Meta-analysis on Gender Differences in Math Tasks

(authors)

Methods

Literature search

xx

Coding of the dataset

xx

Data analysis

Standardized Mean Differences (SMDs) between males and females in math scores were meta-analyzed. Formulas and analytical strategy suggested by Borenstein et al. (2009) were used. Random effects multilevel models were fitted using the "metafor" package of R. Random intercepts were set for studies, samples (some studies had multiple samples), and effect sizes. Heterogeneity was quantified using the following estimates: τ as the standard deviation (*SD*) of the effects across studies, ω_{sample} as the SD across samples, and ω as the SD across underlying effect sizes. Moderators were mean age of the sample, year of publication, and type of math task. Effect sizes larger than d = 2.00 (absolute value) have provisionally been excluded for double-check and because they are considered implausible and too influent in the analysis.

Results

A total number of 440 studies, 785 samples, and 1198 effect sizes were included in the quantitative synthesis. The estimated overall number of individuals involved is 11344712. The median sample is 268. The overall meta-analytic effect estimated via multilevel random effect model (with intercept only) was d = 0.09 [95%CI: 0.06, 0.11], with substantial heterogeneity across studies (as compared to the mean effect), = 0.22, no estimated heterogeneity across samples, sample = 0.00, and some heterogeneity across individual effect sizes, = 0.15. The overall heterogeneity is significant, Q(1210) = 43,869.16, $p < 0.001$. Due to the number of studies and effects, the forest plot is impossible to represent, but the funnel plot is shown below in Figure 1. Figure 1. Forest plot of all effect sizes.

Moderators Mean age of the sample could be coded for only 153 studies, 197 samples, and 349 effect sizes, so many fewer than the original overall analysis. Mean age of sample was found a significant moderator of the effect size, Q(1) = 13.71, $p < 0.001$, B = 0.014 (for age in years), meaning that the SMD is expected to increase by 0.14 per decade. The predicted SMD as a function of age is shown below in Figure 2. It is about zero at 6 years of age, about d = 0.10 at 12 years of age, and about 0.20 at 18 years of age. Figure 2. Meta-regression with Age (years) as moderator of the effect size. Dots are effect sizes. Blue straight line represents the estimated meta-regression curve. Waving blue line represent a non-linear loess smoother across individual points, which was added as a double-check to ensure that it followed the meta-regression estimate. Shaded areas represent 95% confidence bands.

Year of publication was tested in a separate moderator analysis because could be coded for virtually all studies (unlike Age), so we preferred to test it on the entire dataset. Also, covarying the two moderators appears unnecessary because they are weakly and non-significantly correlated at the between-sample level, r = -0.13, p = 0.09. Year of publication also was a significant moderator of the effect size. It could be coded for 441 studies, 777 samples, and 1175 effect sizes. Its moderating effect was significant, Q(1) = 6.66, p = 0.01, B = 0.001, meaning that the SMD has increased by about 0.10 per decade (but note that the publication year range is only [2010, 2022]). The predicted SMD as a function of age is shown below in Figure 3. Type of task could not be tested as a moderator because it has not been coded yet. Stay tuned!

Figure 3. Meta-regression with Year of publication as moderator of the effect size. Dots are effect sizes. Blue straight line represents the estimated meta-regression curve. Waving blue line represent a non-linear loess smoother across individual points, which was added as a double-check to ensure that it followed the meta-regression estimate. Shaded areas represent 95% confidence bands.

References