Sex Roles
DOI 10.1007/s11199-013-0269-0

ORIGINAL ARTICLE

Gender-Role Differences in Spatial Ability: A Meta-Analytic Review

David Reilly · David L. Neumann

Abstract Although gender-related differences in highly gender typed cognitive abilities are of considerable interest to educators and cognitive researchers alike, relatively little progress has been made in understanding the psychological processes that lead to them. Nash (1979) proposed a *gender-role mediation* hypothesis for such differences, with particular emphasis on spatial ability. However, changes in gender equality and gender stereotypes in the decades since merit a re-examination of whether a gender-role association still holds (Feingold 1988). A meta-analysis of 12 studies that examined gender-role identity and mental rotation performance was conducted. These included studies from the United Kingdom, Canada, Poland, Croatia, and the United States of America. The mean effect size for masculinity was r = .30 for men and r = .23 for women; no association was found between femininity and mental rotation. This effect size was slightly larger than that found previously by Signorella and Jamison (1986), and exceeds many other factors known to influence spatial ability. The implications of gender-role mediation of gender differences are discussed and future research directions are identified.

Keywords Gender differences · Spatial ability · Gender-role mediation · Gender roles · Mental rotation · Meta-analysis

D. Reilly (✉) · D. L. Neumann
School of Applied Psychology, Griffith University,
Southport, Queensland 4222, Australia
e-mail: d.reilly@griffith.edu.au

D. L. Neumann
Behavioural Basis of Health Program, Griffith Health Institute,
Queensland, Australia
e-mail: d.neumann@griffith.edu.au

Introduction

Though progress has been made in closing gaps in recent decades, women still remain underrepresented in science, technology, engineering and mathematics (STEM)-related fields in the United States with fewer women entering these fields in tertiary education (National Science Foundation 2011). Concerns about the underrepresentation of women are also present in many other countries, including Britain (Brosnan 1998) and Australia (Bell 2010). Although exceptions exist for psychology and medical sciences (Hyde 2007b), in general women are underrepresented in the sciences at a graduate level, as well scoring lower in tests of mathematics and science achievement at school within the U.S. (Gallagher and Kaufman 2005; Hedges and Nowell 1995). These findings are also supported by more recent reviews of mathematics and science literacy in large international assessments of student achievement such as the Programme for International Student Achievement (Else-Quest et al. 2010; Guiso et al. 2008; Reilly 2012), which assesses students worldwide as they reach the end of compulsory schooling. Much of the research in this area, however, draws on samples from America, and all studies cited herein are U.S.-based unless otherwise noted.

A consensus statement issued by major researchers in the area of gender-related cognitive differences identified research into the sources of individual differences in STEM achievement as an important priority (Halpern et al. 2007). When men and women are compared at the population level, reviews find no evidence of gender differences in general intelligence (Halpern and Lamay 2000; Neisser et al. 1996). However, researchers have frequently observed gender differences in more *specific* components of cognitive ability (Boyle et al. 2010a, b; Neumann et al. 2007, 2010). The size of such differences ranges from small to large, as a function of the cognitive component under investigation (Halpern et al. 2011). The largest and most consistent gender differences are found in spatial ability (Halpern 2011; Kimura

Published online: 03 April 2013

2000; Maccoby and Jacklin 1974), where reviews find effect sizes ranging from medium to large (Linn and Petersen 1985; Voyer et al. 1995). Gender differences in spatial ability are also found cross-culturally in large international studies with young-adult samples (Peters et al. 2006; Silverman et al. 2007).

The present review explores one such contribution to the development of spatial ability, that of gender roles. This term has been previously referred to in the literature as sex roles (Bem 1981; Constantinople 1973), but the term gender roles is preferred as it is broader and encompasses sociocultural factors as well as biological explanations for observed differences (Frieze and Chrisler 2011). Relationships between spatial ability, quantitative reasoning and gender roles are discussed before reviewing empirical support for gender-role associations with spatial ability. All studies cited herein are U.S.-based unless otherwise noted.

Spatial Ability and Quantitative Skills

Many researchers (e.g. Wai et al. 2009) have proposed that spatial ability provides a foundation for the development of quantitative reasoning such as science and mathematics (Nuttall et al. 2005; Serbin et al. 1990). Factor analyses of cognitive ability tests show high loadings for mathematical performance against a spatial factor (Carrol 1993; Halpern 2000). Furthermore, measures of spatial ability have predictive validity, in that they can predict future performance in quantitative fields (Williams and Ceci 2007). For example, Shea et al. (2001) followed a large group of intellectually talented boys and girls over a 20 year longitudinal study, from seventh grade until age 33. They found that individual differences in spatial, verbal, and quantitative reasoning in adolescence predicted educational and vocational outcomes two decades later. Further, spatial ability made a significant unique contribution even after controlling for verbal and mathematical ability (Shea, et al. 2001). Spatial ability is also predictive of college mathematical entrance scores (Casey et al. 1995, 1997), which are an important prerequisite for entry to further education in science and mathematics disciplines (Ceci et al. 2009).

Factors that influence spatial ability during development hold promise for educational interventions that seek to reduce the gender gap in science and mathematics in adulthood (Halpern 2007; Newcombe 2007). Hyde and Lindberg (2007, p. 29) argued that even mild improvement in spatial ability may have "multiplier effects in girls' mathematical and science performance". Additionally, higher levels of spatial ability are associated with attitudinal changes towards mathematics and self-confidence in mathematical ability from elementary school (Eccles et al. 1993) to high school and college (Eccles 1987; Eccles et al. 1990). Thus the contribution of spatial ability to later cognitive development may be in part social as well as intellectual (Crawford et al. 1995; Nash 1979). Academic domains where one feels competent and are seen as being socioculturally valued for one's gender are more likely to be pursued than those that are not (Eccles et al. 1990).

Although medium to large gender differences in spatial ability performance are found in most reviews of studies (Linn and Petersen 1985; Voyer et al. 1995), Hyde (2005) notes that *within*-gender variation is larger than *between*-gender differences. Since gender alone explains only a portion of individual variation in spatial ability (Caplan and Caplan 1994), identifying other developmental factors which promote spatial ability is an important research goal (Halpern et al. 2007; Hyde and Lindberg 2007). Neisser et al. (1996, p. 97) argued that understanding the source of such differences is critical, and that such questions are "socially, as well as scientifically important". One potential source of individual differences is that of gender-role identity.

Although the exact mechanisms contributing to the emergence of gender differences in spatial ability are debated (see Caplan and Caplan 1994 and Halpern 2011 for a discussion) they are believed to be influenced by a network of biological and sociocultural contributions (Ceci et al. 2009; Crawford et al. 1995; Eagly and Wood 1999; Halpern and Tan 2001). One such contribution is that of gender-role identity.

Though boys and girls typically differ in early socialisation experiences (Eccles et al. 1990; Emmott 1985; Lytton and Romney 1991), there is considerable individual variation in the degree to which they develop and acquire stereotypically *masculine* and *feminine* personality traits, behaviors and interests (Bem 1974; Constantinople 1973; Kagan 1964a). This process is referred to as *gender typing* (Kohlberg 1966; Kohlberg and Ullian 1974), and holds implications for the development of gender-role identity and integration of masculinity and femininity into an individual's self-concept and gender schema (Bem 1981; Knafo et al. 2005; Spence 1993). Highly gender typed individuals are motivated to keep their behavior and self-concept consistent with traditional gender norms (Bem1975; Bem and Lenney 1976; Maccoby 1990; Martin and Ruble 2004), and this also applies to academic domains (Nosek et al. 2002; Oswald 2008; Steffens and Jelenec 2011). Others may integrate aspects of both masculine and feminine identification into their self-schema, termed androgyny (Bem1984; Spence 1984).

Gender-Role Mediation of Spatial Ability

Nash (1979) proposed a *gender-role mediation* explanation for gender differences in which it is argued that gender-role identity can either promote or inhibit optimum development of cognitive ability in highly gender-typed domains, such as spatial and verbal ability. Specifically, Nash (1979) theorized that masculine identification leads to cultivation of spatial, mathematical, and scientific skills, whereas feminine identification facilitates verbal and language abilities.

In a review of gender-role influences on cognitive ability, Nash (1979, p. 263) wrote "For some people, cultural myths are translated into personality beliefs which can affect cognitive functioning in gender-typed intellectual domains". This argument was based on earlier work by Sherman (1967) into differential learning and practice experiences of boys and girls. In doing so, Nash extended Sherman's theory by placing cognitive development of spatial ability in a *social* context, where gender-role identity encourages or discourages optimum development of spatial potential. Nash identified several mechanisms that contribute to spatial development, including gender typing of intellectual domains, gender-role conformity and self-efficacy beliefs.

Differential Spatial Experiences

Sherman (1967) hypothesized a causal explanation for the presence of gender differences in spatial ability, based on a child's differential opportunities to develop and refine spatial skills through play and recreational activities. Boys and girls typically differ in their socialisation experiences, and are encouraged by parents to engage in either stereotypically masculine or feminine play appropriate to their gender (Eccles et al. 1990; Lytton and Romney 1991). However, play is also an opportunity for active engagement and cognitive development (Piaget 1968). Caplan and Caplan (1994) argued that traditionally "masculine" typed activities promote the development of spatial ability by encouraging the practice and application of spatial skills (Connor and Serbin 1977). In contrast, traditionally "feminine" activities do not require the use of spatial skills, but reinforce other socially valued skills (Lever 1976).

What distinguishes Sherman's (1967) explanation from other explanations (such as Caplan and Caplan 1994) is that it focuses specifically on gender roles, rather than solely on biological gender, as explaining *individual* differences in spatial ability. Differential practice of skills promoting spatial development occur through gender typing of activities and interests (Serbin and Connor 1979; Serbin et al. 1990). Rather than assuming that the lives of boys and girls do not overlap, or that *all* boys engage in a high level of activity and receive equal opportunities to practise and develop spatial ability, it accounts for individual differences and gender typing. There is evidence to support this argument. Retrospective studies have shown that an association exists between spatial ability and activity preferences in young adult college-level samples (Baenninger and Newcombe 1989; Signorella et al. 1989).

Gender Typing of Intellectual Domains

Kagan (1964b) noted that objects in the everyday world, social activities, and even intellectual pursuits become gender typed as either masculine or feminine, based on shared consensual beliefs that emerge very early in childhood. For example, reading and language is regarded as being feminine (Dwyer 1973, 1974), whereas mathematics, science and technology are regarded as masculine (Li 1999; Nash 1975). Both at an implicit (Lane et al. 2012; Nosek et al. 2009; Steffens and Jelenec 2011) and an explicit level (Benbow 1988; Halpern and Tan 2001), cultural beliefs about specific cognitive tasks as being inherently *masculine* or *feminine* prevail - even for generations growing up with increased gender equality (Liben et al. 2002). Recently, Halpern et al. (2011) showed that lay beliefs about cognitive gender differences in student and community samples were firmly entrenched across *both* men and women. Although these stereotypes are not an accurate reflection of reality, Nash (1979) argued they have the potential to shape the self-concepts of boys and girls, and how they see themselves in relation to these academic domains (Hyde and Lindberg 2007).

Gender-role Conformity Pressures

Gender roles and associated stereotypes describe differences between men and women, and prescribe how they should behave in social and occupational settings (Eagly and Mitchell 2004). Highly gender typed persons are motivated to keep their behavior consistent with internalised gender-role standards and norms (Bem and Lenney 1976), whereas those low in gender typing or for whom gender-role identity is less salient show greater cognitive and behavioral flexibility (Arbuthnot 1975; Bem 1975; Stein and Bailey 1973). Conformity cues as to who should engage in certain behaviors, and what activities are permissible for boys or girls, come from peers, parents, and the media (Martin and Ruble 2004; Matthews 2007), and this has implications for intellectual domains that are masculine or feminine dominated (Eccles 2007).

Nash (1979) argued that the increased saliency of gender and gender typing of academic subjects in adolescence may lead to a conflict between the "ideal" image a student holds of himself or herself, and the activities he or she chooses to perform well in and values. Perceived incompatibility between being "feminine" and succeeding in stereotypically "masculine" domains can hinder academic achievement (Rosenthal et al. 2011; Schmader 2002). Thus there is also an attitudinal and motivational component to development of intellectual abilities (Nash 1979).

Self-efficacy Beliefs and Gender Stereotypes

During childhood when gender-role saliency is low, boys and girls show relatively little difference in intellectual abilities, and what differences exist often favors girls (Halpern 2000; Nash 1979). However, gender typing of intellectual pursuits quickly emerges in adolescence (Dwyer 1974; Kagan 1964b), and leads to several negative psychological consequences for some children (Nash 1979). Firstly, girls and boys receive different messages about occupational aspirations and the usefulness of specific academic skills (Fennema and Sherman 1977; Hyde and Lindberg 2007). Secondly, as noted earlier, gender typing of intellectual tasks is often seen as being incompatible with a feminine gender-role identity at a time when conformity pressure increases (Eccles 2007; Hoffman 1972; Rosenthal et al. 2011). This can result in lowered self-esteem and reduced self-efficacy beliefs for gender-typed tasks (Pajares and Miller 1994). Gender stereotypes suggest that men and women are better at some tasks than others, and this is reflected in self-estimations of intelligence in gender-typed domains (for a review see Szymanowicz and Furnham 2011). Additionally, a large body of research has observed that a feminine gender typing is associated with considerably lower self-esteem than masculine or androgynous individuals (Spence et al. 1975; Whitley 1983, 1988), including academic self-esteem (Alpert-Gillis and Connell 1989; Lau 1989; Robison-Awana et al. 1986).

Evidence for a Spatial-Gender-Role Association

A prior meta-analysis by Signorella and Jamison (1986) found support for Nash's hypothesis in spatial ability. However there have been major and potentially relevant changes in gender roles and stereotypes in the intervening decades (Auster and Ohm 2000; Hyde and Lindberg 2007) which Feingold (1988) has argued are responsible for declining gender differences in cognitive ability. This view is supported by Hyde (2005, 2006, 2007) and colleagues across a range of intellectual abilities (Hyde 2007a; Lindberg et al. 2010). These changes question the validity of Nash's theory in contemporary society and whether such gender-role associations still exist today. For this reason, we aimed to conduct a meta-analysis of studies published since Signorella and Jamison's (1986) review, to see whether the gender-role mediation hypothesis still holds. Although these studies are primarily based on research conducted in the USA, studies from other nations (e.g., Poland, Croatia, United Kingdom, Canada) are also examined for a broader test of Nash's theory.

Meta-analysis provides researchers with a way to critically evaluate the cumulative evidence of empirical evidence (Rosenthal 1984), and the technique is becoming increasingly common in psychology (Hyde 1990; Rosenthal and DiMatteo 2001). Although individual studies taken in isolation might show that a relationship between factor *X* on ability *Y* may be present or absent, factors such as random sampling error and lack of statistical power may result in erroneously rejecting the null hypothesis (Type I error) or failing to detect an effect that is real (Type II error). The technique of meta-analysis allows one to draw firmer conclusions about the existence of an association (Rosenthal and DiMatteo 2001), as well to arrive at an estimate of its size that is more accurate and reliable than could be determined from a single empirical study.

A requirement of meta-analysis is that empirical studies measure a similar construct drawn from similar samples (R. Rosenthal 1984, 1995), and that there are a sufficient number of studies to make meaningful conclusions. Spatial ability is not a unitary construct; it encompasses at least three separate processes – spatial perception, visualisation, and mental rotation (Linn and Petersen 1985). Mental rotation is one of the most widely researched areas of cognitive gender differences (Halpern and Lamay 2000), due in part to the fact comparisons of men and women in mental rotation show the largest effect sizes of all spatial tasks (Voyer et al. 1995). Some researchers regard mental rotation to be a representation of general spatial reasoning (Casey et al. 1995; Halpern 2000; Vandenberg and Kuse 1978), and there is evidence that performance in mental rotation prospectively predicts later development of quantitative reasoning (Casey et al. 1997; Nuttall et al. 2005). Therefore this review is confined to studies that investigated performance in mental rotation tasks. In addition, gender differences are larger after late adolescence when gender roles become particularly salient (Nash 1979). There are also issues of reliability and validity when assessing gender roles in younger samples. For this reason, only studies using high school, college or young adult samples were considered for inclusion in the reported meta-analysis. Studies using younger samples, such as that by Titze et al. (2010), were not considered.

In sum, the present review involved a meta-analysis of studies that have investigated gender-role associations with mental rotation task performance. It was hypothesized (Hypothesis 1) that masculinity would be positively associated with greater mental rotation performance in men and women. The influence of femininity was also investigated as a research question. It was hypothesised (Hypothesis 2) that there would be a negative association between femininity and mental rotation performance for both genders. Since the magnitude of gender differences typically varies with the type and level of difficulty of mental rotation task (Voyer et al. 1995), we also examined the type of mental rotation instrument as a potential moderator. Similarly, because there have been debates over which measures of masculinity and femininity are the best predictor of behavior (Bem 1984; Spence and Buckner 2000), we examined the type of gender-role instrument as a potential moderating variable.

Method

Search Strategy

To access as many studies as possible, a number of search strategies were used. Firstly, a Web of Science citation search for articles citing either Nash (1979) or Signorella and Jamison (1986) was performed, as any study published that is relevant to the meta-analysis would be likely to cite these key articles. Secondly, GoogleScholar and PsycINFO searches were performed for studies containing the keywords "spatial ability" or "mental rotation" and any combination with the keywords "masculine", "masculinity", "androgynous", or "androgyny". This second method identified a number of additional studies that were not *specifically* testing a gender-role mediation hypothesis, but merely included a gender-role measure and mental rotation task as part of a larger battery of neuropsychological tests (e.g., Rahman et al. 2004). Furthermore, an attempt to locate unpublished studies was made by searching the *Dissertation Abstracts* and *ERIC* databases for studies, locating one additional study. The search was performed in September, 2012.

Selection Criteria

The following inclusion criteria were used:

- peer-reviewed empirical studies published after 1986 or unpublished manuscripts and reports dated after 1986
- gender-role identity was measured using a psychometrically valid and reliable gender-role instrument, such as the Bem Sex Role Inventory (BSRI; Bem 1974) or the Personal Attributes Questionnaire (PAQ; Spence et al. 1974)
- participants sampled were either an adult or high school aged adolescent, from a non-clinical sample

Requests to authors ($n = 5$) for additional information were made where a masculinity and mental-rotation association was not explicitly tested or reported. Three studies could not be included due to insufficient information to determine an effect size (Evardone and Alexander 2009; Tuttle and Pillard 1991; Vonnahme 2005). One practice sometimes adopted is to consider all studies missing an effect size to have an association with an absolute value of zero, a practice that Rosenthal (1995) considers overly conservative and leads to inaccurate estimates. This practice was considered at length by Hedges and Becker (1986) who caution against missing value substitution. Accordingly the decision was made to exclude these missing studies. Following application of the selection and exclusion criteria, there were 12 available studies examining mental rotation and gender roles. However it should be noted that the possibility of unpublished null studies (commonly termed the "file drawer problem") is addressed using meta-analytic techniques that test for publication bias (Orwin 1983; Rosenthal 1979).

Sample Characteristics

The characteristics of all studies identified in the literature search are presented in Table 1. Several of the studies recruited participants from different countries, making for a broader test of Nash's hypothesis than would be possible if analysing only data from the USA. It should be noted that in most studies, samples were drawn almost exclusively from student subject pools, limiting generalisability somewhat to a young-adult, college-level educated sample.

Procedure

Comprehensive Meta Analysis (CMA) V2 software was used for the calculation of statistics (Borenstein and Rothstein 1999). A random-effects model was chosen (Borenstein et al. 2009) because spatial ability is subject to a large number of psychosocial moderators, and a variety of different gender-role instruments and mental rotation tasks were used over multiple decades. The random effects model gives slightly wider confidence intervals than a fixed-effects model (Field 2001; Rosenthal and DiMatteo 2001), but gives a more appropriate estimate of how much variability is present in empirical studies (Kelley and Kelley 2012).

The focus of the review was the relationship between gender-role identity and mental rotation, which can be represented by Pearson's product moment correlation, r. Gender-role instruments offer separate masculinity and femininity scales, allowing us to consider the effect of masculinity independently of femininity, and to test both for a mental rotation association.

Where the direct product–moment correlation between gender-role masculinity scale and mental rotation was reported, this was used because it represents the direct association independent of a subject's femininity scale. However, two studies reported only the mean values for masculine, feminine, and androgynous groups. Since androgyny represents a "special case", and some theorists argue that such participants cannot be legitimately combined with either the masculine or feminine group (Taylor and Hall 1982), the androgynous participants were excluded as per Signorella and Jamison's (1986) recommendation. Such an approach is the most conservative strategy available, and may lead to an underestimation of the true effect size in cases where androgynous participants (high masculinity, high femininity) score higher than their masculine or feminine counterparts (e.g. Hamilton 1995). By doing so, however, it affords a simple comparison between masculine and feminine participants only, allowing for the use of Cohen's d and then conversion to r as the common effect size unit using the formula given by Rosenthal (1984). Several studies recruited male or female participants only, and in several cases examined only masculinity associations. Calculations were performed using the CMA software.

Table 1 Characteristics of the studies identified from the literature search on mental rotation performance and gender roles

Study	Country	Sample Type	Sample Age	N	Gender Role	Mental Rotation	Males		Females	
							Masculinity (r)	Femininity (r)	Masculinity (r)	Femininity (r)
Jamison and Signorella (1987)	USA	High school students	8th grade	10 females 19 males	BSRI	CRT	$.45^{*}$	.04	.27	.14
Signorella et al. (1989)	USA	Subject pool	n/r	132 females 156 males	BSRI	CRT	.08	.01	$.24^{*}$	.05
Gilger and Ho (1989)	USA	Subject pool	M=19.0	52 females 38 males	BSRI	TSRT	.00	−.17	.00	−.17
Voyer and Bryden (1990) [a]	Canada	Subject pool	M=21.0	65 females 65 males	BSRI	VMRT	$.50^{**}$	-	.21	-
Tuttle and Pillard (1991)	USA	Community	Range 25–40	88 females 101 males	CPI	TSRT	n/r			
Newcombe and Dubas (1992)	USA	Longitudinal	16	61 females; attrition rate = 29 %	PAQ	TSRT	-	-	.15	-.10
Hamilton (1995) [b]	United Kingdom	Community, school and college	M=18.0	122 females 54 males	BSRI	SMRT	.12	−.12	.14	−.14
Jagieka and Herman-Jeglinska (1998) [a e]	Poland	Subject pool	n/a	30 males	BSRI	SMRT	$.34^{*}$	-	-	-
Saucier et al. (2002)	Canada	Subject pool	M=22.8	54 females 41 males	PAQ	VMRT	$.45^{**}$	−.02	$.45^{***}$	−.02
Rahman et al. (2004) [d e]	United Kingdom	Community	Range 18–40	120 females 120 males	EPP	VMRT	$.41^{***}$	-	$.23^{**}$	-
Ritter (2004) [c]	United Kingdom	Subject pool	M=21.0	37 females 42 males	BSRI	SMRT	$.34^{*}$	−.26	−.18	−.14
Scarbrough and Johnston (2005)	USA	Subject pool	M=19.6	41 females	BSRI	CSMRT	-	-	$.40^{**}$	.00
Vonnahme (2005)	USA	Subject pool	M=21.2	46 males	BSRI	CMRT	n/r	n/r		
Hromatko et al. (2008)	Croatia	Unspecified	M=24.8	26 females	BSRI	TSRT	-	-	$.64^{***}$	.03
Evardone and Alexander (2009)	USA	Subject pool	M=20.0	52 females 58 males	BSRI	VMRT	n/r	n/r	n/r	n/r

[a] Calculated from p value

[b] Androgynous (high masculinity, high femininity) eliminated * $p < .05$; ** $p < .01$; *** $p < .001$ two-tailed

[c] Calculated from median split of masculinity/femininity

[d] Data provided by author

[e] Data only available for masculinity n/r = data not reported

CPI California Psychological Inventory; *EPP* Eysenck Personality Profile (EPP; Eysenck et al. 1996); *CRT* Card Rotation Test (French et al. 1963); *TSRT* Thurstone Spatial Relations Test (Thurstone 1958); *SMRT* Shepard and Metzler Mental Rotation Test (Shepard and Metzler 1971); *VMRT* Vandenberg Mental Rotation Test (Vandenberg and Kuse 1978); *CMRT* Cooper and Shepard MRT (Cooper and Shepard 1973)

Meta-analytic Results

Study characteristics and effect sizes are presented in Table 1. Since empirical studies using gender roles frequently find gender × gender-role interactions, the associations with masculinity and femininity are reported separately for men and women. Forest plots are provided when gender-role associations are statistically significant. A forest plot conveys a visual representation of the effect size estimates of individual studies and their variability (Lewis and Clarke 2001); one can see the amount of variation between individual studies as well as the overall trend. In the centre of each study's confidence interval is a square; the size of the square corresponds to the sample size used in each study. The diamond symbol represents the overall estimate of the sample, with the centre of the diamond being the point estimate and its horizontal tips representing the confidence interval.

Girls and Women

Figure 1 presents a forest plot of the association between masculinity and mental rotation performance for girls and women, and effect sizes are given in Table 1. Hypothesis 1 predicted that masculinity would be positively associated with greater mental rotation performance. As shown in Fig. 1, most studies with female samples were in a direction consistent with this hypothesis with the exception of two studies: Gilger and Ho (1989) found no association, whereas Ritter (2004) found a weak negative association. The distribution of effect sizes across studies was heterogenous, $Q(10) = 21.13$, $p = .020$, $I^2 = 52.67$ indicating moderate variability across studies. It is also noteworthy that the two largest associations were found in the non-USA samples of Croatia ($r = .64$) and Canada ($r = .45$). However, the size of the correlation is unlikely to be culturally related given that the third largest association was found in a USA sample ($r = .40$) and that small associations were also found in non-USA samples (e.g., $r = -.18$ for Ritter 2004; $r = 14$ for Hamilton 1995).

In support of Hypothesis 1, the combined masculinity effect size for women was $r = .23$ (95 % CI lower = .11, upper = .34), $Zma = 3.72$, $p < .001$. This correlation for women was only slightly larger than that found by Signorella and Jamison (1986), who found a significant association of $r = .19$ between masculinity and mental rotation for girls and women using androgyny measures. To put these findings into perspective, we employed Rosenthal's Binomial Effect Size Difference (BESD; R. Rosenthal and Rubin 1982), a metric that represents effect size in a format suitable for interpretation by non-statisticians (R. Rosenthal and DiMatteo 2001). Represented in the BESD format, the likelihood of being average or higher in mental rotation performance increases from 38.5 % for feminine women to 61.5 % for masculine or androgynous women.

The possibility of unpublished null studies (referred to as the "file drawer problem") was also addressed by the calculation of Orwin's Fail-Safe N, which estimates the number of null studies required to reduce mean effect sizes to a specific cutoff-point (Borenstein et al. 2009; Orwin 1983). Employing Orwin's calculation, it would take only two more null studies to reduce the association to that found previously by Signorella and Jamison (1986); therefore the stronger association in these studies should be taken only tentatively.

Hypothesis 2 predicted that there would be a significant negative association between femininity and mental rotation performance. This hypothesis was not supported, $r = -.05$, $p =$ *n.s.* Such a finding is also consistent with the findings of Signorella and Jamison (1986) who failed to find any association between femininity and mental rotation performance.

Boys and Men

The forest plot of the association between masculinity and mental rotation performance for boys and men is shown in Fig. 2 and it presents the second test of Hypothesis 1. As can

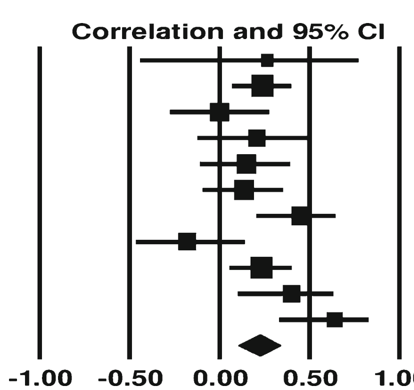

Fig. 1 Forest plot of masculinity association with mental rotation performance for girls and women. Positive associations indicate better mental rotation performance as masculinity increases. The combined effect size is represented as a *diamond shaped* correlation

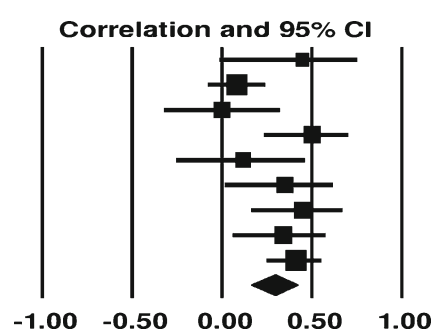

Fig. 2 Forest plot of masculinity association with mental rotation performance for boys and men. Positive associations indicate better mental rotation performance as masculinity increases. The combined effect size is represented as a *diamond shaped* correlation

be seen from the figure, the scores of men were slightly wider in variability than for women, with many studies showing relatively large associations while three showed relatively weak or non-significant correlations. Similar to the results for women, there did not appear to be a strong relationship between the country the study was conducted in and the size of the association. The largest association was found in a sample of men from Canada (r = .50), but the equal second largest association was found in a USA male sample (r = .45). However, it is noteworthy that the two remaining studies with USA samples did not find any significant association (r = .08 and r = .00). The distribution of effect sizes across all studies was heterogenous, $Q(8) = 17.92$, $p = .022$, $I^2 = 55.36$, indicating moderate variability between studies.

In support of Hypothesis 1, the association between masculinity and mental-rotation performance for men was significant, r = .30, (95 % CI lower = .16 upper = .42), Zma = 4.25, $p < .001$. Again, the association is slightly larger than that estimated by Signorella and Jamison (1986), who reported an r = .15 between masculinity and mental rotation performance for boys and men. Orwin's Fail-safe N showed that it would take an additional eight unpublished studies with a mean association of zero to reduce this correlation to the size found in the earlier review (r = .15). Represented in the BESD format, the likelihood of being average or higher in mental rotation performance increases from 35 % for feminine boys and men to 65 % for those with a masculine or androgynous gender-role identity. Finally, in contrast to Hypothesis 2, no association was found between femininity and mental rotation for boys and men, $r = -.06$, p = *n.s.*

Moderating Variables

Since there was moderate between-study heterogeneity in the masculinity association for both men and women, it is important to determine potential moderators that may be responsible such as the type of gender-role instrument used to classify participants, or the nature of the mental rotation task. Alternately, instruments might vary in their predictive validity for men and women, and this information might be useful in planning future research. Accordingly, effect sizes and heterogeneity were examined for men and women separately across gender-role instrument.

Tables 2 and 3 present associations across type of gender-role instrument for men and women respectively. While the BSRI was used most frequently, the strongest gender-role associations were found with the PAQ for both men and women. However with an insufficient number of studies employing gender-role measures other than the BSRI, any conclusions made about the predictive validity of these instruments are tentative.

Another potential source of heterogeneity is the nature of the mental rotation task employed. Meta-analytic reviews have found that the magnitude of gender differences differs across instruments (Voyer et al. 1995). It seems likely, therefore, that similar variation would be present when considering gender-role associations. Table 4 presents effect sizes for studies grouped by mental rotation instrument. Instruments were grouped into four categories. These groupings reduced heterogeneity, suggesting that much of the variability observed across studies was the result of using different instruments for measuring mental rotation. It should also be noted that the Vandenberg instrument also produced the highest gender-role effect size of any mental rotation task. This may reflect the increased difficulty of this instrument which allows for greater differentiation between high and low ability (Voyer et al. 1995).

Discussion

The present meta-analysis examined evidence for Nash's (1979) gender-role mediation hypothesis of spatial ability, as measured by performance on mental-rotation tasks. In a previous review, Signorella and Jamison (1986) found a small but statistically significant association between gender role and mental rotation performance. The present results support the conclusions drawn by Signorella and Jamison (1986). There is a significant and medium sized association between masculinity and mental rotation in research conducted in the past 25 years. The size of the association did not appear to be strongly related to the country in which the study was conducted, although there was some evidence that the type of mental rotation task and gender-role measure used in the study was a factor. The present meta-analysis also showed that there was no association between femininity and mental rotation performance.

Table 2 Effect size and heterogeneity by gender-role instrument for men

Type of instrument	N of studies	Effect size *(r)*	*Zma, p-value*	Heterogeneity
Bem Sex-Role Inventory	7	.25	$Z = 3.07$, $p = .002$	$Q(6) = 11.73$, p = *n.s.*
Personal Attributes Questionnaire	1	.45	Z = 2.20; $p = .028$	*N/A*
Eysenck Personality Profiler	1	.41	Z = 2.50; $p = .013$	*N/A*

Total heterogeneity within-groups, $Q(6) = 11.74$, $p = .068$; between-groups, $Q(2) = 6.18$, $p = .045$

Table 3 Effect size and heterogeneity by gender-role instrument for women

Type of instrument	*N* of studies	Effect size *(r)*	*Zma, p-value*	Heterogeneity
Bem Sex-Role Inventory	8	.21	$Z = 2.43$, $p = .015$	$Q(7) = 17.07$, $p = .017$
Personal Attributes Questionnaire	2	.30	Z = 1.97; $p = .049$	$Q(1) = 3.02$, $p =$ *n.s.*
Eysenck Personality Profiler	1	.23	Z = 1.17; $p =$ *n.s.*	*N/A*

Total heterogeneity within-groups, $Q\ (8) = 20.09$, $p = .010$; between-groups, $Q(2) = 1.04$, $p =$ *n.s.*

The results of this meta-analysis demonstrate three important things. Firstly, it upholds the claims made by Nash (1979) that, at least for mental rotation tasks, masculine gender roles contribute to the development of spatial ability. Although only correlational in nature, the inclusion of the longitudinal study by Newcombe and Dubas (1992) shows that gender roles have predictive validity for later development of spatial ability. Secondly, this review demonstrates the persistence of gender roles over a larger span of time, in that studies reviewed are drawn from three decades of research; it would appear that the empirical findings of Nash and others were not a statistical quirk, or an artefact of prevailing gender inequalities of the past. Thirdly, the review shows that the magnitude of the gender-role association may be somewhat larger than previously thought by researchers, especially for men.

A possible explanation for finding a stronger association between gender roles and spatial ability than Signorella and Jamison (1986) is the quality of instruments used across studies. Many of the earlier studies reviewed by Signorella and Jamison (1986) used instruments that operationalised masculinity and femininity as bipolar opposites of a unidimensional construct (Constantinople 1973) rather than orthogonal aspects of gender-role identity (Bem 1981). This leads to misclassification of masculine, feminine, and androgynous participants (Bem 1974, 1977) and an attenuation of effect size due to imprecision (Cooper 1981). It is difficult to suggest a theoretical reason why gender roles might influence cognitive development *more* strongly now in men than in previous decades, but this possibility cannot be ruled out entirely.

A growing trend in empirical research is a move away from levels of statistical significance towards evaluations of the magnitude of effect sizes (Wilkinson 1999), to assess their practical impact and importance. Cohen (1988) provides a good rule of thumb to gauge associations by: correlations of .10 or higher are regarded as small, .30 or higher as medium, and correlations higher than .50 are considered large. Frequently these yardsticks are used rather rigidly, and some researchers regard differences that are "small" as "trivial" or non-existent (Hyde 1996, 2005). Cooper (1981) warns against this practice, as the magnitude of effects that may be found can differ greatly from one field of psychological research to another. Similarly, in a review of effect sizes and practical importance for research with children, McCartney and Rosenthal (2000) advise against such yardsticks, and caution that effect sizes should be compared to those found in that particular research domain. For this reason, comparisons to a range of other effects deemed previously to be influential in spatial ability may be better able to put the results of this meta-analysis in context (Hyde 1990).

The present results showed a gender-role association of $r = .30$ for men and $r = .23$ for women. Two areas previously documented to contribute to spatial ability are prior spatial activity preferences in childhood (Signorella et al. 1989) and socioeconomic status (Levine et al. 2005). A meta-analysis by Baenninger and Newcombe (1989) produced an $r = .10$ between spatial activity preferences and spatial ability. Levine et al. (2005) found that spatial ability differences are found between low, medium, and high socioeconomic status groups for adolescents with an effect size $r = .23$ for mental rotation. When compared to these factors, which researchers have previously argued to be important and have a meaningful impact on spatial ability, the contribution of gender role and mental rotation is greater, and may go some

Table 4 Effect size and heterogeneity by mental rotation instrument

Type of instrument	*N* of studies	Effect size *(r)*	*Zma, p-value*	Heterogeneity
Card Rotations Task (French et al. 1963)	2	.22	$Z = 1.81$, $p = .071$	$Q(1) = 1.44$, $p =$ *n.s.*
Thurstone Spatial Relations (Thurstone 1958)	3	.21	Z = 1.83; $p = .058$	$Q(2) = 10.38$, $p = .006$
Vandenberg MRT (Vandenberg and Kuse 1978)	3	.38	Z = 4.29; $p < .001$	$Q(2) = 1.41$, $p =$ *n.s.*
Generic Mental Rotation Tasks	4	.22	Z = 2.43; $p = .015$	$Q(3) = 3.91$, $p =$ *n.s.*

Total heterogeneity within-groups, $Q\ (8) = 17.14$, $p = .029$; between-groups, $Q(3) = 10.47$, $p = .015$

way to explaining existing gender differences in spatial ability (Nash 1979; Sherman 1967).

Implications for Spatial Development

One possible intervention being considered for individuals most at risk of forestalled spatial development is that of spatial training. In a review article, Newcombe and Frick (2010) stress the importance of early intervention in the development of spatial abilities during early childhood. Although it would be desirable to offer spatial instruction and training for *all* students to address this gender-gap (Hyde and Lindberg 2007), competing interests in an ever crowded curriculum make the likelihood of this practice being adopted rather bleak; indeed few schools incorporate spatial ability specifically into the curriculum during elementary school (Mathewson 1999; Newcombe and Frick 2010). A more practical measure might be for limited intervention programs to target at-risk students, in the same way that reading and literacy interventions are offered for students struggling in these areas.

While screening directly for spatial deficits may be possible, large gender differences do not typically emerge until adolescence (Linn and Petersen 1985; Voyer et al. 1995). Early intervention is desirable before such differences emerge (Newcombe and Frick 2010). The assessment of gender roles might serve as a more useful risk factor to consider than gender, and it has the advantage of not necessarily being limited to one gender. Nuttall et al. (2005) describe gender-role appropriate intervention programs that develop spatial expertise, but as of yet, there are no longitudinal studies of such programs. Educators may wish to be mindful to include a range of opportunities that encourage spatial development as well as stressing their importance and relevance to *both* boys and girls.

Newcombe and Frick (2010) also advocate early intervention by parents, in providing children with activities and opportunities outside the classroom to develop spatial awareness, perception and visualisation. Rigidly held gender roles restrict children's self-selection of activities (Ruble et al. 2006; Tracy 1987), and parents may wish to encourage a broader repertoire in their children including sports and toys that encourage spatial development (Doyle et al. 2012). The continuing failure to find a negative relationship between femininity and spatial ability for both genders is also noteworthy. Feminine identification should not be discouraged in order to develop spatial and quantitative ability.

Future Directions for Research

Although gender differences in cognitive ability are frequently debated, many researchers note there is greater within-gender variability than *between* men and women (Hyde 1990; Priess and Hyde 2010). Gender-role identity appears to be an important, but previously underestimated contributor to these individual differences in spatial ability, which in turn is a key foundation for higher-level quantitative skills such as mathematics (Casey et al. 1997; Delgado and Prieto 2004) and STEM related fields (Halpern 2007; Newcombe 2007). Indeed, Halpern (2007, p. 125) has claimed that spatial ability is "essential" for success in STEM-related subjects. As such, the emergence of gender roles as a factor that meets or exceeds other factors that contribute to spatial ability is important, both as a potential diagnostic indicator for interventions as well as a focus for future investigation. By better understanding the psychosocial processes associated with gender roles and intellectual development, one might be able to identify strategies - such as self-efficacy training or challenging of gender stereotypes - that would help negate performance impairments.

Additionally, this meta-analysis affirms the merit of considering gender roles, rather than just biological gender, in studies of individual differences in cognition. Though this review was confined to only mental-rotation, it remains to be seen whether the results can be generalised more widely to other spatial ability tasks such as spatial perception and visualisation (Linn and Petersen 1985). For example, is there something specific about a masculine or androgynous gender role that leads to improved ability to perceive spatial objects and mentally rotate them, or can it be generalised to other spatial tasks? This would allow us to test whether gender-role differences in perception are chiefly responsible, or whether there are differences in the actual cognitive processes underlying such tasks, for example a general cognitive style (Arbuthnot 1975; Milton 1957). A limited number of studies with adolescents and young adults have considered the Piaget water-level task (Jamison and Signorella 1980; Kalichman 1989; Popiel and De Lisi 1984; Signorella and Jamison 1978) or the Embedded Figures Test (Bernard et al. 1990; Brosnan 1998; Hamilton 1995), with some inconsistencies, but larger studies are required. Furthermore, as Signorella and Jamison (1986) note, Nash's (1979) hypothesised associations between gender-role identity and verbal ability remain largely untested, which future studies should pursue.

Conclusion

We have seen many changes in society's beliefs about gender equality in the intervening decades since Nash (1979) proposed her gender-role mediation hypothesis of intellectual development. However, for spatial ability at least, this association seems as relevant today as when the claim was first made. The results from our meta-analysis support Nash's hypothesis for the development of spatial ability, and this provides strong support for calls to conduct further research in this area to investigate the cognitive and social processes that underlie the association between gender-roles and cognitive abilities.

Acknowledgments This research was supported in part by a Griffith University Postgraduate Research Scholarship. Thanks go to Dr Heather Green, Dr Michael Steele, Dr Elizabeth Conlon, and Dr Margaret Signorella for early revisions of this manuscript.

References

Alpert-Gillis, L. J., & Connell, J. P. (1989). Gender and sex role influences on children's self-esteem. *Journal of Personality, 57*, 97–114. doi:10.1111/j.1467-6494.1989.tb00762.x.

Arbuthnot, J. (1975). Sex, sex-role identity, and cognitive style. *Perceptual and Motor Skills, 41*, 435–440. doi:10.2466/pms.1975.41.2.435.

Auster, C. J., & Ohm, S. C. (2000). Masculinity and femininity in contemporary American society: A reevaluation using the Bem Sex-Role Inventory. *Sex Roles, 43*, 499–528. doi:10.1023/A:1007119516728.

Baenninger, M., & Newcombe, N. S. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles, 20*, 327–344. doi:10.1007/BF00287729.

Bell, S. (2010). Women in science: The persistence of gender in Australia. *Higher Education Management and Policy, 22*, 1–19. doi:10.1787/hemp-v22-art3-en.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162. doi:10.1037/h0036215.

Bem, S. L. (1975). Sex role adaptability: One consequence of psychological androgyny. *Journal of Personality and Social Psychology, 31*, 634–643. doi:10.1037/h0077098.

Bem, S. L. (1977). On the utility of alternate procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology, 45*, 196–205. doi:10.1037/0022-006X.45.2.196.

Bem, S. L. (1981). Gender Schema Theory: A cognitive account of sex typing. *Psychological Review, 88*, 354–364. doi:10.1037/0033-295X.88.4.354.

Bem, S. L. (Ed.). (1984). *Androgyny and gender schema theory: A conceptual and empirical integration*. Lincoln: University of Nebraska Press.

Bem, S. L., & Lenney, E. (1976). Sex typing and the avoidance of cross-sex behavior. *Journal of Personality and Social Psychology, 33*, 48–54. doi:10.1037/h0078640.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *The Behavioral and Brain Sciences, 11*, 169–232. doi:10.1017/S0140525X00049670.

Bernard, M. E., Boyle, G. J., & Jackling, B. F. (1990). Sex-role identity and mental ability. *Personality and Individual Differences, 11*, 213–217. doi:10.1016/0191-8869(90)90234-I.

Borenstein, M., & Rothstein, H. R. (1999). *Comprehensive meta-analysis: A computer program for research synthesis*. Englewood: BioStat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: John Wiley & Sons, Ltd.

Boyle, G. J., Neumann, D. L., Furedy, J. J., & Westbury, H. R. (2010a). Combining the methods of differential and experimental psychology to study sex differences in human cognitive psychological functions. *Perceptual and Motor Skills, 110*, 392–410. doi:10.2466/pms.110.2.396-410.

Boyle, G. J., Neumann, D. L., Furedy, J. J., & Westbury, H. R. (2010b). Sex differences in verbal and visual-spatial tasks under different hemispheric visual-field presentation conditions. *Perceptual and Motor Skills, 110*, 396–410. doi:10.2466/pms.110.2.396-410.

Brosnan, M. J. (1998). The implications for academic attainment of perceived gender-appropriateness upon spatial task performance. *British Journal of Educational Psychology, 68*, 203–215. doi:10.1111/j.2044-8279.1998.tb01284.x.

Caplan, P. J., & Caplan, J. B. (1994). *Thinking critically about research on sex and gender*. New York: Harper Collins.

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

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*, 697–705. doi:10.1037/0012-1649.31.4.697.

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33*, 669–680. doi:10.1037/0012-1649.33.4.669.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*, 218–261. doi:10.1037/a0014412.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Earlbaum Associates.

Connor, J. M., & Serbin, L. A. (1977). Behaviorally based masculine-and feminine-activity-preference scales for preschoolers: Correlates with other classroom behaviors and cognitive tests. *Child Development, 48*, 1411–1416. doi:10.1111/j.1467-8624.1977.tb03947.x.

Constantinople, A. (1973). Masculinity-femininity: An exception to a famous dictum? *Psychological Bulletin, 80*, 389–409. doi:10.1037/h0035334.

Cooper, H. M. (1981). On the significance of effects and the effects of significance. *Journal of Personality and Social Psychology, 41*, 1013–1018. doi:10.1037/0022-3514.41.5.1013.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75–176). Oxford: Academic.

Crawford, M., Chaffin, R., & Fitton, L. (1995). Cognition in social context. *Learning and Individual Differences, 7*, 341–362. doi:10.1016/1041-6080(95)90006-3.

Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence, 32*, 25–32. doi:10.1016/S0160-2896(03)00061-8.

Doyle, R. A., Voyer, D., & Cherney, I. D. (2012). The relation between childhood spatial activities and spatial abilities in adulthood. *Journal of Applied Developmental Psychology, 33*, 112–120. doi:10.1016/j.appdev.2012.01.002.

Dwyer, C. A. (1973). Sex differences in reading: An evaluation and a critique of current theories. *Review of Educational Research, 43*, 455–467. doi:10.2307/1170076.

Dwyer, C. A. (1974). Influence of children's sex role standards on reading and arithmetic achievement. *Journal of Educational Psychology, 66*, 811–816. doi:10.1037/h0021522.

Eagly, A. H., & Mitchell, A. A. (2004). Social role theory of sex differences and similarities: Implications for the sociopolitical attitudes of men and women. In M. A. Paludi (Ed.), *Praeger guide to the psychology of gender* (pp. 183–206). Westport: Praeger Publishers.

Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist, 54*, 408–423. doi:10.1037/0003-066X.54.6.408.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135–172. doi:10.1111/j.1471-6402.1987.tb00781.x.

Eccles, J. S. (2007). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci (Ed.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 199–210). Washington, D.C.: American Psychological Association.

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46*, 183–201. doi:10.1111/j.1540-4560.1990.tb01929.x.

Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*, 830–847. doi:10.2307/1131221.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103–127. doi:10.1037/a0018053.

Emmott, S. (1985). Sex differences in children's play: Implications for cognition. *International Journal of Women's Studies, 8*, 449–456.

Evardone, M., & Alexander, G. (2009). Anxiety, sex-linked behaviors, and digit ratios (2D:4D). *Archives of Sexual Behavior, 38*, 442–455. doi:10.1007/s10508-007-9260-6.

Eysenck, H. J., Wilson, G. D., & Jackson, C. (1996). *Eysenck Personality Profiler*. Surrey: PSI Press.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*, 95–103. doi:10.1037/0003-066X.43.2.95.

Fennema, E., & Sherman, J. A. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal, 14*, 51–71. doi:10.3102/00028312014001051.

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed-and random-effects methods. *Psychological Methods, 6*, 161–180. doi:10.1037//1082-989X.6.2.161.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Manual for kit of reference tests for cognitive factors*. Princeton: Educational Testing Service.

Frieze, I. H., & Chrisler, J. C. (2011). Editorial policy on the use of the terms "Sex" and "Gender". *Sex Roles, 64*, 789–790. doi:10.1007/s11199-011-9988-2.

Gallagher, A. M., & Kaufman, J. C. (Eds.). (2005). *Gender differences in mathematics*. New York: Cambridge University Press.

Gilger, J. W., & Ho, H.-Z. (1989). Gender differences in adult spatial information processing: Their relationship to pubertal timing, adolescent activities, and sex-typing of personality. *Cognitive Development, 4*, 197–214. doi:10.1016/0885-2014(89)90015-4.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science, 320*, 1164–1165. doi:10.1126/science.1154094.

Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah: Erlbaum.

Halpern, D. F. (2007). Science, sex, and good sense: Why women are underrepresented in some areas of science and math. In S. J. Ceci (Ed.), *Why aren't more women in science?: Top researchers debate the evidence* (pp. 121–130). Washington, D.C.: American Psychological Association.

Halpern, D. F. (2011). *Sex differences in cognitive abilities* (4th ed.). Mahwah: Erlbaum.

Halpern, D. F., & Lamay, M. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review, 12*, 229–246. doi:10.1023/A:1009027516424.

Halpern, D. F., & Tan, U. (2001). Stereotypes and steroids: Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cognition, 45*, 392–414. doi:10.1006/brcg.2001.1287.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–51. doi:10.1111/j.1529-1006.2007.00032.x.

Halpern, D. F., Beninger, A. S., & Straight, C. A. (2011). Sex differences in intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 253–272). Cambridge: Cambridge University Press.

Halpern, D. F., Straight, C. A., & Stephenson, C. L. (2011). Beliefs about cognitive gender differences: Accurate for direction, underestimated for size. *Sex Roles, 64*, 336–347. doi:10.1007/s11199-010-9891-2.

Hamilton, C. J. (1995). Beyond sex differences in visuo-spatial processing: The impact of gender trait possession. *British Journal of Psychology, 86*, 1–20. doi:10.1111/j.2044-8295.1995.tb02542.x.

Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14–50). Baltimore: Johns Hopkins University Press.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*, 41–45. doi:10.1126/science.7604277.

Hoffman, L. W. (1972). Early childhood experiences and women's achievement motives. *Journal of Social Issues, 28*, 129–155. doi:10.1111/j.1540-4560.1972.tb00022.x.

Hromatko, I., Tadinac, M., & Vranic, A. (2008). Femininity and masculinity across the menstrual cycle: A relation to mate value. *Collegium Antropologicum, 32*, 81–86.

Hyde, J. S. (1990). Meta-analysis and the psychology of gender differences. *Signs, 16*, 55–73. doi:10.1086/494645.

Hyde, J. S. (1996). Where are the gender differences? Where are the gender similarities? In D. M. Buss & N. M. Malmouth (Eds.), *Sex, power and conflict* (pp. 107–118). New York: Oxford University Press.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581–592. doi:10.1037/0003-066X.60.6.581.

Hyde, J. S. (2006). Gender similarities still rule. *American Psychologist, 61*, 641–642. doi:10.1037/0003-066X.61.6.641b.

Hyde, J. S. (2007a). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science, 16*, 259–263. doi:10.1111/j.1467-8721.2007.00516.x.

Hyde, J. S. (2007b). Women in science: Gender similarities in abilities and sociocultural forces. In S. J. Ceci (Ed.), *Why aren't more women in science?: Top researchers debate the evidence* (pp. 131–145). Washington, D.C.: American Psychological Association.

Hyde, J. S., & Lindberg, S. M. (2007). Facts and assumptions about the nature of gender differences and the implications for gender equity. In S. S. Klein (Ed.), *Handbook for achieving gender equity through education* (2nd ed., pp. 19–32). Mahwah: Lawrence Erlbaum Associates.

Jagieka, A., & Herman-Jeglińska, A. (1998). The effect of masculinity and sex hormones on a mental rotation task in right-and left-handed males. *International Journal of Psychophysiology, 30*, 111–112. doi:10.1016/S0167-8760(98)90282-3.

Jamison, W., & Signorella, M. L. (1980). Sex-typing and spatial ability: The association between masculinity and success on Piaget's water-level task. *Sex Roles, 6*, 345–353. doi:10.1007/BF00287356.

Jamison, W., & Signorella, M. L. (1987). Relations of masculinity and femininity in self-concept to spatial performance in adolescents. *The Journal of Genetic Psychology, 148*, 249–251. doi:10.1080/00221325.1987.9914554.

Kagan, J. (1964a). Acquisition and significance of sex-typing and sex-role identity. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research* (Vol. 1, pp. 137–168). New York: Russell Sage.

Kagan, J. (1964b). The child's sex role classification of school objects. *Child Development, 35*, 1051–1056. doi:10.2307/1126852.

Kalichman, S. C. (1989). Sex roles and sex differences in adult spatial performance. *The Journal of Genetic Psychology, 150*, 93–100. doi:10.1080/00221325.1989.9914579.

Kelley, G., & Kelley, K. (2012). Statistical models for meta-analysis: A brief tutorial. *World Journal of Methodology, 2*, 27–32. doi:10.5662/wjm.v2.i4.27.

Kimura, D. (2000). *Sex and cognition*. Cambridge: MIT Press.
Knafo, A., Iervolino, A. C., & Plomin, R. (2005). Masculine girls and feminine boys: Genetic and environmental contributions to atypical gender development in early childhood. *Journal of Personality and Social Psychology, 88*, 400–412. doi:10.1037/0022-3514.88.2.400.
Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby & R. G. D'Andrade (Eds.), *The development of sex differences* (pp. 82–172). Stanford: Stanford University Press.
Kohlberg, L., & Ullian, D. (1974). Stages in the development of psycho-sexual concepts and attitudes. In R. C. Friedman, R. M. Richart, & R. Vande Wiele (Eds.), *Sex differences in behavior* (pp. 209–222). New York: Wiley.
Lane, K., Goh, J., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles, 66*, 220–234. doi:10.1007/s11199-011-0036-z.
Lau, S. (1989). Sex role orientation and domains of self-esteem. *Sex Roles, 21*, 415–422. doi:10.1007/BF00289600.
Lever, J. (1976). Sex differences in the games children play. *Social Problems, 23*, 478–487. doi:10.1525/sp.1976.23.4.03a00100.
Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16*, 841–845. doi:10.1111/j.1467-9280.2005.01623.x.
Lewis, S., & Clarke, M. (2001). Forest plots: Trying to see the wood and the trees. *BMJ, 322*, 1479–1480. doi:10.1136/bmj.322.7300.1479.
Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research, 41*, 63–76. doi:10.1080/0013188990410106.
Liben, L. S., Bigler, R. S., & Krogh, H. R. (2002). Language at work: Children's gendered interpretations of occupational titles. *Child Development, 73*, 810–828. doi:10.1111/1467-8624.00440.
Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*, 1123–1135. doi:10.1037/a0021276.
Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479–1498. doi:10.2307/1130467.
Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin, 109*, 267–296. doi:10.1037/0033-2909.109.2.267.
Maccoby, E. E. (1990). The role of gender identity and gender constancy in sex-differentiated development. *New Directions for Child and Adolescent Development, 1990*, 5–20. doi:10.1002/cd.23219904703.
Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford: Stanford University Press.
Martin, C. L., & Ruble, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science, 13*, 67–70. doi:10.1111/j.0963-7214.2004.00276.x.
Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. *Science Education, 83*, 33–54. doi:10.1002/(SICI)1098-237X(199901)83:1<33::AID-SCE2>3.0.CO;2-Z.
Matthews, J. (2007). Hidden sexism: Facial prominence and its connections to gender and occupational status in popular print media. *Sex Roles, 57*, 515–525. doi:10.1007/s11199-007-9276-3.
McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173–180. doi:10.1111/1467-8624.00131.
Milton, G. A. (1957). The effects of sex-role identification upon problem-solving skill. *Journal of Abnormal and Social Psychology, 55*, 208–212. doi:10.1037/h0044103.
Nash, S. C. (1975). The relationship among sex-role stereotyping, sex-role preference, and the sex difference in spatial visualization. *Sex Roles, 1*, 15–32. doi:10.1007/BF00287210.
Nash, S. C. (1979). Sex role as mediator of intellectual functioning. In M. A. Wittig & A. C. Petersen (Eds.), *Sex-related differences in cognitive functioning: Developmental issues* (pp. 263–302). New York: Academic.
National Science Foundation. (2011). *Women, minorities, and persons with disabilities in science and engineering: 2011*. Arlington: National Science Foundation. Retrieved from http://www.nsf.gov/statistics/wmpd/pdf/wmpd2011.pdf.
Neisser, U. C., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77–101. doi:10.1037/0003-066X.51.2.77.
Neumann, D. L., Fitzgerald, Z. T., Furedy, J. J., & Boyle, G. J. (2007). Sexually dimorphic effects of acute nicotine administration on arousal and visual-spatial ability in non-smoking human volunteers. *Pharmacology Biochemistry and Behavior, 86*, 758–765. doi:10.1016/j.pbb.2007.03.001.
Neumann, D. L., Sturm, A. C., Boyle, G. J., & Furedy, J. J. (2010). Effects of nicotine administration via a sublingual tablet on arousal and verbal ability in non-smokers. *Australian Journal of Psychology, 62*, 75–81. doi:10.1080/00049530902795458.
Newcombe, N. S. (2007). Taking science seriously: Straight thinking about spatial sex differences. In S. J. Ceci (Ed.), *Why aren't more women in science?: top researchers debate the evidence* (pp. 69–77). Washington, D.C.: American Psychological Association.
Newcombe, N. S., & Dubas, J. S. (1992). A longitudinal study of predictors of spatial ability in adolescent females. *Child Development, 63*, 37–46. doi:10.1111/j.1467-8624.1992.tb03593.x.
Newcombe, N. S., & Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. *Mind, Brain, and Education, 4*, 102–111. doi:10.1111/j.1751-228X.2010.01089.x.
Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of Personality and Social Psychology, 83*, 44–59. doi:10.1037//0022-3514.83.1.44.
Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*, 10593–10597. doi:10.1073/pnas.0809921106.
Nuttall, R. L., Casey, M. B., & Pezaris, E. (2005). Spatial ability as a mediator of gender differences on mathematics tests: A biological-environmental framework. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 121–142). Cambridge: Cambridge University Press.
Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8*, 157–159.
Oswald, D. L. (2008). Gender stereotypes and women's reports of liking and ability in traditionally masculine and feminine occupations. *Psychology of Women Quarterly, 32*, 196–203. doi:10.1111/j.1471-6402.2008.00424.x.
Pajares, F., & Miller, D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193–203. doi:10.1037/0022-0663.86.2.193.
Peters, M., Lehmann, W., Takahira, S., Takeuchi, Y., & Jordan, K. (2006). Mental rotation test performance in four cross-cultural samples (N= 3367): Overall sex differences and the role of academic program in performance. *Cortex, 42*, 1005–1014. doi:10.1016/S0010-9452(08)70206-5.
Piaget, J. (1968). *Six psychological studies*. New York: Vintage Books.
Popiel, E. M., & De Lisi, R. (1984). An examination of spatial ability in relation to factors from the Bem Sex-Role Inventory. *Perceptual and Motor Skills, 59*, 131–136.

Priess, H. A., & Hyde, J. S. (2010). Gender and academic abilities and preferences. In J. C. Chrisler & D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 297–316). New York: Springer.

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Biosocial factors, sexual orientation and neurocognitive functioning. *Psychoneuroendocrinology, 29*, 867–881. doi:10.1016/S0306-4530(03)00154-9.

Reilly, D. (2012). Gender, culture and sex-typed cognitive abilities. *PLoS One, 7*, e39904. doi:10.1371/journal.pone.0039904.

Ritter, D. (2004). Gender role orientation and performance on stereotypically feminine and masculine cognitive tasks. *Sex Roles, 50*, 583–591. doi:10.1023/B:SERS.0000023077.91248.f7.

Robison-Awana, P., Kehle, T. J., & Jenson, W. R. (1986). But what about smart girls? Adolescent self-esteem and sex role perceptions as a function of academic achievement. *Journal of Educational Psychology, 78*, 179–183. doi:10.1037//0022-0663.78.3.179.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*, 638–641. doi:10.1037/0033-2909.86.3.638.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills: Sage.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192. doi:10.1037/0033-2909.118.2.183.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*, 59–82. doi:10.1146/annurev.psych.52.1.59.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166–169. doi:10.1037/0022-0663.74.2.166.

Rosenthal, L., London, B., Levy, S., & Lobel, M. (2011). The roles of perceived identity compatibility and social support for women in a single-sex STEM program at a co-educational university. *Sex Roles, 65*, 725–736. doi:10.1007/s11199-011-9945-0.

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology (6th ed., Vol. 3. Social, emotional, and personality development, pp. 858–932)*. Hoboken: Wiley.

Saucier, D. M., McCreary, D. R., & Saxberg, J. K. J. (2002). Does gender role socialization mediate sex differences in mental rotations? *Personality and Individual Differences, 32*, 1101–1111. doi:10.1016/S0191-8869(01)00106-4.

Scarbrough, P. S., & Johnston, V. S. (2005). Individual differences in women's facial preferences as a function of digit ratio and mental rotation ability. *Evolution and Human Behavior, 26*, 509–526. doi:10.1016/j.evolhumbehav.2005.03.002.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194–201. doi:10.1006/jesp.2001.1500.

Serbin, L. A., & Connor, J. M. (1979). Sex-typing of children's play preferences and patterns of cognitive performance. *Journal of Genetic Psychology, 134*, 315–316. doi:10.1016/0022-0965(79)90050-X.

Serbin, L. A., Zelkowitz, P., Doyle, A. B., Gold, D., & Wheaton, B. (1990). The socialization of sex-differentiated skills and academic performance: A mediational model. *Sex Roles, 23*, 613–628. doi:10.1007/BF00289251.

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. doi:10.1037//0022-0663.93.3.604.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701–703. doi:10.1126/science.171.3972.701.

Sherman, J. A. (1967). Problem of sex differences in space perception and aspects of intellectual functioning. *Psychological Review, 74*, 290–299. doi:10.1037/h0024723.

Signorella, M. L., & Jamison, W. (1978). Sex differences in the correlations among field dependence, spatial ability, sex role orientation, and performance on Piaget's water-level task. *Developmental Psychology, 14*, 689–690.

Signorella, M. L., & Jamison, W. (1986). Masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin, 100*, 207–228. doi:10.1037/0033-2909.100.2.207.

Signorella, M. L., Jamison, W., & Krupa, M. H. (1989). Predicting spatial performance from gender stereotyping in activity preferences and in self-concept. *Developmental Psychology, 25*, 89–95. doi:10.1037/0012-1649.25.1.89.

Silverman, I., Choi, J., & Peters, M. (2007). The Hunter-Gatherer theory of sex differences in spatial abilities: Data from 40 countries. *Archives of Sexual Behavior, 36*, 261–268. doi:10.1007/s10508-006-9168-6.

Spence, J. T. (1984). Gender identity and its implications for masculinity and femininity. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motiuation: Psychology and Gender* (pp. 59–95). Lincoln: University of Nebraska Press.

Spence, J. T. (1993). Gender-related traits and gender ideology: Evidence for a multifactorial theory. *Journal of Personality and Social Psychology, 64*, 624–635. doi:10.1037/0022-3514.64.4.624.

Spence, J. T., & Buckner, C. (2000). Instrumental and expressive traits, trait stereotypes, and sexist attitudes: What do they signify? *Psychology of Women Quarterly, 24*, 44–62. doi:10.1111/j.1471-6402.2000.tb01021.x.

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). The Personal Attributes Questionnaire: A measure of sex role stereotypes and masculinity-femininity. *JSAS Catalog of Selected Documents in Psychology, 4*, 127.

Spence, J. T., Helmreich, R. L., & Stapp, J. (1975). Ratings of self and peers on sex-role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology, 32*, 29–39. doi:10.1037/h0076857.

Steffens, M., & Jelenec, P. (2011). Separating implicit gender stereotypes regarding math and language: Implicit ability stereotypes are self-serving for boys and men, but not for girls and women. *Sex Roles, 64*, 324–335. doi:10.1007/s11199-010-9924-x.

Stein, A. H., & Bailey, M. M. (1973). The socialization of achievement orientation in females. *Psychological Bulletin, 80*, 345–366. doi:10.1037/h0035261.

Szymanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences, 21*, 493–504. doi:10.1016/j.lindif.2011.07.001.

Taylor, M. C., & Hall, J. A. (1982). Psychological androgyny: Theories, methods, and conclusions. *Psychological Bulletin, 92*, 347–366. doi:10.1037/0033-2909.92.2.347.

Thurstone, T. G. (1958). *SRA Primary Mental Abilities (test)*. Chicago: Science Research Associates.

Titze, C., Jansen, P., & Heil, M. (2010). Mental rotation performance in fourth graders: No effects of gender beliefs (yet?). *Learning and Individual Differences, 20*, 459–463. doi:10.1016/j.lindif.2010.04.003.

Tracy, D. M. (1987). Toy-playing habits, spatial ability, science and mathematics achievement: Are they related? *Sex Roles, 17*, 115–138. doi:10.1007/BF00287620.

Tuttle, G. E., & Pillard, R. C. (1991). Sexual orientation and cognitive abilities. *Archives of Sexual Behavior, 20*, 307–318. doi:10.1007/BF01541849.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604. doi:10.2466/pms.1978.47.2.599.

Vonnahme, P. (2005). *Facial preferences of Caucasian males as a function of 2D:4D digit ratio and mental rotation ability (Unpublished dissertation)*. New Mexico State University, Las Crues, New Mexico.

Voyer, D., & Bryden, M. (1990). Mental rotation: A study of gender differences and sex role identity. *Canadian Psychology, 31*, 253.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270. doi:10.1037//0033-2909.117.2.250.

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*, 817–835. doi:10.1037/a0016127.

Whitley, B. E. (1983). Sex role orientation and self-esteem: A critical meta-analytic review. *Journal of Personality and Social Psychology, 44*, 765–778. doi:10.1037/0022-3514.44.4.765.

Whitley, B. E. (1988). Masculinity, femininity, and self-esteem: A multitrait-multimethod analysis. *Sex Roles, 18*, 419–431. doi:10.1007/bf00288393.

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604. doi:10.1037/0003-066X.54.8.594.

Williams, W. M., & Ceci, S. J. (2007). Introduction: Striving for perspective in the debate on women and science. In S. J. Ceci (Ed.), *Why aren't more women in science?: Top researchers debate the evidence* (pp. 3–23). Washington, D.C.: American Psychological Association.