1980 (HSB80)

High School and Beyond surveyed 30,000 high school seniors and 28,000 sophomores in 1980 (Sebring *et al.*, 1987). They used a two-stage stratified random sample of schools and students (1,100 schools sampled in stage 1). The academic assessments included reading and math achievement tests, and they administered a 16-item 3D visualization task (identical to the assessment included in PT60) to the seniors in the sample.

Questions related to academic (dis)engagement included that courses were too hard, they had poor study habits, or they were unsatisfied with their education. Variables related to behavioural issues included questions about being interested in school, being suspended or put on probation, and cutting classes.

National Longitudinal Study of Youth 1997 (NLSY97)

The NLSY97 cohort included 8,984 youth ages 12–18 (US Bureau of Labor & Statistics, 2006). The sample was designed to be United States nationally representative with an over-sampling of Latino and African American respondents to ensure representation. Students completed the Armed Services Vocational Aptitude Battery (ASVAB), which provided verbal measures (Word Knowledge and Paragraph Comprehension) and mathematics knowledge measures (Arithmetic Reasoning and Math Knowledge). ASVAB includes a measure of spatial reasoning called 'Object Assembly' (a 2D visualization task, see Figure 3) and a mechanical comprehension test (similar to the test administered for PT60).

What NLSY97 provides that is unique from the other data sets is parental reports of behavioural, learning, and emotional problems. This provides a number of variables

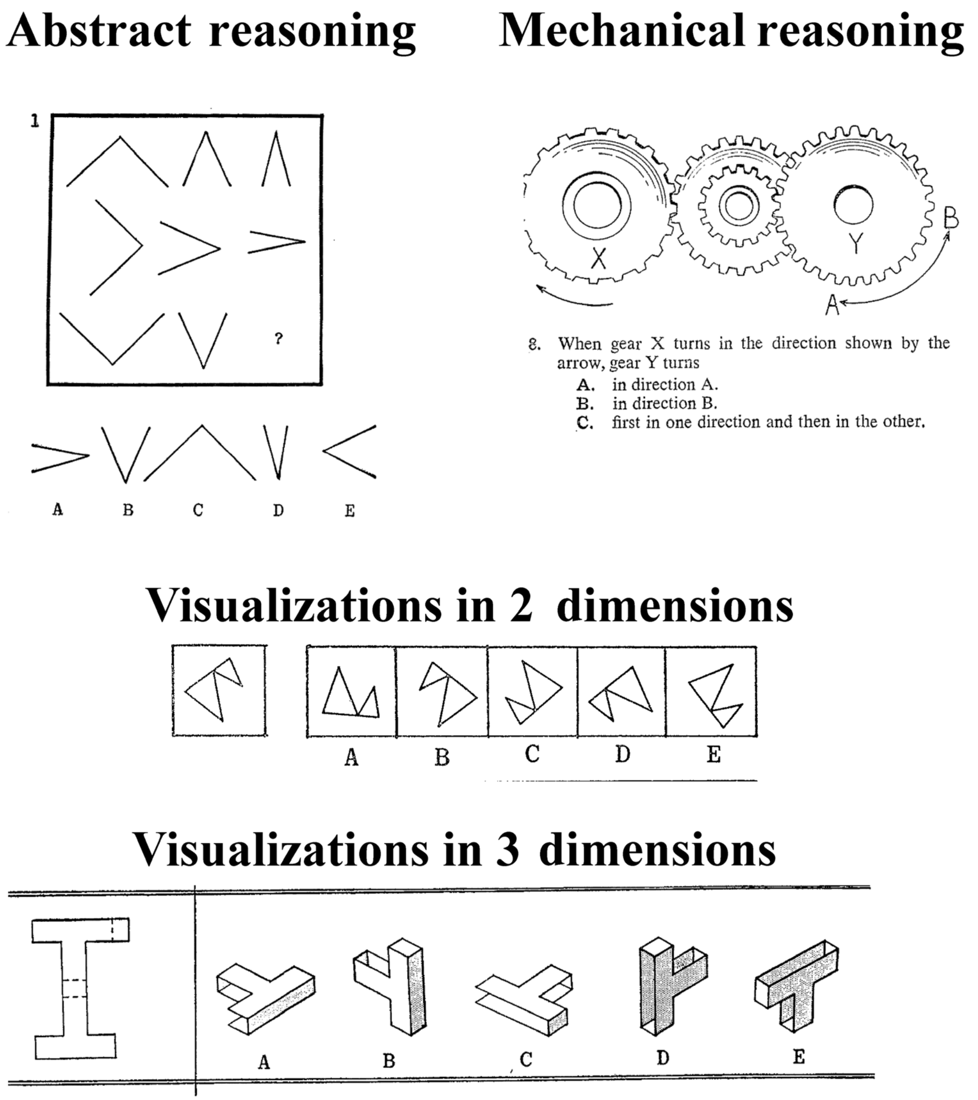

Figure 2. PT60 measures of spatial reasoning (Flanagan & American Institutes for Research in the Behavioral Sciences, 1972).

related to struggling with the teen years and schooling that may correlate to having talents in spatial rather than verbal or quantitative domains.

Identifying factors and calculating battery scores

Project Talent results

Of all the data sets, PT60 (Flanagan *et al.*, 1962) had the most measures of achievement and reasoning related to our three domains of interest (verbal, quantitative/mathematics, and spatial). We were able to identify more reasoning than achievement measures for verbal and spatial reasoning, but no strong reasoning measures were administered in the quantitative domain, so achievement measures were used.

Table 3. PT60 factor structure (pattern loadings)

	Factors		
	Academic disengagement	Reading difficulty	Academic engagement
Rotated factor per cent of variance	19.4%	4.8%	4.6%
Difficulty expressing self in reports exams assignments	0.215	0.264	−0.041
Speed reading helps complete lessons quickly	−0.012	0.277	−0.194
Grades accurately reflect ability	−0.087	0.000	**0.402**
Understand what to do before starting assignments	−0.156	0.021	**0.460**
Little accomplishment for time spent studying	**0.300**	0.237	0.032
Lack of interest increases distraction	**0.526**	0.125	−0.045
Enjoy writing reports and compositions	0.082	−0.165	**0.341**
Inattention in class causes lower marks	**0.620**	−0.022	0.025
Difficult assignments are challenges to ability	0.130	−0.108	**0.458**
Too quick to do best work	**0.556**	−0.026	0.092
Inattention caused missed assignments and announcements	**0.697**	−0.074	−0.002
Teachers critical of sloppy assignments	**0.562**	−0.085	−0.030
Do just enough to get by unless really like course	**0.549**	0.038	−0.050
Difficulty with mechanics of English composition	**0.357**	0.176	−0.029
Attention strays in class	**0.576**	0.125	−0.044
Get behind in assignments	**0.613**	−0.024	−0.136
Careless errors lower grades	**0.423**	0.108	−0.017
Slow reading holds me back	0.229	**0.318**	−0.065
Pronounce words to self while reading them	0.053	0.109	0.261
Courses not much help to occupation after school	0.239	0.124	0.082
Can pick out important points when studying for test	−0.086	−0.074	**0.510**
Easily distracted when reading	0.252	**0.460**	0.033
Keep up to date and do work every day	**−0.324**	0.228	**0.504**
Trouble remembering reading	0.034	**0.744**	0.030
Read and reread without comprehension	0.109	**0.681**	0.000

Note. Loadings at or above .30 in bold.

To create composite scores, we ran exploratory factor analyses (EFAs) using maximum likelihood estimation and retained factor scores based on the first factor (which in each case explained the vast majority of variance). Verbal reasoning was based on reading comprehension, creativity (a verbal fluency measure), and word functions (the first factor explained 71% of variance). Mathematics reasoning was based on the three math tests (Math I + II + III) plus computation (the first factor explained 76% of variance). The spatial reasoning factor was based on 2D visualization, 3D visualization, mechanical reasoning, and abstract reasoning (consistent with prior work with this data set, Wai *et al.*, 2009; the first factor explained 62% of variance).

To condense the measures of academic disengagement, a factor analysis with maximum likelihood estimation and Oblimin orthogonal rotation was run on a selection of 25 items that were most relevant to our interests. Five factors had eigenvalues above 1, but there was a clear break in the scree plot at three factors and the two additional factors were not theoretically meaningful. See Figure 4a. Therefore, three factors were retained, accounting for 36% of the variance. Factor 1 was defined by general academic disengagement (e.g., ‘lack of interest increases distraction’). Factor 2 was defined by specific difficulty with reading (e.g., ‘read and reread without comprehension’). Factor 3

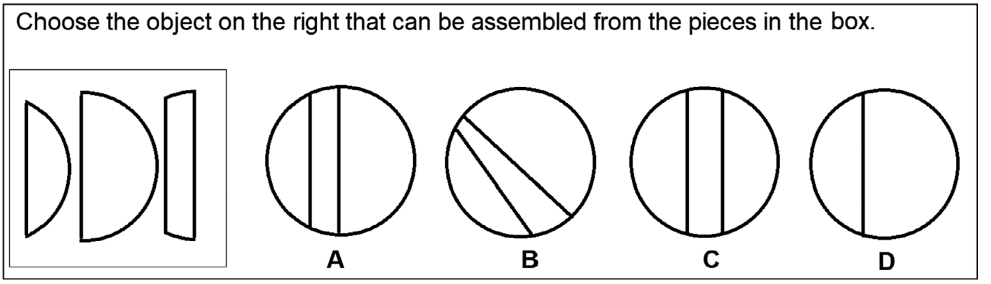

Figure 3. Example of an Assembling Objects item.

was characterized by more academic engagement statements (e.g., 'Difficult assignments are challenges to ability').

Weighted composites were estimated for each factor using item loadings of ≥0.30 to weight each variable included in the composite. Costello and Osborne (2011) recommend 0.30 as a cut-off for acceptable item loading. Items shown in Table 3 with no loadings at or above 0.30 were omitted from the composite factor score calculations.

PT60 also included a number of personality scales (Pozzebon, Damian, Hill, Lin, Lapham, & Roberts, 2013). Previous work (Major, Johnson, & Deary, 2014) found that the Tidiness and Mature Personality scales align with Conscientiousness. Because this personality trait is associated with academic success (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), we included it as another possible covariate with spatial giftedness.

HSB80 results

Unlike PT60, HSB80 had only achievement-type measures of verbal and quantitative domains. Therefore, the standardized scale scores for Reading total and Mathematics total were used. For spatial, only the 3D visualization task from PT60 was administered. Therefore, the standardized score from the visualization task was used as our indicator of spatial reasoning.

Not all of the behavioural variables were administered to students who also took the visual task (some tests seem to have been administered to different halves of the sample). From the measures available, we selected those that were (1) collected alongside the visual–spatial ability measures and (2) were most pertinent to academic engagement and behavioural challenges. Three factors had eigenvalues above 1 and were interpretable. See Figure 4b. They were extracted using maximum likelihood and Oblimin orthogonal rotation. These three factors explained about 30% of the variance. The first factor can be described as general attitude towards school and working hard in school. The second factor could be described as items reflecting serious behavioural issues (mainly defined by disciplinary problems, suspensions, and trouble with the law). The third factor mainly reflected perceptions of interference in success (e.g., 'Poor study habits' interfere) that loaded on this factor, in contrast to the item stating they are satisfied with their schooling. Weighted composites were estimated for each factor using item loadings of ≥.30 to weight each variable included in the composite factor scores (see Table 4).

NLSY97 results

We used Armed Services Vocational Aptitude Battery (ASVAB) scale scores to estimate the three domains for NLSY97. Verbal Expression is a commonly used composite in other

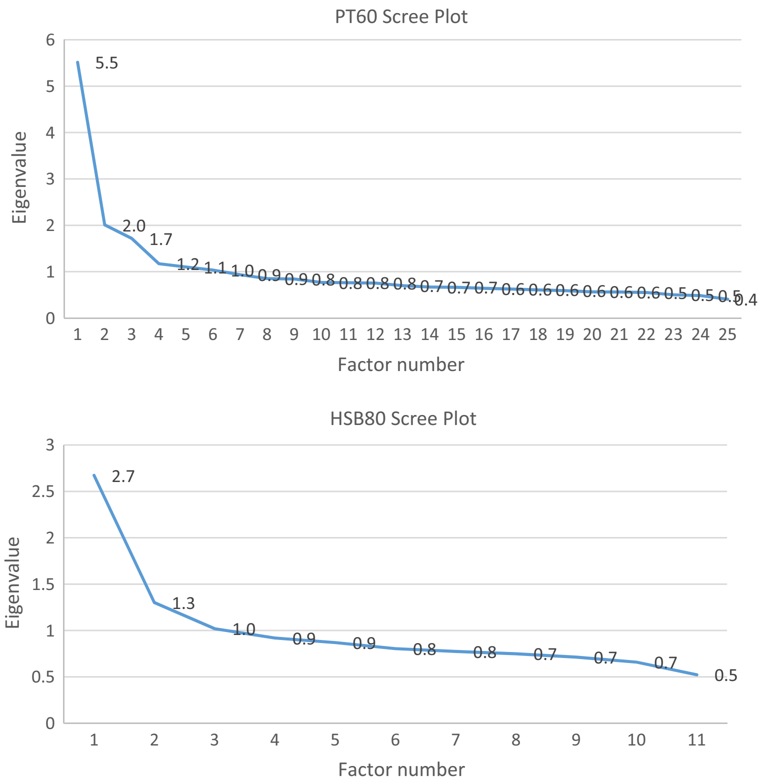

Figure 4. Scree plots for exploratory factor analysis of (a) PT60 and (b) HSB80.

research, so we used that as our measure of verbal reasoning (Barto *et al.*, 2010). This composite equally weights the Word Knowledge and Paragraph Comprehension scales. Likewise, the Quantitative Composite equally weights Arithmetic Reasoning and Math Knowledge.

ASVAB provides both a 2-D visualization task known as Assembling Objects (AO) and a Mechanical Comprehension (MC) task (similar to other Mechanical Reasoning [MR] type tests). We felt that including mechanical reasoning in the spatial score would weight the composite too heavily against females (Feingold, 1992) and contaminate the measure of spatial reasoning when it is not offset by several stronger measures of spatial reasoning (Lohman, 1979). In contrast, MR was just one of four measures in the PT60 battery and is therefore less weighted in that composite. In sum, we decided to use AO scale scores as our indicator of spatial ability and did not use MC scores in these analyses.

Unlike the other data sets, NLSY97 provided mainly scale composites of relevant personality and academic engagement variables. Therefore, factor analysis was not appropriate. Also, several variables were only available for male or female participants, so overall factor solutions were not possible. We selected the 11 scales listed in Table 7 as providing scores that were most relevant to our interests.

Talent pools

For each data set, students were assigned to three categories in the 'talent pool'. Students with verbal or quantitative scores at or above the 95th percentile (compared to other

Table 4. HSB80 factor structure (pattern loadings)

	Factors		
	School attitude	Serious behaviour issues	Interfering with success
Rotated factor per cent of variance	13.7%	11.2%	4.9%
Satisfied with education	0.026	−0.024	**−0.449**
Interested in school	**0.303**	−0.113	−0.236
Like working hard in school	**1.037**	0.044	0.059
Interference: courses too hard	0.015	−0.006	**0.355**
Interference: hard to adjust to school	0.012	0.148	**0.453**
Interference: poor teaching	0.043	−0.049	**0.463**
Interference: poor study habits	−0.143	−0.029	**0.481**
Disciplinary problems in school	0.024	**0.565**	0.011
Suspended or probation in school	0.001	**0.616**	−0.083
Cut classes now and then	−0.123	**0.233**	0.129
Serious trouble with the law	0.007	**0.313**	0.025

Note. Loadings at or above .30 in bold.

students in the data set) were classified as ‘Verbal or Quantitative (V/Q) talented’, regardless of their spatial scores. Students not included in that group were identified as ‘Spatial talented only’ if their scores on the spatial domain were at or above the 95th percentile. All other students were classified as ‘non-identified’. This system of identification allowed us to compare the group with only spatial talents to students who have that strength but would likely already be identified by traditional (i.e., verbal or quantitative) measures. Talent searches and academic selection broadly rely on standardized tests that use math and verbal reasoning measures so this helps examine the talent pool missed by most academic selection procedures as they exist in US education.

Analyses

To address our research questions, we used the factor scores and composite variables described above. Research question 1 was addressed with descriptive statistics. Research question 2 was addressed by a series of one-way independent ANOVAs comparing the behavioural and engagement factors available for each data set across groups defined by the ‘talent pool’ variable. Finally, the degree attainment data was tested with χ^2 comparison of proportions.

Results

Our first research question was, how many students have a relative strength in spatial ability? Each of the data sets provides a tentative estimate. The PT60 data estimate the proportion of students with scores in the top 95% on verbal or quantitative (or both) scales was 8%, whereas those in the top 95% of spatial (and *not* in the top of verbal or quantitative) was 4%. Extrapolated to today’s population of roughly 56.6 million K-12 students (National Center for Educational Statistics, 2018), this suggests that there are about 2.21 million students with spatial strengths that would not be identified by talent searches that look at only verbal and quantitative talents (or general ability underlying those measures).

HSB80 identified 15% of the sample with scores in the top 95% on verbal or quantitative (or both) scales.[2] Again, 4% of students had a score in the top 95% of spatial alone. HSB80 suggests about 2.27 million students with spatial strengths overlooked.

NLSY97 identified 9% of the sample with high verbal or quantitative (or both) scores. In NLSY97, 6% of students had a spatial strength but not math or verbal. This suggests 3.16 million students could be overlooked. Across the three estimates, at least 2 million students are currently overlooked for their primarily spatial talents. Many more would have been overlooked in the past.[3]

Association with academic struggles

PT60

We compared the three talent pool levels (not identified, spatial-only, V/Q identified) in terms of the scores on these academic struggles factors created through EFA analyses. Significant group differences were found on all three factors with small to medium effect sizes (Cohen, 1988). See Table 5 and Figure 5.

Spatial-only identified students had the highest academic disengagement. This differed significantly from non-identified students (Cohen's $d = 0.05$) and V/Q-identified students ($d = 0.49$).

Scores on the reading challenges factor ('reading holds me back') showed significant differences. Although spatially identified students had lower scores than non-identified students ($d = -0.10$), they reported more struggles with reading than other identified students ($d = 0.29$). Spatial-identified students reported the lowest engagement of the three groups, with averages similar to non-identified students ($d = -0.06$), but substantially lower than V/Q-identified students ($d = -0.46$). On the personality scales, spatial-identified students showed similar Tidiness + Mature Personality scores compared to non-identified ($d = 0.03$), but lower scores than V/Q-identified students ($d = -0.28$). Overall, our hypotheses were confirmed in PT60. Students with spatial talents tended to struggle in school much more than V/Q-identified students and their struggles were similar in magnitude to non-identified students.

HSB80

Our analysis using HSB80 data also showed that spatially talented students had significantly lower academic engagement than V/Q-talented students ($d = -0.17$, marginal), significantly more serious trouble ($d = 0.16$, marginal), and a greater sense of study habits and course difficulty interfering with success ($d = 0.23$, small). See Table 6 and Figure 6. These scores were similar to levels reported by the non-identified students.

NLSY97

We used Holm's sequential Bonferroni adjustments (Abdi, 2010) to maintain type I error rates due to the large number of outcomes in NLSY97 (see Table 7 and Figure 7).

---

[2] *Some estimates of percentile ranking were inflated due to tied scores and limited score ranges on the underlying measures.*

[3] *The typical pool that participates in the 7th-grade talent search across the United States is roughly the top 5% of ability (Wai, Pullataz, & Makel, 2012), so these findings directly pertain to the pool of students who are able to participate in the talent search itself. This is important, given that students who fail to be a part of the initial selection procedure are unlikely to have their talents developed further in this academic development pipeline.*

Table 5. PT60 academic struggle ANOVA results

	F	*df*	Significance	η^2 effect size	Cohen's *d* Spatial versus V/Q	Cohen's *d* Spatial versus not identified
Academic disengagement	1034.87	2, 96,660	$p < .0001$	.021	0.49	0.05
Reading holds student back	909.37	2, 96,668	$p < .0001$	.018	0.29	−0.11
Academic engagement	906.57	2, 96,669	$p < .0001$	.018	−0.46	−0.06
Tidiness + Mature personality	838.33	2, 101,180	$p < .0001$	.016	−0.28	0.03

Note. $\alpha = .05$.

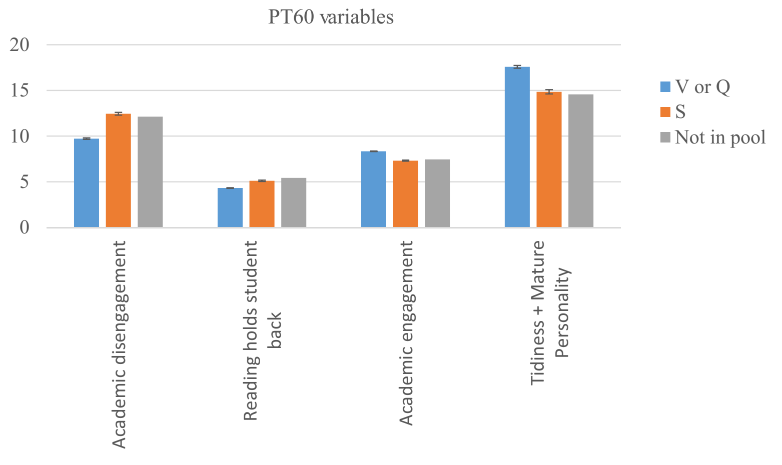

Figure 5. Mean differences in PT60 composites.

Significant differences were found for the personality scales related to 'amount of work' and 'breaks school rules'. A scale related to girls' 'School work is poor', boys' 'trouble paying attention', and boys' 'Behavioural/Emotional Scale' as rated by the parent was also significant.

The key differences between the V/Q-identified and the spatial-only students were in girls' poor school work, breaking school rules, where spatial students had higher scores, boys behavioural/emotional scores, and trouble paying attention. Of the significant effects, only 'amount of work' showed similar levels between the V/Q and spatial-identified students.

Association with academic and career outcomes

PT60

PT60 collected information about students' degree completion 11 years after high school graduation. No patterns were found across groups for high school or college degree completion (the chi-square test was not significant). See Table 8.

Table 6. HSB80 academic struggle ANOVA results

					Cohen's *d*	
	F	*df*	Significance	η^2 effect size	Spatial versus V/Q	Spatial versus not identified
Academic engagement	50.630	2, 22,685	$p < .001$	.004	−0.17	NS
Serious trouble	86.998	2, 22,961	$p < .001$	.008	0.16	NS
Factors interfering with success	108.330	2, 22,784	$p < .001$	.009	0.23	NS

Note. $\alpha = .05$; NS = non-significant.

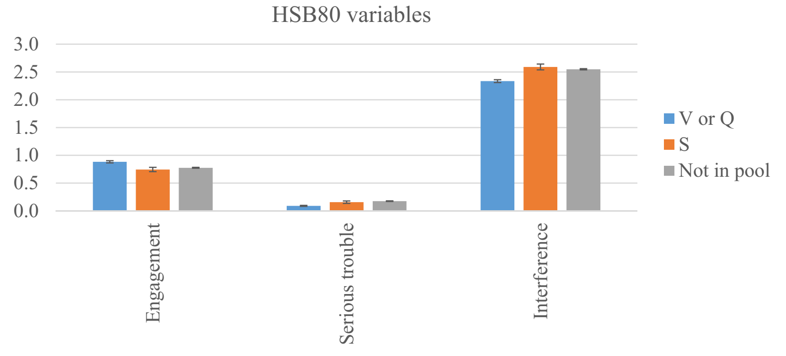

Figure 6. Mean differences in HSB80 composites.

HSB80

Seniors from the original HSB80 sample were surveyed again in 1986.[4] There was an association of talent identification and degree completion, where spatial-only students were much more likely to have no degree or only a high school degree compared to other identified students, $\chi^2(2) = 29.6$, $p < .0001$. See Table 8.

NLSY97

NLSY97 provided the same degree information around 7 and 18 years after initial data collection. Chi-square tests showed that the differences in proportions were significant ($p < .01$). See Table 8. In both 2004 and 2015, we found substantial differences in degree completion between spatial-only and VQ-identified students, with spatial-identified students completing high school and college at lower rates than V/Q-identified students.

Discussion

In this paper, we sought to investigate the importance of identifying spatially talented students. We first documented the probable size of this population and found it substantial. Next, we showed that students with these specific talents were more likely to

[4] *There were additional follow-up surveys, but the 1986 follow-up had a markedly larger response rate.*

Table 7. NLSY97 academic struggle ANOVA results

	Year, source	*F*	*df*	Adjusted α^{a}	Significance	η^2 effect size	Cohen's d^{b} Spatial versus V/Q	Cohen's d^{b} Spatial versus not identified
Amount of Work	2008, self-report	24.659	2, 6,016	.05	$\boldsymbol{p < .001}$	.008	NS	−0.16
Girls' Poor School Work	1997, parent-report	18.382	2, 2,057	.025	$\boldsymbol{p < .001}$	.017	0.27	−0.23
Breaks School Rules	2008, self-report	11.745	2, 6,017	.0125	$\boldsymbol{p < .001}$	.004	0.18	NS
Boys' Behavioural/Emotional Scale	1997, parent-report	10.46	2, 1,393	.00625	$\boldsymbol{p < .001}$	.014	0.27	NS
Boys, Trouble Paying Attention	1997, self-report	5.958	2, 2,206	.00313	$\boldsymbol{p = .003}$	.005	0.12	−0.09
Girls' Behavioural/Emotional Scale	1997, parent-report	5.506	2, 1,272	.00156	$p = .004$	.008		
Boys' Behavioural/Emotional Scale	1997, self-report	4.054	2, 2,206	.00078	$p = .017$	.004		
Girls' Behavioural/Emotional Scale	1997, self-report	3.261	2, 2,057	.00039	$p = .039$	.003		
Thorough or Careless	2002, self-report	2.09	2, 3,926	.0002	$p = .124$	.001		
Hard Worker	2008, self-report	1.888	2, 6,021	9.77E-05	$p = .151$	.001		
Disorganized or Organized	2002, self-report	0.902	2, 3,924	4.88E-05	$p = .406$	<.001		

Note. NS = non-significant.
[a]Holm's sequential Bonferroni adjustment (Abdi, 2010) creates increasingly stringent α for each additional test conducted. This column provides the adjusted α for each test;; [b]Positive Cohen's *d* means spatial talent group has higher mean. *p* values in bold are significant.

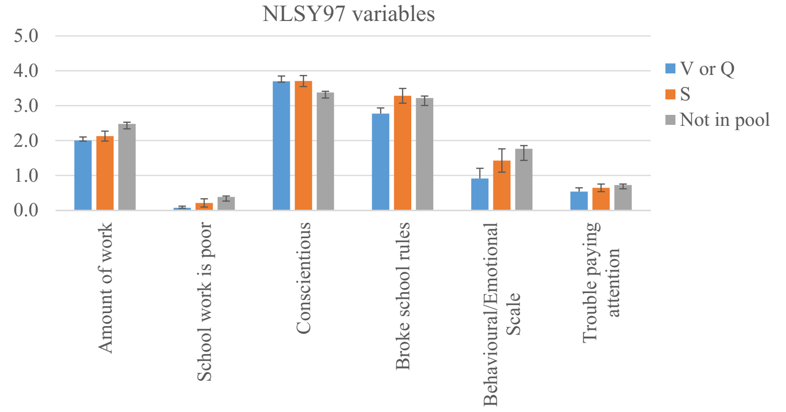

Figure 7. Mean differences in NLSY97 composites.

struggle academically than other gifted students, specifically in terms of lower academic engagement, behavioural difficulties, and lower degree completion rates. This combined evidence demonstrates the importance of finding and serving spatially talented students.

Limitations and future directions

Each of the databases (PT60, HSB80, and NLSY97) varied in how and what they measured for each of the key constructs examined in this study. Consistent with the concept of constructive replication (Lykken, 1968), we sought to maximize the comparability of our analyses across data sets despite their inherent differences in both measures collected and cohort. Even so, our results are impacted to an unknown degree by cohort and measurement differences. We hope that future longitudinal survey programmes will include measures of spatial reasoning and a wide array of behavioural and academic outcomes to see whether our findings further replicate.

Each of the data sets has its own limitations. First, PT60 is older and was collected around the time of US racial integration of public schools when demographic and school quality differed substantially from the present time. The strength of the PT60 data is that it continues to be among the most comprehensive databases including a wide array of baseline spatial measures and later outcomes. HSB80 is also dated and has a less comprehensive battery than PT60. However, the consistency in measures across those programmes is a benefit. Finally, NLSY97 has by far the most comprehensive battery measuring behavioural and academic outcomes. Thus, this newer data set, while utilizing just one spatial measure, provides powerful evidence of predictive validity.

Another limitation of our study is that the longitudinal databases we utilized are powerful for detecting correlational patterns in independent data sets across time, but are not able to test causal explanations for our findings. Ultimately, although we showed associations between spatial talents and academic struggles, more research is needed into *why* students with spatial strengths tend to have academic struggles as indicated by school difficulty, behavioural issues, and lower academic credential completion. Lohman (1994) offered several potential explanations for why spatial talents seem associated with verbal weaknesses and academic struggles. One possibility is that spatial talents are connected to specific learning disabilities surrounding language and communication. Another possibility is that school systems, in the United States in particular, simply are not set up to accommodate and educate students with spatial strengths when their verbal ability is

Table 8. Advanced degree outcomes (percentage)

	PT60			
	Verbal or quantitative	Spatial only	Not identified	χ^2 (V/Q compared to Spatial)
No degree	3.3%	2.6%	2.8%	$p > .05$
HS diploma	66.1%	67.9%	66.1%	
College degree or beyond	30.6%	29.5%	30.6%	
HSB80				
No degree	7.6%	16.7%	13.6%	$p < .001$
HS diploma	15.8%	23.5%	36.2%	
College degree or beyond	76.6%	59.7%	50.2%	
NLSY97 2004 Follow-up				
No degree	2.2%	5.9%	26.5%	$p < .001$
HS diploma	90.1%	89.9%	72.3%	
College degree or beyond	7.7%	4.2%	1.1%	
NLSY97 2015 Follow-up				
No degree	0.6%	1.6%	8.6%	$p < .001$
HS diploma	23.9%	36.7%	57.3%	
College degree or beyond	75.5%	61.8%	34.1%	

relatively low. Future research investigating explanations for these findings using novel approaches, perhaps to isolate causal effects, is important for helping this population of students.

The finding regarding degree completion also warrants further study. Given the many systemic and social changes in high schools from 1960 to 1997, we cannot be certain why the three samples yield different outcomes. One possibility is that because more students are encouraged to finish high school, and there are fewer work alternatives to finishing, students in 1997 were more likely to leave high school because of ability-related or academic reasons rather than needing to work or provide for a family. Exploring the academic trajectory of spatially talented students longitudinally could help to explain these patterns of degree completion and determine whether it is a modern phenomenon.

An ongoing concern in schools is the development of social–emotional learning (known as SEL). Our study focused on the adjustment of spatially talented students suggests that one possibility to enhance SEL may be better matching between academic talents and development rather than direct training of SEL itself, which remains unclear in its efficacy (Whitehurst, 2019). Future research may investigate the idea that improving SEL or adjustment of students may be first about ensuring their academic needs are fully met.

A final limitation and future direction is that this study did not explore whether effects of area of talent would be replicated across demographic groups. One reviewer was particularly interested in whether the effects of spatial talents would be similar for males and females. Looking at ANOVAs by sex, we found that the patterns for PT60 and HSB80 were identical for both sexes. For NLSY97, not all scales were administered to both sexes, but we did find that the effects of spatial talents were not present for females in terms of later reports of 'amount of work' and 'break school rules'. These differences could be a function of statistical power as there were about half as many spatially talented girls as boys. This suggests that future work should explore whether our findings replicate within different groups of students.

K-12 talent development implications

What continues to be a challenge in differentiating instruction through gifted and talented education programmes in K-12 schools is the need to identify a broader base of gifted students from low-income and underrepresented groups and help them develop those talents to the best of their capacity (Subotnik, Olszewski-Kubilius, & Worrell, 2011). Our findings show that spatially talented students with relative math and verbal reasoning weaknesses are falling through the cracks in the educational system due to a lack of attention and education tailored to their needs and that this pattern has persisted for the last 60 years. As Wai and Worrell (2016) have shown using similarly representative data sets, selecting on spatial reasoning would greatly improve the proportional representation of a pool of disadvantaged but gifted students who could then be matched with appropriate gifted education opportunities.

If the aim of public education is to help each student reach their full potential, students with spatial strengths should not need to wait until college for their talents to be valued, especially because talent development is a long and drawn out process which means that many students with primarily spatial strengths may not even make it onto college campuses (Wai & Worrell, 2016). And as we have shown, these students are less likely to have positive K-12 educational experiences and less likely to even complete college.

Higher education: STEM pipeline and middle skill job implications

Despite the connection between early spatial reasoning talent and prediction of later high-level STEM outcomes (Wai *et al.*, 2009), strategies to broaden the pool of individuals prepared by post-secondary education to fill STEM jobs at every level remain important (Autor, 2015; Holzer, 2015; Kell & Lubinski, 2013). More specifically, spatial reasoning skills may be particularly valuable to 'middle skill jobs' in STEM fields such as electricians, pipe fitters, advanced manufacturing machinists, brick masons, or radiology technicians (Wai & Uttal, 2018). Newman and Winston (2016) argue that 'More than 600,000 jobs remain open in the manufacturing sector alone. These are jobs that provide a middle-class wage without a traditional four year degree'. Preparing students for these middle-skills jobs and a faster path to productive employment could be an important outcome of identifying and serving spatial talents.

Policy implications

Using data from US K-12 schools, we demonstrated that 4–6% of students (over 2 million) are talented in spatial reasoning (in the top 5%) but are missed in academic selection procedures. Around the world, there are millions more who are spatially talented. This clearly indicates a huge pool of untapped talented students, including many who come from low-income backgrounds (Wai & Lakin, 2018; Wai & Allen, 2019). Thus, the first policy implication of our findings is that we need earlier and more systematic spatial talent identification and educational development. Secondly, innovation in curriculum and programme development needs to take place, including research studying which types of educational programming are effective for spatially talented students. Thirdly, teacher training and familiarity regarding spatial talents would help serve this neglected pool of students. This is especially critical because education majors have been shown to have the lowest average spatial ability of college majors (Wai *et al.*, 2009) and thus may not recognize or be prepared to nurture this type of talent. Finally, researchers should invest in explorations of the academic struggles of these talented students and what can be done to help such students through intervention.

Conclusion

Using three large-scale longitudinal databases spanning 60 years, we investigated whether students with strengths primarily in spatial reasoning were more likely to have academic struggles, behavioural problems, and other negative academic outcomes, including being less likely to complete college. We project that, based on these three US-population representative data sets, that roughly 2–3 million children in US K-12 education are currently overlooked for their talents, first at the stage of testing (which focuses on math and verbal measures) and secondarily at the stage of matching curricula. In sum, millions of students in the United States and many millions more around the world may be underperforming in school because their talents are not as valued and schools are not currently set up to serve their unique educational needs.

Acknowledgements

This research was supported by a grant from the American Psychological Foundation.

Author contributions

Joni M Lakin, Ph.D. (Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Writing – original draft; Writing – review & editing); Jonathan Wai, Ph.D. (Conceptualization; Formal analysis; Funding acquisition; Writing – original draft; Writing – review & editing).

Conflicts of interest

All authors declare no conflict of interest.

Data availability statement

The data that support the findings of this study were derived from the following resources available in the public domain:

American Institutes for Research. Project Talent, Base Year Data, 1960. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2013-05-23. https://doi.org/10.3886/ICPSR33341.v2

Bureau of Labor Statistics. National Longitudinal Survey of Youth, 1997. Washington, DC: National Longitudinal Survey Program. Retrieved from https://www.nlsinfo.org/content/cohorts/nlsy97

United States Department of Education. Institute of Education Sciences. National Center for Education Statistics. High School and Beyond, 1980: A Longitudinal Survey of Students in the United States. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-01-12. https://doi.org/10.3886/ICPSR07896.v2

References

Abdi, H. (2010). Holm's sequential Bonferroni procedure. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 573–577). Thousand Oaks, CA: Sage.

American Institutes for Research (n.d.). *Project Talent database*. Retrieve from https://www.air.org/project/project-talent

Andersen, L. (2014). Visual–spatial ability: Important in STEM, ignored in gifted education. *Roeper Review*, *36*(2), 114–121. https://doi.org/10.1080/02783193.2014.884198

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*(3), 3–30.

Barto, E., Gonzalez, M., Alley, W. E., Shore, C. W., Schwartz, K. L., & Weissmuller, J. J. (2010). *ASVAB subtest and composite homogenization* (Report No. AFCAPS-FR-2010-0019). Randolph AFB, TX: Air Force Personnel Center. Retrieved from http://apps.dtic.mil/dtic/tr/fulltext/u2/a553624.pdf

Barton, K., Cattell, R. B., & Silverman, W. (1974). Personality correlates of verbal and spatial ability. *Social Behavior & Personality: An International Journal*, *2*(2), 113–118.

Bicknell, B. (2008). Gifted students and the role of mathematics competitions. *Australian Primary Mathematics Classroom*, *13*, 16–20.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.

Child Trends, Inc, and the Center for Human Resource Research. (1999). *NLSY 97 Codebook Supplement Mainfile Round 1. Appendix 9: Family process and adolescent outcomes measures*. Washington, DC: Department of Labor.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Costello, A. B., & Osborne, J. W. (2011). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*, 1–9.

Coxon, S. V. (2012). The malleability of spatial ability under treatment of a FIRST LEGO League-based robotics simulation. *Journal for the Education of the Gifted*, *35*, 291–316. https://doi.org/10.1177/0162353212451788

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research*, *62*, 61–84. https://doi.org/10.3102/00346543062001061

Flanagan, J. C. & American Institutes for Research in the Behavioral Sciences (1972). *The project TALENT data bank: A handbook*. Palo Alto, CA: Author.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorman, W. A., Orr, D. B., & Goldberg, I. (1962). *Design for a study of American youth*. Boston, MA: Houghton Mifflin.

Gohm, C. L., Humphreys, L. G., & Yao, G. (1998). Underachievement among spatially gifted students. *American Educational Research Journal*, *35*, 515–531. https://doi.org/10.3102/00028312035003515

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: Similarities and differences between mental rotation and mental folding. *Cognitive Processing*, *14*, 105–115. https://doi.org/10.1007/s10339-013-0544-6

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A.Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121–169). New York, NY: Cambridge University Press.

Holzer, H. (2015). *Job market polarization and US worker skills: A tale of two middles*. Washington, DC: The Brookings Institution.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, *78*, 250–261. https://doi.org/10.1037/0021-9010.78.2.250

Kell, H. J., & Lubinski, D. (2013). Spatial ability: A neglected talent in educational and occupational settings. *Roeper Review*, *35*, 219–230. https://doi.org/10.1080/02783193.2013.829896

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*, 148–161. https://doi.org/10.1037/0022-3514.86.1.148

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. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA.

Lohman, D. F. (1994). Spatially gifted, verbally inconvenienced. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development* (Vol. *2*, pp. 251–264). Dayton, OH: Ohio Psychology Press.

Lohman, D. F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and assessment* (pp. 97–116). Hillsdale, NJ: Erlbaum.

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, *49*, 344–351. https://doi.org/10.1016/j.paid.2010.03.022

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*(3, Pt.1), 151–159.

Major, J. T., Wendy, J., & Deary, I. J. (2014). Linear and nonlinear associations between general intelligence and personality in Project TALENT. *Journal of Personality and Social Psychology*, *106*(4), 638–654.

Mix, K. S., Levine, S. C., Yi-Ling, C., Young, C. J., Hambrick, D. Z., & Konstantopoulos, S. (2017). The latent structure of spatial skills and mathematics: A replication of the two-factor model. *Journal of Cognition and Development*, *18*(4), 465–492.

National Center for Educational Statistics. (2018). *The Condition of Education 2018* (Report No. NCES 2018–144). Retrieved from https://nces.ed.gov/pubs2018/2018144.pdf

Newcombe, N. S., & Shipley, T. (2015). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying visual and spatial reasoning for design creativity*. New York, NY: Springer.

Newman, K. S., & Winston, H. (2016). Straight from high school to a career. *The New York Times*. Retrieved from https://www.nytimes.com/2016/04/15/opinion/straight-from-high-school-to-a-career.html

Pozzebon, J., Damian, R. I., Hill, P. L., Lin, Y., Lapham, S., & Roberts, B. W. (2013). Establishing the validity and reliability of the Project Talent Personality Inventory. *Frontiers in Psychology*, *4*, 968. https://doi.org/10.3389/fpsyg.2013.00968

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Sebring, P., Campbell, B., Glusberg, M., Spencer, B., Singleton, M., & Turner, M. (1987). *High school and beyond 1980 sophomore cohort third follow-up (1986) data file user's manual*. Washington, DC: National Center for Education Statistics.

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*, 604–614. https://doi.org/10.1037/0022-0663.93.3.604

Shepard, R. N. (1990). *Mind sights*. New York, NY: W. H. Freeman & Company.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, *12*, 3–54. https://doi.org/10.1177/1529100611418056

Talboy, A. N. (2011). Evaluation of the Armed Services Vocation Aptitude Battery (ASVAB). *Journal of Young Investigators*, *25*(11), 81–84.

US Bureau of Labor & Statistics (2006). *The NLS handbook*. Washington, DC: Author.

Uttal, D. H., Meadow, N. G., Elizabeth, T., Hand, L. L., Alden, A. R., Christopher, W., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402.

Wai, W., & Allen, J. (2019). What boosts talent development? examining predictors of academic growth in secondary school among academically advanced youth across 21 years. *Gifted Child Quarterly*, *63*(4), 253–272.

Wai, J., & Lakin, J. M. (2018). Hidden gems: Uncovering academic strengths among under-identified talent to help narrow achievement gaps. Unpublished manuscript.

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.

Wai, J., Putallaz, M., & Makel, M. C. (2012). Studying intellectual outliers: Are there sex differences, and are the smart getting smarter? *Current Directions in Psychological Science*, *21*(6), 382–390.

Wai, J., & Uttal, D. (2018). *Why spatial reasoning matters for education policy*. American Enterprise Institute (AEI) Education Policy Paper. Retrieved from https://www.aei.org/research-products/report/why-spatial-reasoning-matters-for-education-policy/

Wai, J., & Worrell, F. C. (2016). Helping disadvantaged and spatially talented students fulfill their potential: Related and neglected national resources. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 122–128.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, *99*, 397–420. https://doi.org/10.1037/0022-0663.99.2.397

West, T. G. (1991). *In the mind's eye*. Buffalo, NY: Prometheus Books.

Whitehurst, G. J. (2019). *A prevalence of "policy-based evidence-making"*. Education Next, 19. Retrieved from https://www.educationnext.org/prevalence-policy-based-evidence-making-forum-should-schools-embrace-social-emotional-learning/

Yoon, S. Y. (2011). *Psychometric properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations* (the Revised PSVT: R). (Unpublished Doctoral Dissertation). Texas A&M University, College Station, TX.

Received 26 August 2019; revised version received 18 December 2019