CHILD DEVELOPMENT

Child Development, November/December 2018, Volume 89, Number 6, Pages 1943–1955

The title for this Special Section is **Meta-analysis and Individual Participant Data Synthesis in Child Development**, edited by Glenn I. Roisman and Marinus H. van IJzendoorn

The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies

David I. Miller, Kyle M. Nolla, Alice H. Eagly, and David H. Uttal
Northwestern University

This meta-analysis, spanning 5 decades of Draw-A-Scientist studies, examined U.S. children's gender-science stereotypes linking science with men. These stereotypes should have weakened over time because women's representation in science has risen substantially in the United States, and mass media increasingly depict female scientists. Based on 78 studies (*N* = 20,860; grades K-12), children's drawings of scientists depicted female scientists more often in later decades, but less often among older children. Children's depictions of scientists therefore have become more gender diverse over time, but children still associate science with men as they grow older. These results may reflect that children observe more male than female scientists in their environments, even though women's representation in science has increased over time.

Adults in many nations associate science with men much more than with women (e.g., Miller, Eagly, & Linn, 2015; Smyth & Nosek, 2015). To investigate the origins of these associations, researchers have studied children's perceptions of scientists over several decades. For instance, children were asked to draw a scientist in a landmark study of nearly 5,000 elementary school students who were mostly from the United States and Canada (Chambers, 1983). The drawings, collected from 1966 to 1977, almost exclusively depicted male scientists, often with lab coats, eyeglasses, and facial hair, working indoors with laboratory equipment. Only 28 children drew a female scientist (0.6% of the sample), suggesting strong gender-science stereotypes linking science with men. This limited view of scientists might have restricted children's science-related educational and career aspirations, to the extent that children did not identify with such depictions.

Since Chambers (1983) collected data in the 1960s and 1970s, however, women's representation in science has risen substantially in the United States. For instance, women earned 19% of U.S. chemistry bachelor's degrees in 1966 but 48% of such degrees in 2015 (National Science Foundation, 2017). Female scientists are also now more often depicted in popular children's television shows (Long et al., 2010), science textbooks (Pienta & Smith, 2012), magazines such as *Highlights for Children* (Previs, 2016), and other mass media products (Steinke, 2013).

We studied how children's gender-science stereotypes have changed over 5 decades by meta-analyzing the expansive literature of U.S. Draw-A-Scientist studies. Individual studies are often uninformative in studying cultural change because they typically include only one point in time or one cohort. Meta-analytic methods, however, can overcome this limitation by comparing children across multiple decades. Moreover, this drawing task has been widely administered in grade levels varying from early elementary school to late high school, allowing us to study both developmental and historical change in the same meta-analysis.

David I. Miller was supported by a National Science Foundation Graduate Research Fellowship (DGE-0824162).

Correspondence concerning this article should be addressed to David I. Miller, Department of Psychology, Northwestern University, Swift Hall, 2029 Sheridan Road, Evanston, IL 60208. Electronic mail may be sent to dmiller@u.northwestern.edu.

DOI: 10.1111/cdev.13039

Gender-Science Stereotypes Across Development and Historical Time

Development in Childhood

Theories from developmental psychology (e.g., Bigler & Liben, 2007; Cvencek & Meltzoff, 2015; Martin & Ruble, 2004) and social psychology (e.g., Eagly & Wood, 2012; Fiske, Cuddy, Glick, & Xu, 2002; Wood & Eagly, 2012) provide a framework for understanding how gender-science stereotypes develop. In general, stereotypes about social groups form based on people's direct observations of group members, such as through social interactions, and indirect observations, such as through mass media (Bigler & Liben, 2007; Koenig & Eagly, 2014). Because gender is a particularly salient social identity, children actively search their environment for cues to what activities are considered appropriate for boys and girls (Arthur, Bigler, Liben, Gelman, & Ruble, 2008; Martin & Ruble, 2004).

Although some gender stereotypes start to form as early as toddlerhood, the stereotype of scientists as male should emerge somewhat later because preschoolers likely have limited knowledge of scientists (Lee, 2010; Newton & Newton, 1992, figures 1, 2). Formal science instruction is sporadic in early elementary school and often does not begin systematically until late elementary school to middle school (Smith, Banilower, McMahon, & Weiss, 2002). During formal science instruction, children could form gender stereotypes about science and mathematics partly based on teachers' attitudes and beliefs, as some correlational studies have suggested (Keller, 2001; for a review, see Gunderson, Ramirez, Levine, & Beilock, 2012). As children grow older, their opportunities for learning stereotypes about scientists also expand beyond school. For instance, middle school children have reported basing their drawings of scientists on media sources such as television shows, movies, books, and magazines (Fort & Varney, 1989; Steinke et al., 2007). These mass media products have usually depicted many more male than female scientists, at least in prior decades (for a review, see Steinke, 2013). Children may then learn to associate science with men as their exposure to male scientists accumulates during development. We therefore predicted that older children in our meta-analysis would draw male scientists more often than younger children.

Change Over Time

Even in recent decades, children would likely learn to associate science with men because women remain underrepresented in many science and engineering fields (Ceci, Ginther, Kahn, & Williams, 2014). However, women's representation in science has also increased substantially in the United States over the past 5 decades (Miller & Wai, 2015). For instance, from 1960 to 2013, the percentage of women among employed U.S. scientists rose from 28% to 49% in biological science, 8% to 35% in chemistry, and 3% to 11% in physics and astronomy (Hill, Corbett, & St. Rose, 2010, figure 11; National Science Board, 2016, appendix table 3–12). Furthermore, female scientists are now depicted more often in children's media. For instance, women and girls were 13% of images of people in science feature stories in the popular magazine *Highlights for Children* in the 1960s, but this percentage rose to 44% by the 2000s (Previs, 2016). Likewise, women and girls were 39% of images of scientists in children's nonfiction trade books published in 1991–2011 (Rawson & McCool, 2014), 44% of images of scientists in middle school science textbooks published in 2008 (Pienta & Smith, 2012), and 42% of scientist characters in television programs popular among middle school-age children in 2006 (Long et al., 2010).

Given this increased representation of female scientists, children might learn to associate science with both men and women. Exposure to female scientists might also weaken already-formed associations of science with men (Farland, 2006; Galdi, Cadinu, & Tomasetto, 2014; Gonzalez, Dunlop, & Baron, 2017), especially if students identify with the female role models (Young, Rudman, Buettner, & McLean, 2013). Consistent with these hypotheses, women's representation among science majors and employed researchers predicted weaker national gender-science stereotypes across 66 nations in one large study (Miller et al., 2015). Hence, children's stereotypes should partly reflect changes in women's representation in science over time. We therefore predicted that U.S. children would draw male scientists less often in later than earlier decades. Our meta-analysis focused on U.S. samples to test this hypothesis because no other nation had enough relevant Draw-A-Scientist studies of children to allow examination of cultural change over several decades.

The Present Meta-analysis

In summary, our main hypotheses were that stereotypes linking science with men would strengthen with children's age, but weaken over time in the United States. The wide array of Draw-

A-Scientist studies, which include data dating back to the 1960s, allowed us to investigate classic questions in developmental science about age and period effects. This literature also provided a unique opportunity to compare studies on a common metric (i.e., percentage of male scientists) because all studies administered the same task. In contrast, other meta-analyses typically include studies with many different measures that might not always be meaningful to compare in analyses.

Several features of the Draw-A-Scientist Test should be noted from the outset (see Finson, 2002 for a review). First, like other indirect measures of children's attitudes and beliefs, the Draw-A-Scientist Test did not require children to consciously report their stereotypes with an introspective verbal response (Cvencek & Meltzoff, 2015). The indirect nature of this task can be advantageous in developmental research because children may first learn stereotypes implicitly before reporting them explicitly (Galdi et al., 2014).

Second, as prior research suggests, children typically draw more male than female scientists because they associate science with men and not because they generally draw males regardless of occupation. For instance, one study asked 206 elementary school children to draw a scientist, a veterinarian, and a teacher (Losh, Wilke, & Pop, 2008). Among drawings in which sex could be determined, 66% of scientists were male, compared to 40% of veterinarians and 25% of teachers. In other words, children drew more males than females only for scientists, but not for the two other occupations in which women's workforce representation was higher. These results suggest that these drawings at least in part reflect children's associations between sex and occupations.

Finally, these drawings may also partly reflect aspects of children's own personal identities. For instance, tasks asking children to draw a generic person have often been interpreted as projective measures of children's self-expression and gender identity (Arteche, Bandeira, & Hutz, 2010). Consistent with this hypothesis, when asked to draw a person, usually 70% or more of both boys and girls have drawn their own sex in prior studies (Arteche et al., 2010; Picard, 2015). Given these results, we predicted that boys in our meta-analysis would draw male scientists more often than girls. This mean difference between boys and girls, however, was not central to our main research goals. Instead, our analyses that considered children's sex focused on investigating if change over age and historical time was present for both boys and girls.

Method

Literature Search and Inclusion Criteria

We searched for the phrases "draw a scientist" OR "draw-a-scientist" in 15 databases of published literature (e.g., ERIC, PsycINFO, Web of Science); see Table S1 for the list of databases. We also systematically searched for unpublished studies using ProQuest Dissertations & Theses Global and Google Scholar. Google Scholar was especially helpful because it searches many sources such as university websites for unpublished studies (see https://scholar.google.com/intl/en/scholar/inclusion.html). ProQuest yielded 317 dissertations and theses and, using the full text search option, Google Scholar yielded 1,890 results. We examined all articles that these 17 databases had indexed by the time we completed our latest keyword searches in March 2017. Using five citation databases (Academic Search Complete, Google Scholar, PsycINFO, Scopus, and Web of Science), we also examined all articles that cited at least one of the following: (a) Chambers' (1983) landmark study, (b) a widely cited narrative review of Draw-A-Scientist studies (Finson, 2002), or (c) at least one of the other articles that ultimately became part of our sample. Search results were imported into EndNote X8, and duplicates were deleted, yielding 3,042 unique references.

From these references, we included studies that met all the following criteria: the study (a) reported a sample of U.S. K-12 children, (b) asked these children to draw a scientist, (c) reported sufficient information to record the number of male and female scientists drawn, and (d) had a sample size of at least 10 drawings. We aimed to study naturalistic change over historical time and therefore wanted to estimate children's baseline stereotypes in the absence of study-specific educational or experimental interventions. Consequently, for pre-post and other longitudinal designs, studies were included only if they reported data separately for the first occasion in which children drew scientists. For posttest-only experimental or quasi-experimental designs, studies were included only if they reported data separately for the control condition or found no significant differences across conditions if data were reported aggregated across conditions.

Applying these inclusion and exclusion criteria identified 93 eligible articles (see Appendix S1 for details on the process for determining eligibility and Appendix S2 for the list of articles). For a sample reported by multiple sources (e.g., in a dissertation and later journal article), we examined all articles for relevant information but created only one row of data for that sample, yielding 79 independent samples. Almost half of these studies (47%) were unpublished, indicating that our search methods were effective at finding large amounts of relevant gray literature (see Table S2 for the sources of unpublished and published studies). Appendix S3 details the coding system that we used to record moderators such as data collection year and average age.

Meta-analytic Procedures

Effect Size Calculations

The main outcome variable was the percentage of male scientists among sex-identifiable drawings. When applicable, we recorded the percentage of all drawings depicting scientists of both sexes or indeterminate sexes (e.g., the drawing only contained stick figures with no sex cues such as long hair) and included that percentage as a moderator in multivariable analyses. We converted percentages to log odds for statistical analysis. This conversion presented challenges for samples (often boys) that drew only male scientists because the logs odds would be infinite. In those cases, we added 0.5 to both cells (female and male counts), consistent with standard practice regarding continuity corrections for log odds (Viechtbauer, 2010).

Statistical Models

Our focal analyses used mixed-effects meta-regression models that assumed that observed variation in effect sizes was due to fixed effects of moderators (e.g., year of data collection), random effects of residual between-study heterogeneity, and within-study sampling variance (Borenstein, Hedges, Higgins, & Rothstein, 2009). Mixed-effects and random effects models were chosen over fixed effects models because we wanted to model heterogeneity of effect sizes and make inferences to a larger population of U.S. children (Hedges & Vevea, 1998). Between-study heterogeneity was quantified by presenting 90% credibility intervals (a measure of the estimated dispersion of true underlying effects) and I^2 statistics (percentage of total variation in effect size due to heterogeneity rather than chance). These models were estimated with the *metafor* R package using restricted maximum likelihood estimation and the Knapp–Hartung adjustment to account for uncertainty in estimating heterogeneity (Viechtbauer, 2010). Although inferential statistical models analyzed log odds, we converted model outputs back into more intuitive percentages when presenting descriptive statistics (e.g., regression trendlines).

Publication Bias

Our analyses tested for publication bias in several ways, including comparing average effects estimated by published versus unpublished studies and testing for asymmetry in the funnel plot of observed effect size plotted by sample size (Borenstein et al., 2009). We also tested for change over time in publication bias by including interaction terms with data collection year in relevant meta-regression models (see Appendix S4 for details). In addition, we computed estimates of mean effect size and change over time adjusted for small-study effects (e.g., mean effect size differing as a function of sample size). Publication bias can be one source of small-study effects. Hence, adjusting for small-study effects using meta-regression models provides one way to correct for publication bias (Stanley & Doucouliagos, 2014; though see Simonsohn, 2017 for critique of this assumption; see Appendix S4 for details).

Outliers

Chambers' (1983) original study was an outlier in terms of its data collection years (1966–1977); all other studies collected data in 1985 or after. That study was therefore unique in estimating children's gender-science stereotypes during a period in which women's representation in U.S. science was especially low. However, Chambers' large effect size (99.4% of drawings were male scientists) and sample size (N = 4,807) could have also disproportionately influenced regression analyses. We therefore repeated all analyses both including and excluding Chambers' study. Appendix S5 details how we identified outliers other than Chambers (1983) using various diagnostics (e.g., studentized deleted residuals) developed for outlier detection in random effects and mixed-effects meta-regression models (Viechtbauer & Cheung, 2010). For reasons detailed in Appendix S5, one outlier (Cavallo, 2007) was excluded from all analyses and another outlier (Flick, 1990) was excluded from analyses of boys' drawings.

Results

Our analyses addressed two main research questions: (a) How do gender-science stereotypes vary across age and historical time? (b) Do the predicted effects of age and time remain after accounting for each other and other moderators? Of the 78 analyzed studies, 30 studies reported data separately for boys and girls, and 11 studies included only girls. No study included only boys. We therefore conducted separate meta-analyses for all children's drawings (k = 78 studies; N = 20,860 sex-identifiable drawings), girls' drawings (k = 41; N = 6,730), and boys' drawings (k = 29; N = 5,941). The Supporting Information contain the data files and R code needed to reproduce all analyses including figures and tables. These materials have also been uploaded to the Open Science Framework (https://osf.io/3awvj/).

All Drawings

Distribution of Effect Sizes

Children overall drew 73% of scientists as male, 95% CI [69, 77], averaged across all 78 analyzed samples using random effects weighting; this percentage was 72% excluding Chambers (1983). However, effect sizes also varied widely across studies ($\tau^2 = 0.698$; $p < .0001$). The middle 90% of true underlying effects (i.e., 90% credibility interval) were estimated to fall within 41%–92%. Furthermore, most variability in observed effect sizes could be attributed to between-study heterogeneity rather than chance ($I^2 = 96\%$), suggesting that moderators (e.g., age, year) may help explain differences in observed effect sizes.

Simple Regression Analysis

Children drew male scientists less often in later than earlier decades (see Figure 1a). Children drew 99.4% of scientists as male in Chambers' (1983) study (data collection years 1966–1977) compared to 72% on average in later studies (years 1985–2016). The regression coefficient for data collection year was significant ($b = -.040$; $p = .0004$). However, when Chambers' study was excluded from analysis, the regression coefficient was 51% smaller in magnitude ($b = -.020$) and the p value was .062 (see also the dashed line in Figure 1a). In other words, the clearest evidence of change over time came from analyses that used data from the 1960s and 1970s (i.e., Chambers' study).

In addition, older children drew male scientists more often than younger children (see Figure 1a), both when including Chambers' (1983) study ($b = .098$; $p = .023$) or excluding it ($b = .136$; $p < .0001$). However, as also shown in Figure 1b, the age effect was smaller in magnitude when Chambers' study was included (see solid line) than excluded (see dashed line) because its large effect size was unexpected based on the young average age of children in it (8.26 years). Given Chambers' (1983) outlying data collection years (1966–1977), we considered analyses excluding Chambers to be more appropriate estimates of average age effects in contemporary U.S. data. The mean percentage of male scientists did not significantly differ from 50% until age 8 when excluding Chambers and age 7 when including Chambers.

Multivariable Regression Analysis

Multivariable meta-regression models tested whether the predicted effects of average age and data collection year remained after controlling for each other and three other moderators: (a) percentage of drawings with scientists of indeterminate sex (see Effect Size Calculations in the Methods section for more details), (b) dummy code for publication status (0 = unpublished, 1 = published), and (c) dummy code for an all-female sample (0 = mixed-sex, 1 = all-female). This multivariable analysis helped address potential confounds between predictors (see Table S3 for the correlation matrix). For instance, the regression coefficient for age estimated comparisons between younger and older children, holding constant other moderators such as data collection year (e.g., 8-year-olds in 2010 vs. 14-year-olds in 2010).

When controlling for other moderators in multivariable models, the age effect remained significant both when Chambers' (1983) study was included ($b = .084$; $p = .024$) or excluded ($b = .119$; $p < .0001$). In the same multivariable models, the historical time effect remained significant when Chambers was included ($b = -.037$; $p = .0004$), but not when excluded ($b = -.013$; $p = .124$). Table 1 summarizes these age and time effects and Table S4 provides the complete multivariable models. Compared to a model with no moderators, between-study heterogeneity was reduced from 0.698 to 0.463 in a multivariable model including Chambers ($R^2 = 34\%$) and 0.392 to 0.226 in a multivariable model excluding Chambers ($R^2 = 42\%$).

One concern about cross-sectional age comparisons is the confound with birth cohort (e.g., 8-year-olds in 2010 were born later in time than 14-year-olds in 2010). For instance, younger children might have drawn fewer male scientists than older children in the same data collection year because younger children were born and grew up later in

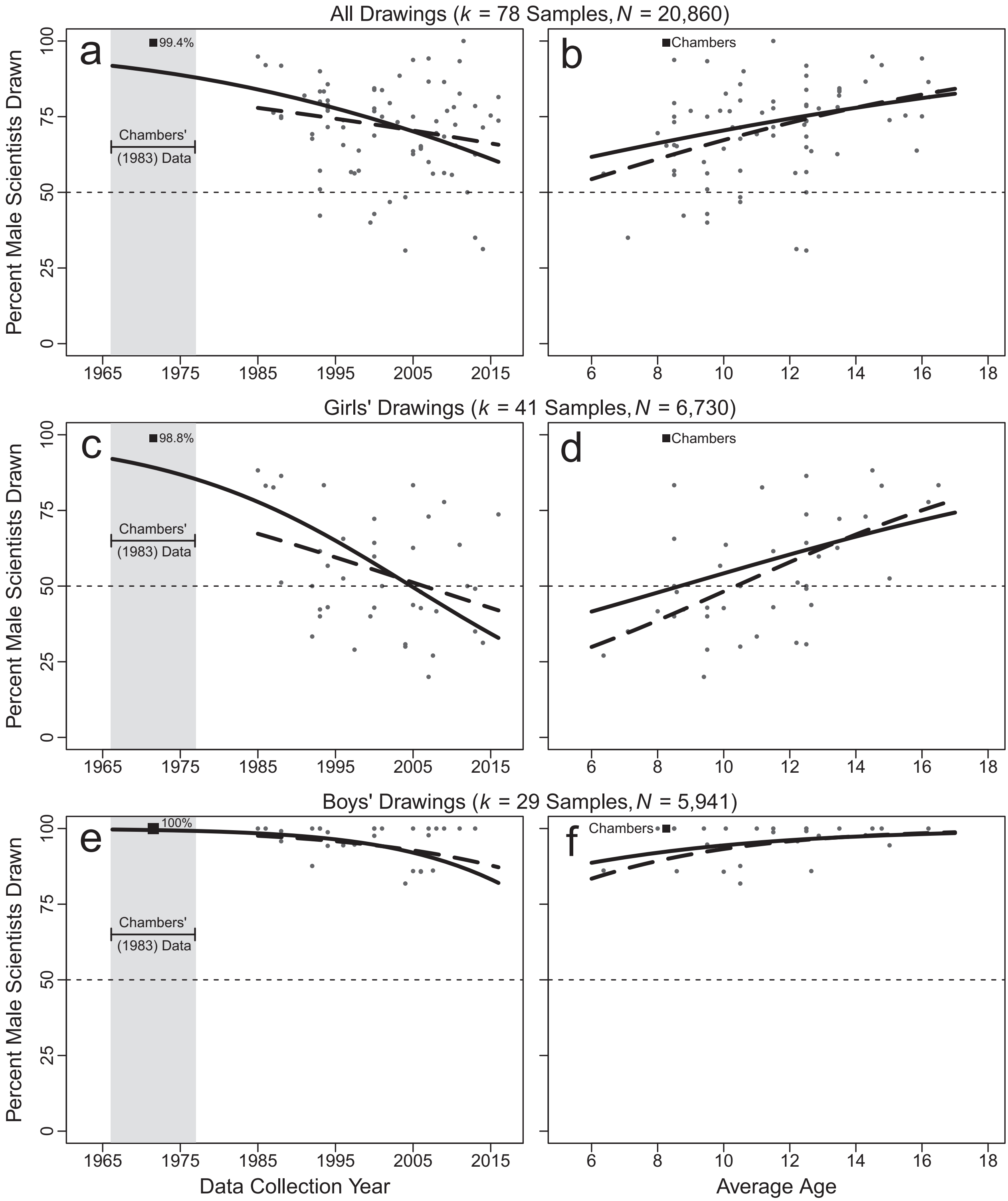

Figure 1. Change over historical time (panels a, c, e) and age (panels b, d, f) in the percentage of scientists drawn as male. The lines represent model predictions converted from log odds to percentages based on including Chambers' (1983) study (solid lines) or excluding it (dashed lines).

historical time. In other words, the estimated effect of age might not represent developmental change but instead a confound with birth cohort. However, this alternative explanation was unlikely because the magnitude of the age effect was much greater than the historical time effect (see Table 1 to

Table 1
Meta-Regressions Estimates of Historical Time and Age Effects

	Including Chambers		Excluding Chambers	
Predictor	*b*	*p*	*b*	*p*
Data collection year				
Full sample (simple)	−.040	< .001	−.020	.062
Full sample (multivariable)	−.037	< .001	−.013	.124
Girls only (simple)	−.063	< .001	−.034	.028
Girls only (multivariable)	−.059	< .001	−.024	.051
Boys only (simple)	−.084	.003	−.058	.035
Boys only (multivariable)	−.081	.004	−.044	.092
Average age				
Full sample (simple)	.098	.023	.136	< .001
Full sample (multivariable)	.084	.024	.119	< .001
Girls only (simple)	.127	.065	.195	< .001
Girls only (multivariable)	.125	.032	.187	< .001
Boys only (simple)	.193	.067	.252	.003
Boys only (multivariable)	.107	.246	.182	.034

Note. The left and right columns display regression coefficients from models including and excluding Chambers (1983) landmark study, respectively. Simple regression models included only one predictor (i.e., data collection year or average age), but multivariable models controlled for data collection year, average age, and other predictors simultaneously (see Table S4 for the complete multivariable models). The unstandardized regression coefficients (*b*s) are the predicted change in log odds of drawing male scientists per unit increase in the predictor variable. The unit for age and data collection year was the same (i.e., 1 year), meaning that their coefficients can be directly compared to contrast the magnitude of age and historical time effects.

compare regression coefficients). For instance, based on data from the 1980s and onwards (i.e., models excluding Chambers, 1983), the mean percentage of male scientists changed from 54% to 82% between ages 6 and 16 (1 decade), but only changed from 74% to 70% between years 1995 and 2005 (also 1 decade); see Figure 1. In other words, change over age happened more rapidly than what historical change would alone predict. These results were therefore consistent with rapid change over children's development in addition to slower change over historical time.

Sex-Specific Analyses

We repeated analyses for girls' drawings (k = 41) and boys' drawings (k = 29) separately. Many studies (k = 37) included both girls and boys but did not report data on the drawn scientists' sex separately by children's sex. Those 37 studies were excluded from the following analyses. In other words, sex-specific analyses used a subset of the full sample of 78 studies. Results could have differed from the prior analyses in part for this reason (e.g., the historical time effect could have been stronger in the subset of studies that reported data separately for boys and girls). The exclusion of those 37 studies also meant the following analyses had reduced statistical power.

Distribution of Effect Sizes

Girls on average drew 58% of scientists as male, 95% CI [50, 66], but boys drew a much higher percentage (96%) of scientists as male, 95% CI [93, 97]. These mean percentages were 55% and 95%, respectively, when excluding Chambers (1983). Boys drew only male scientists in more than half of the samples of boys' drawings (15 of 29 samples). Between-study heterogeneity was also large and significant for both girls (τ^2 = 1.031; p < .0001; I^2 = 96%; 90% credibility interval [21, 88]) and boys (τ^2 = 1.027; p < .0001; I^2 = 80%; 90% credibility interval [80, 99]).

Historical Time and Age Effects

Both girls and boys tended to draw male scientists less often in later than earlier decades (see Figure 1). For instance, girls drew 98.8% of scientists as male in Chambers' (1983) study compared to 55% on average in later studies. The historical time effect was generally significant in both simple and multivariable models for girls and boys (see Table 1). The exceptions to this rule were the multivariable models excluding Chambers for girls (p = .051) and boys (p = .092). The time effects were larger in magnitude when including Chambers (see Table 1).

In addition, both girls and boys tended to draw male scientists more often with increasing age. This age effect was significant in all regression models for girls and boys when Chambers' (1983) study was excluded from analysis (ps < .034). However, when Chambers was included, the age effect was smaller in magnitude and generally not significant (see Table 1). As discussed earlier, we considered analyses based on data from the 1980s and onwards (i.e., models excluding Chambers) to provide more appropriate estimates of average age effects in contemporary U.S. data.

Comparing average percentages on Figure 1 can help interpret the magnitude of these historical time and age effects. For instance, based on a simple meta-regression model excluding Chambers, girls on average drew 30% of scientists as male at age 6 (early elementary school; see Figure 1d). However, girls switched to drawing more male than female scientists between the ages of 10 and 11 (fifth grade; end of elementary school). By age 16 (high school), girls on average drew 75% of scientists as male. In contrast, for boys, the mean percentage of male scientists changed from 83% to 98% between ages 6 and 16, excluding Chambers (see Figure 1f). The magnitude of historical time and age effects did not significantly differ by children's sex (see Appendix S6).

Supplemental Analyses

The Supporting Information details the results for supplemental analyses including those regarding publication bias (see Appendix S4), the within-study effect of age (Appendix S7), and outcomes other than scientists' sex such as the presence of laboratory coats (Appendix S8). We briefly summarize the results of these supplemental analyses here.

Publication Bias

Funnel plot analyses provided no evidence of publication bias or change over time in publication bias (*p*s > .26; see Figure S1 for the funnel plot). Published and unpublished studies also did not significantly differ in mean effect size (*p*s > .50), but the difference between published and unpublished studies changed over time (see Appendix S4 for details). The evidence for change over time in publication bias was therefore inconsistent because the difference between published and unpublished studies changed over time, but funnel plot asymmetry did not.

In addition, we computed estimates of mean effect size and change over time adjusted for small-study effects using meta-regression models that extrapolated to a hypothetical set of studies with infinite sample size (see Appendix S4 for details). The adjusted historical time effect was significant when including Chambers' (1983) study ($b = -.049$; $p = .0008$), but not when excluding it ($b = -.016$; $p = .273$); see Table S5. These results reinforce the earlier conclusion that the evidence for change over time was robust when including Chambers but tentative when excluding Chambers.

Within-Study Effect of Age

Thirteen studies reported data disaggregated by age cohort, allowing us to estimate age effects based on within-study comparisons (i.e., younger vs. older children within the same study). Using hierarchical dependence models (see Appendix S7 for details), we found that that older children drew male scientists more often than younger children, even among children within the same study, demonstrating the robustness of age effects for this literature.

Other Outcomes

Analyses examining outcomes other than the scientist's sex (e.g., presence of laboratory coats) found that older children more often than younger children drew scientists with laboratory coats ($b = .195$; $p = .008$) and eyeglasses ($b = .185$; $p = .0002$). In addition, children in later decades also drew scenes that were indoors or in a laboratory less often than children in earlier decades (see Appendix S8 or Figure S2 for more details). Based on random effects weighting, 50% of drawn scientists had laboratory coats, 38% had eyeglasses or goggles, 78% were indoors or in a laboratory, 18% were middle-aged or older, and 79% were Caucasian on average.

Discussion

This meta-analysis provides the first systematic, quantitative review of studies that have administered the Draw-A-Scientist Test. By combining 78 studies including over 20,000 children, we found changes in children's gender-science stereotypes, as reflected in their drawings of scientists. Consistent with our main hypotheses, the tendency to draw male scientists increased with children's age but decreased over historical time in the United States. In addition, boys drew male scientists more often than girls. These findings provide insight into how children learn to associate science with men and how children respond to changes in their cultural environment such as increases in women's representation in science.

Change Over Children's Age

Results suggested children did not associate science with men until grade school. When children started kindergarten at ages 5–6, they drew roughly equal percentages of male and female scientists, averaged across boys and girls. Children did not draw significantly more male than female scientists until ages 7–8. In contrast, some stereotypes such as those about gender and household chores first emerge much earlier in development including as early as ages 2–3 (Martin & Ruble, 2010). Children may gain knowledge of some occupations such as babysitter and firefighter even in preschool. However, consistent with our findings, children's knowledge of scientists likely develops later as they encounter science in school and in the media.

During elementary and middle school, the tendency to draw male scientists increased rapidly with age. When children started high school at ages 14–15, they drew more male than female scientists by an average ratio of four to one. The tendency to draw scientists with laboratory coats and eyeglasses also increased with age, suggesting that children learn multiple stereotypes about scientists as children mature (Finson, Beaver, & Cramond, 1995). These age-related increases are consistent with children gaining more exposure throughout development to male scientists dressed in archetypal laboratory attire.

Change Over Time

The strongest evidence for change over time came from analyses that included data from the 1960s and 1970s (i.e., Chambers, 1983). Less than 1% of children drew a female scientist in Chambers' landmark study (data collection years 1966–1977), but that percentage rose to 28% on average in later studies (data collection years 1985–2016). The historical time effect was robust in all regression models including Chambers' study (ps < .005 in all six models). Even within studies conducted after Chambers', children tended to draw female scientists more often over time, but this trend was less clear (e.g., the p value was .062 in the simple regression model including all studies except Chambers). In other words, U.S. children have drawn more female scientists since the 1960s and 1970s, but the evidence for change within later decades is tentative.

These findings suggest that children's stereotypes associating science with men have weakened over time in the United States, consistent with increases in women's representation in science. However, women also remain underrepresented in several science fields. For instance, in 2013, women were 49% of biological scientists, 35% of chemists, and 11% of physicists and astronomers in the United States (National Science Board, 2016, appendix table 3–12). Children may glean information about these numerical imbalances through multiple sources such as textbooks, classroom and extracurricular experiences, and extensive media content. Consistent with this hypothesis, children in recent years still drew more male than female scientists on average.

Differences by Children's Sex

Both boys' and girls' drawings showed age-related and historical time-related changes in the predicted directions, confirming our main hypotheses separately for boys and girls. In addition, boys drew male scientists much more often than girls. This sex difference is consistent with these drawings partly reflecting children's own gender identities and in-group preferences. Most children have positive attitudes toward their own sex (Dunham, Baron, & Banaji, 2015; Halim, Ruble, Tamis-LeMonda, Shrout, & Amodio, 2017) and therefore may draw their own sex as an expression of these attitudes. For instance, when asked to draw a generic person, usually 70% or more of both boys and girls draw their own sex (Arteche et al., 2010; Picard, 2015). Our meta-analysis found similar results for the youngest children included in Draw-A-Scientist studies. For instance, based on data from the 1980s and onwards (i.e., models excluding Chambers, 1983), 70% of girls and 83% of boys on average drew their own sex at age 6. Children at this age may have defaulted to drawing their own sex because they had limited knowledge of scientists (Lee, 2010; Newton & Newton, 1992).

As children grew older, both girls and boys drew male scientists more often, likely reflecting their increased gender-science stereotypes. For data from the 1980s and onwards (i.e., models excluding Chambers, 1983), girls switched to drawing more male than female scientists between the ages 10 and 11, roughly corresponding to the end of elementary school. By age 16, girls drew more male than female scientists by an average ratio of three to one. The age-related change for boys was less apparent in graphs of mean percentages (e.g., Figure 1f) because boys started at a much higher mean percentage of male scientists (e.g., 83% at age 6). Likewise, although both boys and girls drew female scientists more often over time, this historical time-related change might appear to some readers to be less dramatic for boys than girls. For instance, between the years 1985 and 2016, the mean percentage of female scientists rose from 33% to 58% for girls (25% points) and 2.4% to 13% for boys (10% points), based on data excluding Chambers. However, differences in baseline rates must be taken in account when interpreting these changes. For instance, viewed another way, boys' likelihood of drawing female scientists rose by over 400% between the years 1985 and 2016, which could be considered a large change. Hence, interpretation of whether changes were larger for girls versus boys is partly a matter of perspective. In fact, the age-related and historical time-related changes tended to be stronger in magnitude for boys than girls on a log odds scale, but these differences were not statistically significant. Most importantly, our main hypotheses about age-related and historical time-related change were confirmed separately for girls and boys.

Methodological Contributions

In addition to providing insight on the development of children's gender-science stereotypes, our research illustrates how developmental scientists can use meta-analytic methods to study age and historical period effects. These effects are often difficult to study in tandem because age-related differences in studies can stem from maturational processes or changes in the sociocultural context in which children are reared and observed. The traditional approach to addressing this issue is to collect longitudinal data, but this approach is limited when studying cultural change because longitudinal studies often include only one cohort of children. In contrast, meta-analyses are well suited to simultaneously investigate developmental and cultural change if the relevant studies include children of varying ages and span several decades of data collection.

Meta-analysts adopting this approach, however, should still be cautious by considering potential confounds. For instance, our approach for studying developmental change focused on cross-sectional age comparisons (e.g., 8-year-olds in 2010 vs. 14-year-olds in 2010). These cross-sectional comparisons controlled for data collection year, but not birth cohort (e.g., 8-year-olds in 2010 were born later in time than 14-year-olds in 2010). Comparing the estimated magnitudes of age and period effects can help evaluate the potential impact of such confounds. For instance, in our meta-analysis, change over children's age was much more rapid than change over data collection year, suggesting that historical differences in children's rearing cannot alone account for the large age effect. In other words, these results suggested genuine change over children's development in addition to slower change over historical time.

Limitations

Several limitations of our research should be noted. First, our meta-analysis included only one measure of children's stereotypes of scientists (i.e., the Draw-A-Scientist Test). Although multiple stereotype measures could have provided additional insight, analyzing a single measure also enabled us to cleanly compare studies on a simple common metric (i.e., percentage of male scientists). Including multiple measures could have presented additional interpretational challenges because different tasks may measure different constructs (for an example of these interpretational issues, see Signorella, Bigler, & Liben, 1993). Second, our meta-analysis included only studies from the United States. Our literature search found some studies from other nations, but only U.S. research provided enough relevant studies of children to study change over several decades of data collection. Investigating differences across nations was therefore beyond the scope of this meta-analysis. Third, most studies were convenience samples, not nationally representative samples. For instance, several researchers were teachers who assessed students in their own classrooms (e.g., Bohrmann & Akerson, 2001). Differences in local sample populations could have added extraneous between-study heterogeneity, making differences across other study features (e.g., data collection year) more difficult to observe. However, in defense of our findings, age-related and historical time-related differences were still found despite this additional heterogeneity.

Conclusion and Implications

In summary, the Draw-A-Scientist literature provided a valuable opportunity to study developmental and cultural change in the same meta-analysis and compare studies on a simple common metric assessing children's associations of science with men. Our meta-analysis is the first systematic, quantitative review of this extensive literature spanning 5 decades of data collection. Based on 78 studies with over 20,000 children, U.S. children's drawings of scientists depicted female scientists more often in later decades but less often among older children. These results suggest both age-related and historical time-related changes in children's gender-science stereotypes. The time-related change was consistent with increases in women's representation in U.S. science. However, even in recent years, children may still learn to associate science with men because women remain underrepresented in some science fields. Consistent with this hypothesis, children in recent samples still drew more male than female scientists on average.

Stereotypes linking science with men might limit girls' interests in science-related activities and careers, as some theories of gender development would predict (e.g., Hyde, 2014; Martin & Ruble, 2010). For instance, girls may avoid activities that they consider appropriate for boys but not girls, as some correlational and experimental studies have suggested (e.g., Bian, Leslie, & Cimpian, 2017; Weisgram, 2016). Girls might also underperform on evaluative tests in male-stereotyped domains such as mathematics and science (Galdi et al., 2014; but see also Flore & Wicherts, 2015, for evidence of publication bias in this stereotype threat literature). Furthermore, children may learn other related traits of scientists that also limit girls' science-related aspirations. For instance, scientists are often seen as agentic (e.g., competitive, independent) but not communal (e.g., helpful, sociable), which conflicts with traits people usually associate with women (Carli, Alawa, Lee, Zhao, & Kim, 2016). This cultural mismatch between traits associated with scientists and women might also dampen girls' interest in science careers (Diekman, Steinberg, Brown, Belanger, & Clark, 2017).

Children's stereotypes of scientists could therefore partly shape sex differences in science-related interests (Gunderson et al., 2012; Hyde, 2014). Girls in recent years may now develop these interests more freely because these stereotypes of scientists have become more androgynous over time. Nevertheless, women remain underrepresented in several science fields, and information about such imbalances is filtered through multiple sources such as mass media and social interactions. Children's drawings of scientists provide one fruitful way to study how children integrate information from these sources to form stereotypes about scientists.

References

Arteche, A., Bandeira, D., & Hutz, C. S. (2010). Draw-A-Person test: The sex of the first-drawn figure revisited. *The Arts in Psychotherapy*, *37*, 65–69. https://doi.org/10.1016/j.aip.2009.09.002

Arthur, A. E., Bigler, R. S., Liben, L. S., Gelman, S. A., & Ruble, D. N. (2008). Gender stereotyping and prejudice in young children. In S. R. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66–86). New York, NY: Oxford University Press.

Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*, 389–391. https://doi.org/10.1126/science.aah6524

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*, 162–166. https://doi.org/10.1111/j.1467-8721.2007.00496.x

Bohrmann, M. L., & Akerson, V. L. (2001). A teacher's reflections on her actions to improve her female students' self-efficacy toward science. *Journal of Elementary Science Education*, *13*, 41–55. https://doi.org/10.1007/BF03176219

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley. https://doi.org/10.1002/9780470743386

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly*, *40*, 244–260. https://doi.org/10.1177/0361684315622645

Cavallo, A. (2007). Draw-A-Scientist/mystery box redux. *Science and Children*, *45*, 37–41. http://www.jstor.org/stable/43175509

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75–141. https://doi.org/10.1177/1529100614541236

Chambers, D. W. (1983). Stereotypic images of the scientist: The Draw-A-Scientist test. *Science Education*, *67*, 255–265. https://doi.org/10.1002/sce.3730670213

Cvencek, D., & Meltzoff, A. N. (2015). Developing implicit social cognition in early childhood: Methods, phenomena, prospects. In S. Flannery Quinn & S. Robson (Eds.), *The Routledge international handbook of young children's thinking and understanding* (pp. 43–53). Abingdon, UK: Routledge.

Diekman, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., & Clark, E. K. (2017). A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in STEM gender gaps. *Personality and Social Psychology Review*, *21*, 142–175. https://doi.org/10.1177/1088868316642141

Dunham, Y., Baron, A. S., & Banaji, M. R. (2015). The development of implicit gender attitudes. *Developmental Science*, *19*, 781–789. https://doi.org/10.1111/desc.12321

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781446249222.n49

Farland, D. (2006). The effect of historical, nonfiction trade books on elementary students' perceptions of scientists. *Journal of Elementary Science Education*, *18*, 31–47. https://doi.org/10.1007/BF03174686

Finson, K. D. (2002). Drawing a scientist: What we do and do not know after fifty years of drawings. *School Science and Mathematics*, *102*, 335–345. https://doi.org/10.1111/j.1949-8594.2002.tb18217.x

Finson, K. D., Beaver, J. B., & Cramond, B. L. (1995). Development and field test of a checklist for the Draw-A-Scientist Test. *School Science and Mathematics*, *95*, 195–205. https://doi.org/10.1111/j.1949-8594.1995.tb15762.x

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Flick, L. (1990). Scientist in residence program improving children's image of science and scientists. *School Science and Mathematics*, *90*, 204–214. https://doi.org/10.1111/j.1949-8594.1990.tb15536.x

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, *53*, 25–44. https://doi.org/10.1016/j.jsp.2014.10.002

Fort, D. C., & Varney, H. L. (1989). How students see scientists: Mostly male, mostly white, and mostly benevolent. *Science and Children*, *26*, 8–13. http://www.jstor.org/stable/43167098

Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The roots of stereotype threat: When automatic associations disrupt girls' math performance. *Child Development*, *85*, 250–263. https://doi.org/10.1111/cdev.1212

Gonzalez, A. M., Dunlop, W. L., & Baron, A. S. (2017). Malleability of implicit associations across development. *Developmental Science*, *20*, e12481. https://doi.org/10.1111/desc.12481

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 152–166. https://doi.org/10.1007/s11199-011-9996-2

Halim, M. L. D., Ruble, D. N., Tamis-LeMonda, C. S., Shrout, P. E., & Amodio, D. M. (2017). Gender attitudes in early childhood: Behavioral consequences and cognitive antecedents. *Child Development*, *88*, 882–899. https://doi.org/10.1111/cdev.12642

Hedges, L. V., & Vevea, J. L. (1998). Fixed and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486. https://doi.org/10.1037/1082-989X.3.4.486

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: American Association of University Women. Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*, 373–398. https://doi.org/10.1146/annurev-psych-010213-115057

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *14*, 165–173. https://doi.org/10.1080/00224540109600544

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology*, *107*, 371–392. https://doi.org/10.1037/a0037215

Lee, T. R. (2010). *Young children's conceptions of science and scientists*. Doctoral dissertation. Available from ProQuest Dissertations and Theses database. (UMI No. 3421874)

Long, M., Steinke, J., Applegate, B., Knight Lapinski, M., Johnson, M. J., & Ghosh, S. (2010). Portrayals of male and female scientists in television programs popular among middle school-age children. *Science Communication*, *3*, 356–382. https://doi.org/10.1177/1075547009357779

Losh, S. C., Wilke, R., & Pop, M. (2008). Some methodological issues with "Draw A Scientist Tests" among young children. *International Journal of Science Education*, *30*, 773–792. https://doi.org/10.1080/09500690701250452

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. https://doi.org/10.1111/j.0963-7214.2004.00276.x

Martin, C. L., & Ruble, D. N. (2010). Patterns of gender development. *Annual Review of Psychology*, *61*, 353–381. https://doi.org/10.1146/annurev.psych.093008.100511

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *107*, 631–644. https://doi.org/10.1037/edu0000005

Miller, D. I., & Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: A 30-year analysis. *Frontiers in Psychology*, *6*, 36. https://doi.org/10.3389/fpsyg.2015.00037

National Science Board. (2016). *Science & engineering indicators 2016* (NSB Publication No. 2016-1). Arlington, VA: National Science Foundation. Retrieved from https://www.nsf.gov/statistics/2016/nsb20161/uploads/1/6/at03-12.pdf

National Science Foundation. (2017). *IPEDS completions survey, integrated science and engineering resources data system (WebCASPAR)*. Retrieved from https://ncsesdata.nsf.gov/webcaspar/

Newton, D. P., & Newton, L. D. (1992). Young children's perceptions of science and the scientist. *International Journal of Science Education*, *14*, 331–348. https://doi.org/10.1080/0950069920140309

Picard, D. (2015). Sex differences in scores on the Draw-A-Person Test across childhood: Do they relate to graphic fluency? *Perceptual and Motor Skills*, *120*, 273–287. https://doi.org/10.2466/10.27.PMS.120v10x1

Pienta, R. S., & Smith, M. (2012). Women on the margins: The politics of gender in the language and content of science textbooks. In H. Hickman & B. J. Porfilio (Eds.), *The new politics of the textbook: Problematizing the portrayal of marginalized groups in textbooks* (pp. 49–68). Rotterdam, The Netherlands: Sense.

Previs, K. K. (2016). Gender and race representations of scientists in *Highlights for Children*: A content analysis. *Science Communication*, *38*, 303–327. https://doi.org/10.1177/1075547016642248

Rawson, C. H., & McCool, M. A. (2014). Just like all the other humans? Analyzing images of scientists in children's trade books. *School Science and Mathematics*, *114*, 10–18. https://doi.org/10.1111/ssm.12046

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review*, *13*, 147–183. https://doi.org/10.1006/drev.1993.1007

Simonsohn, U. (2017, March 21). *The funnel plot is invalid because of this crazy assumption: r(n,d)=0 [Blog post]*. Retrieved from http://datacolada.org/58

Smith, P. S., Banilower, E. R., McMahon, K. C., & Weiss, I. R. (2002). *The national survey of science and mathematics education: Trends from 1977 to 2000*. Chapel Hill, NC: Horizon Research. Retrieved from http://2000survey.horizon-research.com/reports/trends/trends_report.pdf

Smyth, F. L., & Nosek, B. A. (2015). On the gender–science stereotypes held by scientists: Explicit accord with gender-ratios, implicit accord with scientific identity. *Frontiers in Psychology*, *6*, 415. https://doi.org/10.3389/fpsyg.2015.00415

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*, 60–78. https://doi.org/10.1002/jrsm.1095

Steinke, J. (2013). Portrayals of female scientists in the mass media. In A. Valdivia & S. R. Mazzarella (Eds.), *The international encyclopedia of media studies* (pp. 1–18). Oxford, UK: Blackwell.

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2007). Assessing media influences on middle school–aged children's perceptions of women in science using the Draw-A-Scientist Test (DAST). *Science Communication*, *29*, 35–64. https://doi.org/10.1177/1075547007306508

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*, 112–125. https://doi.org/10.1002/jrsm.11

Weisgram, E. S. (2016). The cognitive construction of gender stereotypes: Evidence for the dual pathways model of gender differentiation. *Sex Roles*, *75*, 301–313. https://doi.org/10.1007/s11199-016-0624-z

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. *46*, pp. 55–123). London, UK: Elsevier.

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The influence of female role models on women's implicit science cognitions. *Psychology of Women Quarterly*, *37*, 283–292. https://doi.org/10.1177/0361684313482109

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Figure S1. Funnel Plot of Effect Size Plotted Against Sample Size

Figure S2. Change Over Age and Historical Time in Outcomes Other Than Scientists' Sex

Table S1. List of Literature Databases Searched

Table S2. Sources of Unpublished and Published Studies

Table S3. Correlation Matrix for Predictors in Multivariable Meta-Regression Model

Table S4. Regression Coefficients from Multivariable Meta-Regression Models

Table S5. Mean Effect Size and Change Over Time Adjusted for Small-Study Effects

Appendix S1. Applying Inclusion and Exclusion Criteria

Appendix S2. References for Studies in Meta-analysis

Appendix S3. Coding System

Appendix S4. Publication Bias Analyses

Appendix S5. Outlier Analyses

Appendix S6. Differences by Children's Sex

Appendix S7. Within-Study Effect of Age

Appendix S8. Outcomes Other Than Scientists' Sex

Data S1. Data File Used for Main Analyses

Data S2. Data File Used for Within-Study Analyses

Data S3. R Code for Main Analyses

Data S4. R Code for Figure 1

Data S5. R Code for Publication Bias Analyses

Data S6. R Code for Within-Study Analyses

Data S7. R Code for Outcomes Other Than Scientists' Sex