Do Executive Functions Mediate the Link between Socioeconomic Status and Numeracy Skills? A Cross-Site Comparison of Hong Kong and the United Kingdom

Michelle R. Ellefson[1]

Antonia Zachariou[1,2]

Florrie Fei-Yin Ng[3]

Qian Wang[3]

Claire Hughes[1]

[1]University of Cambridge

[2] University of Roehampton

[3] Chinese University of Hong Kong

This article was accepted for publication by the *Journal of Experimental Child Psychology* on 18th October 2019. This is the accepted version.

Author Notes

Correspondence concerning this article should be addressed to Michelle R. Ellefson, University of Cambridge Faculty of Education, 184 Hills Road, Cambridge, CB2 8PQ, United Kingdom. Email: mre33@cam.ac.uk.

Abstract

In the fields of education, sociology and economics, there is a long-standing connection between socioeconomic status (SES) and school outcomes in a wide variety of cultural settings, but these studies have yet to examine the possible mediating effects of domain-general cognitive factors such as executive functions (EF). Addressing this gap and building on evidence for links between EF and numeracy, the current cross-cultural study used a large sample (N = 835) of 9- to 16-year-old children from Hong Kong and the United Kingdom to the examine the independence and interplay of SES and EF as predictors of numeracy skills. Our analyses yielded three key findings: (1) EF consistently predicts numeracy skills across sites and genders; (2) Associations between SES and EF differ by site and gender; and (3) Associations between numeracy skills and SES/EF differ by site and gender. Together with previous findings, our results suggest culture-specific associations between SES, EF and numeracy, indicating that cultural insights may enable impactful shifts in public policy to narrow the achievement gap between children from affluent and disadvantaged families.

Keywords

Executive Functions, Cross-cultural Research, Socioeconomic status, Numeracy Skills.

Do Executive Functions Mediate the Link between Socioeconomic Status and Numeracy Skills? A Cross-Site Comparison of Hong Kong and the United Kingdom

There is a worrying academic achievement gap for children from families facing high levels of social and/or economic disadvantage. For example, Sirin's (2005) meta-analysis of 58 studies indicated a medium effect size difference (mean .29, median .24, range .005 to .77) in academic achievement between based on low vs. high socioeconomic status. The size and reliability of this gap indicates that identifying the underpinning mechanisms is a vital first step towards developing effective interventions. One mechanism could be executive functions (EF), defined as the set of higher order cognitive skills such as inhibitory control, working memory and cognitive flexibility that underpin flexible and goal-directed behaviour. EF development is protracted in nature, making it especially susceptible to environmental influence (Blair & Raver, 2016; Farah, 2018; Johnson et al., 2016). Numerous studies have reported statistically significant associations between EF and SES (e.g., Blair et al., 2011; Blums et al., 2017; Howard et al., 2019; Hughes & Ensor 2005; Rhoades et al., 2011; Sarsour et al., 2011).

Variation in EF skills reliably predict children's school readiness (e.g., Blair & Razza, 2007; Hughes et al., 2010) and are related to academic achievement in middle childhood and adolescence (e.g., Best et al., 2011; St Clair-Thompson & Gathercole, 2006). In these studies, EF is measured using either researcher-administered tasks or computer-administered tasks. Research-administered tasks tend to be conducted one-to-one by a trained psychologist/researcher and a participant; they usually involve manipulatives (e.g., blocks) and verbal questions/instructions. Computer-administered tasks are those that are presented on computers, or more recently tablets; they can be conducted either one-to-one or in larger groups.

Many studies specifically focus on the link between EF and one type of academic skill - numeracy (for a review, see Bull & Lee, 2014). Numeracy is a broad term that covers the wide range of computational and mathematical skills that students learn in formal educational settings. While these associations are typically cross-sectional, there is some indication that earlier EF predicts later numeracy skills (e.g., Ahmed et al., 2018; Bull et al., 2008; Viterbori et al., 2015). Most studies have focused on pre-school children (e.g., Clark et al., 2013; Fuhs et al., 2014; Welsh et al., 2010), but the relationship between EF and numeracy appears robust across developmental stages (e.g., Cragg & Gilmore, 2014; Cragg et al., 2017; Samuels et al., 2016).

Numeracy is a particularly interesting context to study the interplay between EF, SES and academic skills because of the well-documented gaps in achievement of older children from low compared to higher SES homes, and the consistent links between EF and numeracy. Despite a number of studies of this interplay for young children (e.g., Duncan et al., 2017; Nesbitt, Baker-Ward, Willoughby, 2013), far less is known for older children. A recent NIH study that tracked 336 children and adolescents from the USA across a two-year period showed that EF (but not verbal memory) mediated the association between SES and progress in arithmetic (Lawson & Farah, 2017). It is unknown what underlies this mediation, but it might be that SES is a proxy for family differences in nutrition, stress, or parenting. Taking the example of parenting, Hughes and Devine (2019) reported that, in a United Kingdom (UK) sample, parental scaffolding (age appropriate guidance that is just enough support for a child to complete a difficult task) facilitates child gains in EF skills, while negative parent-child interactions (e.g., harsh tones, criticism, arguments, physical control) have an adverse effect. Parenting effects like these might explain why EF skills mediate the relationship between SES and numeracy.

It is not yet known whether this mediation effect can be generalized to samples outside the USA. Further, it cannot be assumed that findings in the USA (e.g., Lawson & Farah, 2017) will generalize to other English-speaking countries with similar societal values. For example, in a meta-analysis on studies ranging from early childhood to adolescence, Tucker-Drob and Bates (2015) found a stronger influence of SES on intelligence and school achievement are larger for samples from the USA compared with those from Western Europe and Australia. To our knowledge just two studies exploring SES effects on EF have included older children from Asia – with both showing only small effect sizes. Phillipson (2009) gave researcher-administered EF tasks to a Hong Kong (HK) sample of 215 primary school students (X_{age} = 10.7 years) and found a small, but statistically significant correlation with SES (*r* = .21). Wang and colleagues (2016) applied a combination of researcher- and computer-administered EF tasks for a larger model exploring predictors of EF and theory of mind in two different samples comparing children from HK and the UK (Sample 1: 9-16 years, *n* = 118, Sample 2: 10-12 years, *n* = 137) and found small, but statistically significant links between SES (using a proxy score based on school). Likewise, the only existing UK study to examine whether EF helps explain the SES gradient in numeracy links involved a sample of children followed across the transition to school (Devine et al., 2016), such that it is not known whether EF has a similar mediation effect for older children.

Furthermore, international comparisons of SES and numeracy skills suggest that examining cultural universality or specificity could help illuminate environmental influences on cognitive development. Importantly for the current study, children from Eastern Asia consistently outperform their Western counterparts on tests of both EF (e.g. for preschoolers – Imada et al., 2013; Lewis et al., 2009; Sabbagh et al., 2006; for school-aged children – Ellefson et al., 2017; Wang, et al., 2016) and numeracy skills (e.g., TIMMS - Trends in International Mathematics and Science Study, Mullis et al., 2016 and PISA - Programme for International Student Assessment, OECD, 2016). Also of note is the finding that variation in numeracy skills show only weak links with SES for students from HK. Liu and Lu (2008) reported that SES accounted for only .8% of the variance in PISA scores Kalaycıoğlua (2015) found that links between SES and PISA scores were weakest in HK compared to other countries in a sample of adolescents from HK, the Netherlands, Greece, England and the USA (*N* = 8,806). Less well established, however, are issues of (1) cultural universality in the relationship between EF and numeracy; and (2) whether hat EF mediates in the link between SES and numeracy.

Numerous studies report internationally consistent EF-numeracy links (for reviews see Bull & Lee, 2014; Friso-van den Bos et al., 2013). When focusing on older children there are consistent links when comparing samples from the USA (e.g., Lawson & Farah, 2017), UK (e.g., Bull et al., 2008; Cragg et al., 2017; St Clair-Thompson & Gathercole, 2006) Netherlands (van der Sluis et al., 2007), and Singapore (e.g., Lee et al., 2009). Looking outside of North American and Europe, a longitudinal study of school-aged children from China indicates similar EF-numeracy links (Wei et al., 2018). However, very few studies have directly examined whether the link between EF and numeracy skills is similar in magnitude across different cultures. In one exception, Lan and colleagues (2011) reported that, alongside the expected advantage shown by Chinese preschoolers on tests of inhibition and attentional control relative to their American counterparts, variation in EF showed similar associations with numeracy skills in both sites. Ng and colleagues (2015) replicated this result across different ethnic groups with a sample of preschool children living in the USA with Chinese, African American, Dominican and Mexican heritage. Given that EF and numeracy skills are both cognitive processes, we expect that the links between them should continue to be culturally universal for older children and adolescents. However, to our knowledge, this has yet to be demonstrated.

In a recent meta-analysis, Lawson and colleagues (2018) found that although only small to medium in effect size links between SES and EF were very consistent between studies. Notably, 27 of the 33 studies included in this meta-analysis were from North America (25 USA, 2 Canada), with just 3 European studies (1 from Germany, 2 from the Netherlands) and 3 studies from elsewhere (Madagascar, HK and Turkey).To our knowledge, the Lawson and Farah (2017) study is the only published study to show mediation effect of EF on the link between SES and numeracy. However, as stated earlier, it includes only a sample from the USA and it is unknown whether this mediation effect is culturally universal.

At this point it is worth noting that a gender perspective may be valuable in providing a finer-grained picture of possible cultural universality in the processes underpinning individual differences in numeracy skills, not least because gender contrasts in numeracy vary in magnitude across different countries (e.g., Shen et al., 2016; Voyer & Voyer, 2014). Indeed, a meta-analysis of TIMMS and PISA data from 69 countries by Else-Quest and colleagues (2010) found that while gender differences in numeracy skills in mid-adolescence were typically either small or negligible in magnitude, effect sizes differed by country. Reports of gender differences in EF skills are rare, which may well reflect the lack of major gender differences in cognition evident in several meta-analyses (e.g., Grissom & Reyes, 2019; Hyde, 2016).

In a large cross-sequential longitudinal study of children (N = 673, beginning ages 5 to 12 years) from Singapore, Lee & Bull (2016) found that SES, but not gender, plays a predictive role on working memory development, which influenced numeracy skills. A separate large study of adolescents from Finland (N = 619, ages 14 to 15 years) identified significant differences between males and females on many EF measures, especially those that involve classroom-based metrics (Holm et al., 2018). However, these differences did not carry over when looking at those with mathematical difficulties. However, in the UK at least, societal concerns about academic under-achievement in children from low-income families typically focus on males (e.g., Cobb-Clark & Moschion, 2017).

Longitudinal findings from a sample of young children from the USA suggest that links between EF and numeracy skills may be stronger for females than for males (Clark et al., 2013). Although there is a dearth of evidence, given that both are cognitive processes, we would expect that the concurrent (i.e., non-longitudinal) links between EF and numeracy should be culturally universal for both males and females during older childhood and adolescence. Lawson and Farah (2017) have used gender as a control variable. Perhaps sample sizes have been too small to warrant reporting small effect size gender-specific associations. For example, in a study of 453 4- to 10-year-old Portuguese children, Alves and colleagues (2016) found gender differences on measures of cognitive performance (that did not include EF tasks) varied in magnitude across different SES groups, but all were small effects (*partial* η^2 range .013 to .021).

Likewise, in a comparison of TIMMS arithmetic data from two HK cohorts, Guo, Marsh, Parker, Morin and Yeung (2015) found that SES showed similarly small positive associations with males' and females' TIMMS scores (weighted mean effect size = .11). In contrast, a later study of 598 elementary school children from China, Guo et al. (2018) found that links between SES and literacy differed for boys and girls. Specifically, while parental education showed direct effects on literacy for both genders, family income had direct effects on literacy for females, but was indirectly related to literacy (via parent-child communication) for males. Taken together, these findings suggest that SES effects are not consistent for males and females growing up in different countries.

The Current Study

Given the gaps in the literature, the key aim of the current study is to build on the models used by Lawson and Farah (2018) to evaluate the cultural universality of EF as a mediator of the association between SES and numeracy. In addition, given the data that male and female numeracy outcomes seem to be differentially affected by SES, we include a direct investigation of whether these links are consistent across genders.

We focus on an aspect of numeracy skill that is common on many standardized tests of mathematical skill: being able to solve equations at different levels of difficulty ranging from simple addition and subtraction up to basic calculus. We selected this focus because, like our EF measures, it can be administered in different languages without translations except for basic instructions.

Even though there hasn't been a direct study of the cultural universality of links between EF and numeracy, the various studies published across cultures suggest that EF and numeracy links should be consistent across cultures. It is unknown whether there is a cultural universality for links between SES and EF or for the posited mediating effect of EF on the link between SES numeracy. Furthermore, the mediation effects reported by Lawson and Farah (2017) in a sample from the USA cannot be assumed to be generalizable to other countries, whether it be those that share language and many societal norms (e.g., UK, Australia) or those that do not (e.g., HK, Netherlands, Singapore)

To achieve these goals, we evaluated further a large cross-cultural dataset that has be used to report cross-cultural contrasts in EF using a relatively large sample of children from HK and UK (Ellefson et al., 2017). The UK site shares language and some societal norms with the USA, while the HK site shares fewer societal and language norms. The two sites have distinct cultures and educational systems yet show similar gender equity ratings for performance on TIMMS/PISA tests (Else-Quest et al., 2010). Importantly, these similarities sit alongside of overall higher performance of children in HK on EF (Ellefson et al., 2017) and TIMMS/PISA tests (Mullis et al., 2016; OECD, 2016) and some evidence that SES has less influence on their numeracy skills (Kalaycıoğlua, 2015; Liu & Lu, 2008).

Method

Participants

The sample for this study comes from a larger cross-cultural data set including 9- to 16-year-old children (*N* = 930) and their parents from HK and the UK. In the current study, we identified 835 children who had:

(1) at least one of four EF tasks measuring inhibition, working memory, cognitive flexibility and planning had been completed (from the UK sample - one child completed one EF task, one child completed two EF tasks, seven children completed three EF tasks and the rest of the UK sample and all of the HK sample completed all four EF tasks);

(2) a standardized arithmetic skills assessment that included a variety of mathematical computations and the family affluence scale had each been completed;

(3) parent-provided highest level of educational attainment and occupation information in a survey or it could be estimated based on school data; and

(4) gender and birthday/age reported;

Within this group, 15 participants from HK and 42 from the UK reported home languages other than those taught in school. The ethics committees from both universities running the project reviewed and approved this research protocol. All parents provided consent and children verbal assent.

As reported in Ellefson et al. (2017), we recruited participants from state schools and community events in ways that would allow for samples representing the sites. The local geographical area for the researchers based in the UK is affluent, so although there are some participants from this area, we specifically recruited in less affluent areas to be more representative of the UK. The full dataset includes measures of parenting and parent-child relationships, self-regulated learning, general cognitive ability, verbal comprehension (in English for the UK site, Cantonese for the HK site), and English vocabulary skills. Due to content and space limitations, those data will be reported separately (for full dataset see http://reshare.ukdataservice.ac.uk/852570/)

Materials and Procedures

Numeracy skills. We assessed numeracy skills using the mathematics subtest of the Wide Range Achievement Test (WRAT; Wilkinson, 1993). We selected this task because it can be easily administered in group settings, is culturally robust (Snelbaker et al., 2001), has high immediate re-test reliability (.88; Dell et al., 2008), shows good levels of validity (.66; Wilkinson & Robertson, 2006) and covers a wide range of ages. The assessment itself measures computational skills from early childhood to adolescence. The easiest items involve counting. After that there are a variety of equations that increase in complexity. For example, the initial equations are simple addition or subtraction and the later equations include fractions, algebra and basic calculus. The task itself does not need translating, but the instructions were delivered in Cantonese for the HK site and English for the UK site.

EF skills. As described in Ellefson et al. (2017; open access materials from [http://reshare.ukdataservice.ac.uk/852570/]

]), participants completed four EF tasks measuring inhibition (Stop Signal task, Logan, 1994), working memory (Spatial Span task, Corsi, 1971), cognitive flexibility (Color-Shape Switching task, Ellefson et al., 2006; Rogers & Monsell, 1995) and planning (Tower of Hanoi, Welsch, 1991). Accuracy and Reaction Times (RTs) were recorded for each trial. Efficiency scores for each task were calculated by dividing total number of accurate trials by the RT (in seconds).

SES. Parents reported on highest level of education and current occupation, providing the same information about their spouse / partner, where appropriate. These responses were coded into socioeconomic metrics using Hollingshead's Four-Factor Index of Social Status (1975; see also Adams & Weakleim, 2011). A parent's highest educational qualification is converted into scores ranging from 1 to 7. For example, 1 is for primary/elementary school, 4 is for high school, 6 for an undergraduate degree, and 7 for a graduate degree. Parental occupation is converted into scores ranging from 1 to 9. For example, 1 is for cleaners or farm laborers, 5 is for clerical and sales workers, 7 is for owners of small businesses, managers or journalists, 9 for executives, scientists, engineers, or large business owners. For single/widowed parents, then only their scores were used in the analyses, otherwise mean education and occupation scores was computed based on both parents. In addition, if one parent was a full-time homemaker, then the occupation score was based on the parent in-work.

In addition, children answered five questions from the Family Affluence Scale (Wardle et al., 2002), which has good reliability and validity (Boudreau & Poulin, 2009. Specifically, children reported if they had their own bedroom, the number of bedrooms in their home, as well as the number of family vehicles, desktop/laptop computers, and family vacations in the last year. A composite score was calculated by summing the scores for each question together for each participant in which higher scores indicate more affluence (range of scores = 1 to 18).

Results

Preliminary Analyses and Analytic Plan

Data screening indicated that some of the variables violated normality; we made appropriate adjustments to our statistical analyses to accommodate this. A preliminary χ^2 analysis indicated no site difference in the number of males and females (χ^2 = 0.69, *p* = .41). A preliminary factorial ANOVA with age as the dependent variable indicated that HK participants were older (*M* = 12.21 years, *SD* = 0.99) than UK participants (*M* = 11.92 years, *SD* = 0.93; F (1, 831) = 19.61, p < .001, η_p^2 = .02), but males (*M* = 12.11 years, *SD* = 0.99) and females (*M* = 11.99 years, *SD* = 0.94) did not differ in age (F (1, 831) = 2.54, p = .11, η_p^2 = .003).

Next, we ran factorial ANCOVAs using the factors site (HK, UK) and gender (Female, Male) on each of the key measures in the study to identify overall group differences in numeracy skills, SES, and EF skills using the car/effects (Fox & Weisberg, 2011); lsr (Navarro, 2015); psych (Revelle, 2018) packages for R. Age was included as a co-variate in each of these analyses and effect sizes were evaluated using partial eta-squared (η_p^2). We ran non-parametric analyses for each of the non-normally distributed variables; those results matched the parametric results that we report below.

Finally, to explore the links among EF, SES and numeracy, we tested a theoretical model using structural equation modeling (see Figure 1) using the lavaan (Rosseel, 2012). and semTools (Jorgensen et al., 2018) packages for R. In the hypothesized model, the three SES scores load onto an SES latent variable and the four EF scores load onto an EF latent variable. As there is just one numeracy task, the WRAT standardized score is a measured variable.

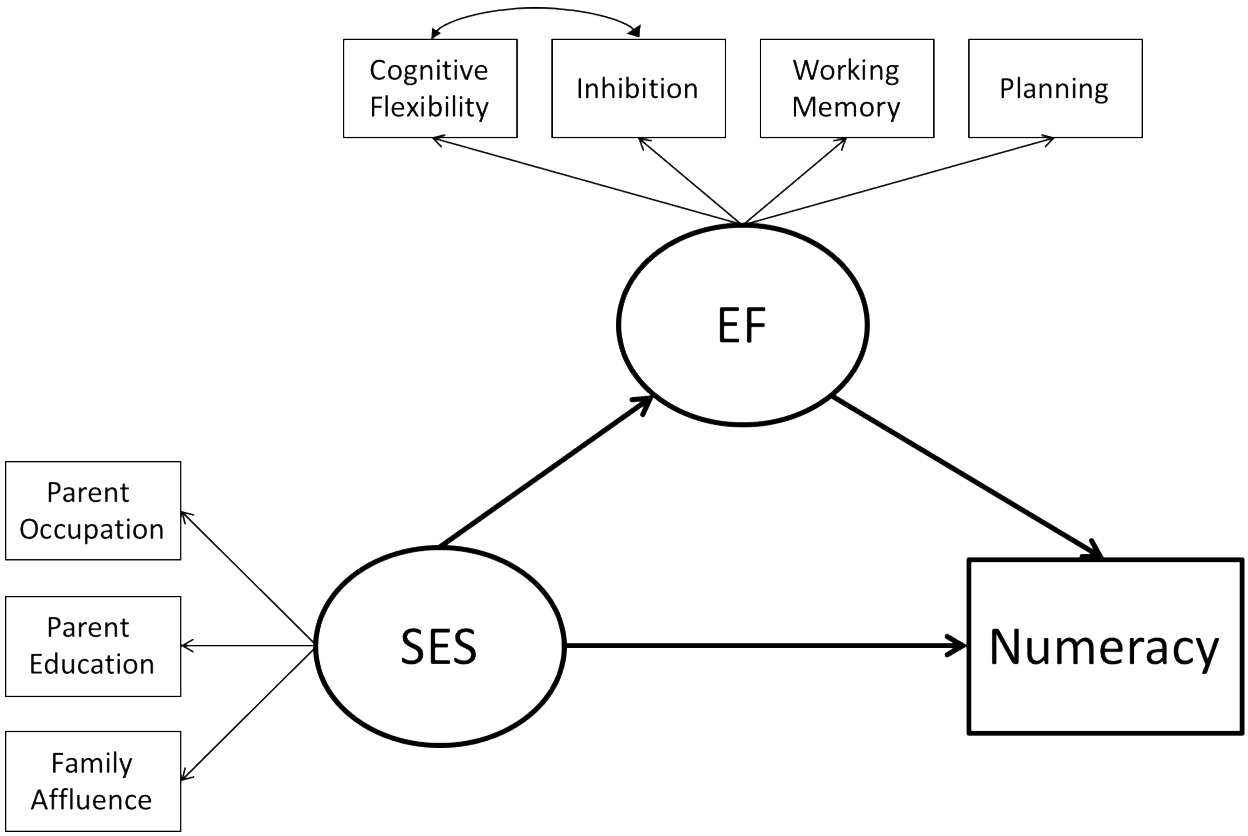

Figure 1. The full theoretical model tested across the four groups (UK male, UK female, HK male, HK female). The correlated residual was included based on preliminary analyses of the data. All other connections were determined a priori.

Before running the correlations and models, we took a few preliminary steps. Following the procedures adopted by Lawson and Farah (2017), we used normed scores from the WRAT and created age-adjusted EF and SES variables, by taking residuals from a linear regression with age as the predictor and each individual measure of EF / SES variable as an outcome. Finally, following procedures adopted by Ellefson, et al. (2017) and Wiebe, Espy and Clark (2008), we converted all measured variables into full-sample *z* scores. We calculated robust standard errors and adjusted for non-normal distribution in some of the measured variables using the Satorra-Bentler estimation (Brown, 2015).

Correlating residuals can inflate a model's fit, when measured variables share modalities. In such cases, it is best for models to include correlated residuals (e.g., Cole et al., 2007). Here, the EF tasks are all computerized and have shared modalities. We tested for correlated residuals and found that cognitive flexibility and inhibition were reliably correlated. As such, we correlated these two residuals in the full model. We removed this correlation in the nested models to test for parsimony and to make sure that including it did not unnecessarily inflate the fit of the model.

To determine whether it was more appropriate to run a single versus multiple models across site and gender, we used confirmatory factor analysis to test for measurement invariance of the measurement model (see Table 1). When comparing across sites there was configural (equal form) and metric invariance (equal factor loadings). However, there was not scalar invariance (equal intercepts) or strict measurement invariance (equal residuals). When comparing across gender x site there was again configural and metric invariance, but not scalar or strict measurement invariance.

Taken together, these preliminary analyses indicated that the sites and genders had a similar factor structure, but there were other heterogeneities across sites and genders. As such we ran models separately for UK males, UK females, HK males and HK females. After running the full model, we tested various nested models to identify the most parsimonious model for each group. Generally, the measured variables part of each latent variable are assumed to be independent, with their factor loadings and error variance free to vary. Following Brown (2015), we ran a tau equivalence test for the EF and SES latent variables as part of these nested model analyses to confirm this independence.

Table 1.

Measurement Invariance results comparing across sites and genders.

	χ^2	df	p	$\Delta\chi^2$	Δdf	p	RMSEA	SRMR	CFI	TLI
Site										
Equal form	51.06	34	.03				.04	.03	.98	.97
Equal loadings	58.49	39	.02	7.03	5	.22	.04	.04	.98	.97
Equal intercepts	181.33	44	<.001	127.15	5	<.001	.09	.07	.84	.79
Equal residuals	328.30	51	<.001	76.68	7	<.001	.11	.10	.67	.64
Gender										
Equal form	119.39	34	<.001				.08	.05	.93	.89
Equal loadings	121.67	39	<.001	2.30	5	.81	.07	.05	.94	.91
Equal intercepts	132.74	44	<.001	11.47	5	.04	.07	.05	.93	.91
Equal residuals	149.69	51	<.001	11.21	7	.13	.07	.06	.93	.92
Site x Gender										
Equal form	93.60	68	.02				.04	.04	.97	.95
Equal loadings	103.86	83	.06	10.04	15	.82	.04	.05	.98	.97
Equal intercepts	240.79	98	<.001	147.79	15	<.001	.08	.08	.83	.81
Equal residuals	422.44	119	<.001	104.53	21	<.001	.11	.11	.64	.66

Sample Characteristics

Numeracy skills differ by site, but not gender. As indicated in Table 2, there were large main effects of site (favouring children from HK) on the WRAT raw scores (F (1, 831) = 970.53, p < .001, η_p^2 = .54), standardised scores (F (1, 831) = 824.19, p < .001, η_p^2 = .50) and grade equivalent scores (F (1, 831) = 1115.53, p < .001, η_p^2 = .57). In contrast, main effects of gender and site by gender interactions were not statistically significant for any of the WRAT variables (all $\eta_p^2 < .01$).

EF skills differ by site, and sometimes by gender. Table 2 shows main effects of site (favouring children from HK) for all EF raw efficiency scores before being adjusted for age: inhibition (F (1, 830) = 114.84, $p < .001$, $\eta_p^2 = .12$); cognitive flexibility (F (1, 829) = 95.87, $p < .001$, $\eta_p^2 = .10$); working memory (F (1, 824) = 98.60, $p < .001$, $\eta_p^2 = .11$); and planning (F (1, 828) = 40.24, $p < .001$, $\eta_p^2 = .05$). In addition, females had higher cognitive flexibility scores than males (F (1, 828) = 5.09, $p = .02$, $\eta_p^2 = .01$). No other main effects of gender or site by gender interactions were statistically significant (all $\eta_p^2 < .01$).

SES differs by site, but not gender. In contrast with the between-site differences in children's cognitive performances noted above, main effects of site for all SES scores (before being adjusted for age) were in the opposite direction (i.e., favouring children from UK, see Table 2): parental education (F (1, 831) = 264.74, $p < .001$, $\eta_p^2 = .24$), parental occupation (F (1, 831) = 111.13, $p < .001$, $\eta_p^2 = .12$) and the family affluence score (F (1, 831) = 303.41, $p < .001$, $\eta_p^2 = .28$). Main effects of gender and site by gender interactions were not statistically significant for any of the SES scores (all $\eta_p^2 < .01$). There is a larger range for all of the SES scores from the HK site, which had scores at the lower range that did not occur in the UK data (see Table 3). Some parents from HK reported that their highest educational qualification was equivalent to elementary school (score = 1) or early secondary school (score = 2). The first group are likely immigrants from mainland China, the second group likely reflect that free, compulsory education beyond secondary school was only recently introduced in Hong Kong in recent decades. Importantly for this study, the patterns of SES scores for males and females were consistent within each site.

Table 2.

Means and standard deviations on the raw data (unstandardized) for key variables for females, males and the overall sample from each site before adjusting for age.

	HK			***UK***		
	All	***Female***	***Male***	***All***	***Female***	***Male***
N	371	174	197	464	231	233
WRAT Raw Score (max = 55)	45.50 (3.85)	45.43 (3.70)	45.56 (3.99)	34.53 (5.81)	34.08 (5.36)	35.07 (6.20)
WRAT Standardized Score	128.80 (11.31)	128.63 (9.99)	128.95 (12.39)	99.62 (16.72)	98.79 (15.40)	100.45 (17.94)
WRAT Grade Equivalent	12.57 (1.83)	12.53 (1.71)	12.61 (1.93)	6.52 (3.10)	6.16 (2.70)	6.88 (3.41)
EF: Inhibition	145.54 (54.57)	149.98 (56.98)	141.62 (52.18)	109.44 (43.37)	109.49 (42.15)	109.39 (44.63)
EF: Cognitive Flexibility	113.91 (27.31)	116.86 (27.20)	111.30 (27.20)	95.84 (26.27)	97.27 (25.53)	94.42 (26.67)
EF: Working Memory	3.71 (1.27)	3.75 (1.08)	3.68 (1.41)	2.93 (1.01)	2.89 (0.99)	2.96 (1.02)
EF: Planning	1.44 (0.50)	1.44 (0.48)	1.44 (0.52)	1.23 (0.45)	1.20 (0.44)	1.27 (0.45)
SES: Parental Education	4.65 (1.18)	4.68 (1.25)	4.63 (1.12)	5.73 (0.71)	5.77 (0.76)	5.69 (0.65)
SES: Parental Occupation	5.93 (1.64)	5.85 (1.69)	6.00 (1.60)	6.95 (1.16)	6.96 (1.10)	6.94 (1.22)
SES: Family Affluence Score	9.59 (3.56)	9.74 (3.58)	9.45 (3.55)	13.61 (3.08)	13.45 (3.03)	13.77 (3.12)

Note – The WRAT (Wide Range Achievement Test) generated three different scores. All scores are reported here to provide information on the sample, but only the standardized score was used in the analyses. Parental Education and Occupation scores were based on a mean of scores from parents. EF data are efficiency scores with the raw scores reported here. These scores were then age-adjusted, standardized scores in the models.

Table 3.

Frequency of the SES variable scores

	HK			***UK***		
	All	***Female***	***Male***	***All***	***Female***	***Male***
Total *N*	371	174	197	464	231	233
<u>Mean Parental Education Score</u>						
1	13	6	7	0	0	0
2	9	6	3	0	0	0
3	26	10	16	4	1	5
4	164	73	91	30	13	17
5	88	41	47	275	135	140
6	63	32	31	118	56	62
7	8	6	2	36	23	13
<u>Mean Parental Occupation Score</u>						
1	5	2	3	0	0	0
2	15	11	4	2	1	1
3	19	10	9	9	4	5
4	33	10	23	15	6	9
5	90	39	51	14	6	8
6	89	46	43	180	90	90
7	72	35	37	153	82	71
8	37	18	19	60	30	30
9	11	3	8	31	12	19
<u>Family Affluence Score</u>						
1 to 2	4	0	4	0	0	0
3 to 4	22	11	11	2	2	0
5 to 6	57	27	30	4	2	2
7 to 8	78	33	45	33	17	16
9 to 10	71	31	40	47	24	23
11 to 12	62	34	28	73	34	39
13 to 14	46	22	24	120	66	54
15 to 16	20	12	8	111	58	53
17 to 18	11	4	7	74	28	46

Notes – This count includes mean scores from .00 to .99, so for 1 that would be 1.00 to 1.99

Main Analysis

Correlations between measures are not consistent across site and gender. Correlations between numeracy skills and the other measures are less frequently statistically significant for the HK data than the UK data (see Table 4; all correlations control for age). The inter-correlations for SES measures were largely consistent for HK males and females. but less so for UK males and females where only parental education and occupation was significant for both. Inter-correlations amongst EF measures were consistent for UK males and females, but less so for HK males and females.

Links among EF, numeracy and SES differ by site for males, but not females. As shown in Figure 2 and Table 5, the most parsimonious model was different for UK males compared to HK females, HK males, and UK females. The most parsimonious UK male model was the full model tested, with EF mediating the link between SES and numeracy (*Sobel test* = 2.48, *SE* = 0.08, *p* = .01. In contrast, EF and SES were independent predictors of numeracy for females from both sites and HK males.

Table 4.

Correlations (controlling for age) of all measured variables by site and gender (males in black, females in gray).

	1	2	3	4	5	6	7	8
HK								
1. Numeracy		**.31*****	**.39*****	.08	.00	.12	.21	.15
2. SES: Parental Education	**.21**		**.59*****	**.31*****	.04	.07	.16	-.03
3. SES: Parental Occupation	**.28****	**.52*****		**.29****	-.05	.06	.06	-.01
4. SES: Family Affluence	.05	**.24****	**.29*****		.14	.16	.10	.17
5. EF: Inhibition	.08	.03	-.04	-.01		**.31*****	.07	.12
6. EF: Cognitive Flexibility	.21	.04	.01	-.06	**.26****		**.24***	.00
7. EF: Working Memory	**.37*****	-.10	-.00	.02	**.24***	.20		.16
8. EF: Planning	.20	.00	.05	.08	.08	**.23***	.16	
UK								
1. Numeracy		**.29*****	**.27*****	.18	.17	**.24****	.20	.04
2. SES: Parental Education	**.39*****		**.47*****	.07	.10	.09	.04	-.02
3. SES: Parental Occupation	**.34*****	**.61*****		.19	-.10	-.02	.00	.02
4. SES: Family Affluence	.10	.19	**.23****		.04	.14	.00	.04
5. EF: Inhibition	**.26*****	**.26*****	**.20***	-.01		**.38*****	.19	.08
6. EF: Cognitive Flexibility	**.40*****	**.23****	.17	.07	**.43*****		**.28*****	**.22***
7. EF: Working Memory	**.34*****	.11	.14	.05	**.28*****	**.42*****		**.20***
8. EF: Planning	**.24****	.03	.07	.04	**.18**	**.27*****	**.41*****	

Notes - bolded items are statistically significant, * p < .05, ** p < .01, *** p < .001; Holm correction used to account for multiple testing.

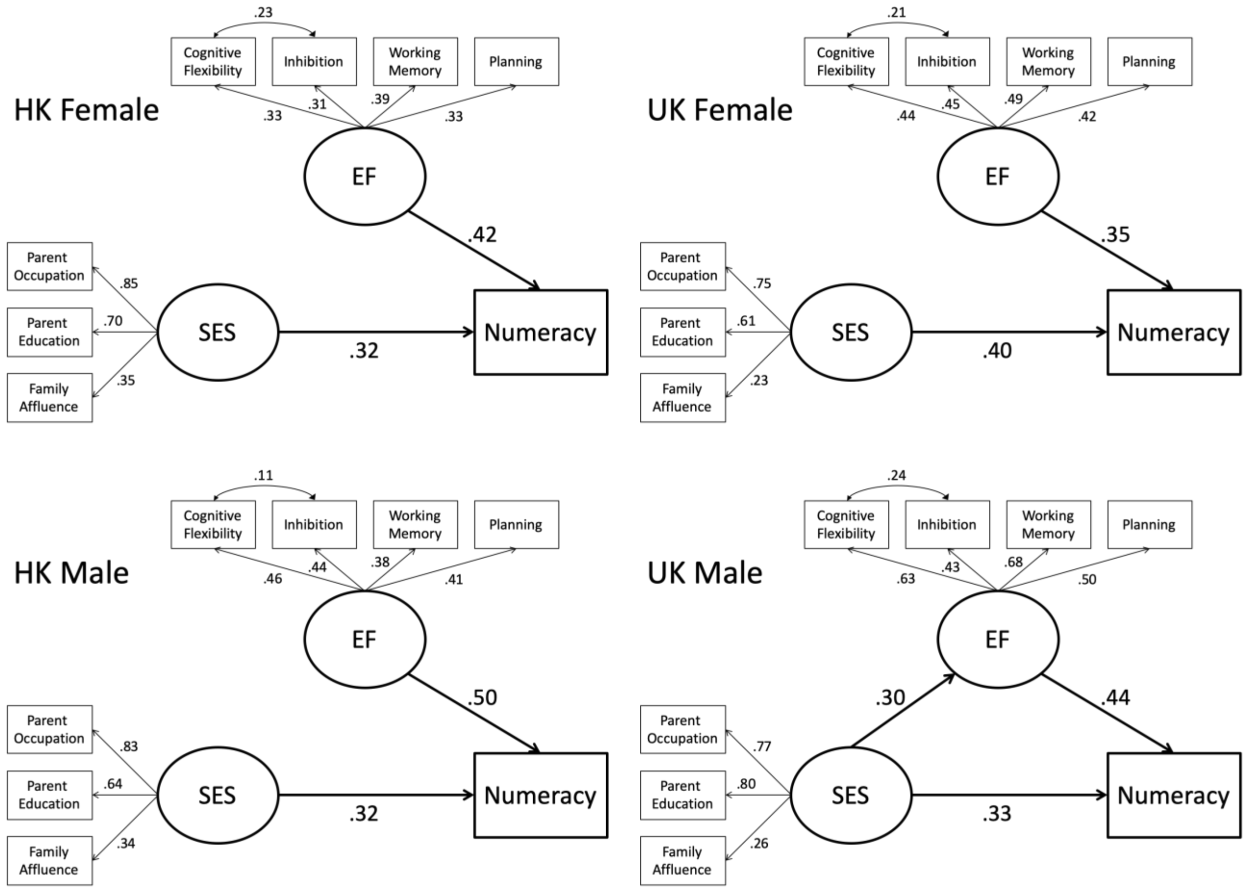

Figure 2. The most parsimonious models resulting for males and females in HK and UK for the main analysis, with the standardized parameter estimates.

Table 5.

Main Analysis: Full and nested models for each site and gender.

	HK Female	HK Male	UK Female	UK Male
Model 1: Full model				
χ^2 (17)	28.15, p = .04	18.82, p = .34	25.10, p = .09	21.53, p =.20
Model 2: Tau equivalence for SES				
χ^2 (19)	59.95, p < .001	43.33, p = .001	43.10, p = .001	52.34, p < .001
$\Delta\chi^2$ (2)	20.84, p < .001	15.63, p < .001	13.7, p = .001	29.24, p < .001
Model 3: Tau equivalence for EF				
χ^2 (20)	33.29, p = .03	30.36, p = .06	32.22, p = .04	33.38, p =.03
$\Delta\chi^2$ (3)	4.81, p = .19	8.65, p = .03	6.81, p = .08	14.07, p = .003
Model 4: Remove correlated residual				
χ^2 (18)	37.75, p = .004	24.20, p = .15	27.06, p = .08	28.23 p = .06
$\Delta\chi^2$ (1)	6.89, p = .01	6.61, p = .01	2.70, p = .10	8.12, p = .004
Model 5: Remove SES → EF				
χ^2 (18)	30.62, p = .03	18.53, p = .40	25.25, p = .12	31.59, p = .03
$\Delta\chi^2$ (1)	2.05, p = .15	0.03, p = .86	0.19, p = .66	8.54, p = .003
Model 6: Tau equivalence for EF, remove SES → EF				
χ^2 (21)	34.74, p = .03	30.37, p = .09	32.28, p = .06	45.15, p = .002
$\Delta\chi^2$ (4)	6.59, p = .16	9.00, p = .06	7.21, p = .12	25.38, p < .001
Model 7: Tau equivalence for EF, remove SES → EF and SES → Numeracy				
χ^2 (22)	60.54, p < .001	47.79, p = .001	57.91, p < .001	75.40, p < .001
$\Delta\chi^2$ (5)	27.49, p < .001	23.88, p < .001	26.33, p < .001	64.99, p < .001
Model 8: Tau equivalence for EF, remove correlated residual and SES → EF				
χ^2 (23)	68.18, p < .001	31.90, p = .08	39.92, p = .01	53.44, p < .001
$\Delta\chi^2$ (5)	34.30, p < .001	11.10, p = .049	14.34, p = .01	31.56, p < .001

Notes – all χ^2values do not use robust methods to allow for $\Delta\chi^2$ comparisons; all model comparisons ($\Delta\chi^2$) were made with Model 1.

Table 6.

Main Analysis: Goodness of Fit Indices for each site and gender.

	HK Female	HK Male	UK Female	UK Male
Preferred Model	6	6	6	1
χ^2	34.74	30.37	32.28	21.53
CFI	.92	.94	.94	.99
TLI	.89	.92	.92	.98
AIC	3618	4264	4217	4326
BIC	3618	4265	4222	4330
RMSEA	.06	.05	.05	.03
SRMR	.07	.06	.06	.05

Notes - CFI: Comparative Fit Indie, TLI: Tucker-Lewis Index; AIC: Akaike; BIC: sample-size adjusted Bayesian; RMSEA: Root Mean Square Error of Approximation; SRMR: Standardized Root Mean Square Residual.

Follow Up Analysis

The main analyses were a test of the cultural universality of Lawson and Farah (2017). After running these analyses, we were encouraged to consider whether general cognitive ability (sometimes referred to as intelligence and not included in Lawson & Farah's models) played a role in our findings. This suggestion was motivated by the long-standing debate about the distinctive role that EF might play above and beyond general cognitive ability (e.g., Blair, 2006; Engelhardt, et al., 2016; Friedman et al., 2006; Royall & Palmer, 2014) and the long history of studying effects of SES on intelligence or the effects and interactions that both can have on academic achievement (e.g., Strenze, 2007; Tucker-Drob & Bates, 2016; von Stumm, 2017; von

Stumm & Plomin, 2015). Furthermore, Corso and colleagues (2016) found that executive functions play an important role when looking at SES influences on intelligence and reading comprehension with a sample of 110 children from Brazil (ages 9 to 12 years). We could not find another study that explored these mediation effects with numeracy skills.

As part of the larger study, we collected general cognitive ability using the Raven's Standard Progressive Matrices (Raven, 1998; 2000) and conducted follow-up analyses to test whether our initial results hold when adding in general cognitive ability (see Figure 3). Five UK participants (3 males) had missing Ravens data, they were excluded from the follow-up analyses. As with the main analyses, correlations and models use only age-adjusted scores.

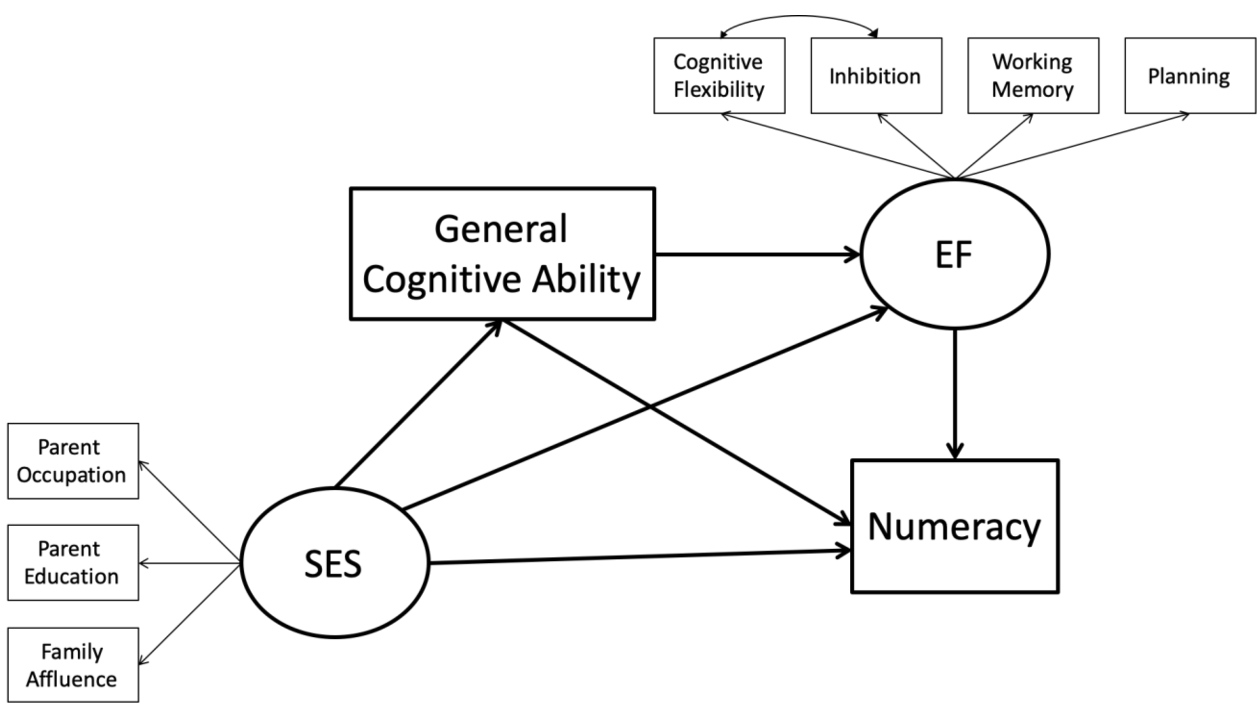

Figure 3. The theoretical model tested across the four groups (UK male, UK female, HK male, HK female) as part of the follow-up analysis. The correlated residual was included based on preliminary analyses of the data. All other connections were determined a priori.

General Cognitive Ability scores differ by site and gender. As indicated in Table 7, there were small main effects of site (favouring children from HK) on the Ravens raw scores (F (1, 826) = 48.32, p < .001, η_p^2 = .06), percentile scores (F (1, 826) = 18.86, p < .001, η_p^2 = .02) and age equivalent scores (F (1, 826) = 35.96, p < .001, η_p^2 = .04). In addition, there were small main effects of gender (favouring females) on the Ravens raw scores (F (1, 826) = 13.88, p < .001, η_p^2 = .02), percentile scores (F (1, 826) = 12.07, p = .001, η_p^2 = .01) and age equivalent scores (F (1, 826) = 14.57, p < .001, η_p^2 = .02). The gender by site interaction was not statistically significant for any of the Ravens variables (all η_p^2 < .005).

General Cognitive Ability correlations differ by site and gender. Overall, there were fewer statistically significant correlations for the participants from HK compared to UK (see Table 7, correlations controlled for age). Ravens was correlated with numeracy for all groups (r = .26 to .61) and working memory for both UK genders and HK males (r = .27 to .36).

Again, links among EF, numeracy and SES differ by site. After adding general cognitive ability UK females now look like UK males (see Figure 4, Tables 8 and 9). The key variation between the follow-up and main analysis is that the direct link between SES and EF has been replaced by an indirect link through general cognitive ability. Again, HK males and females were very similar to each other and differed from UK males and females. In addition, SES predicts general cognitive ability for UK males and females, where it does not for HK males and females. Although general cognitive ability does play an important role in the follow-up models, strong EF-numeracy links are still present for all groups.

Table 7.

Means (standard deviations) and correlations by site and gender for the general cognitive ability and the main analysis measures.

	HK			***UK***		
	All	***Female***	***Male***	***All***	***Female***	***Male***
N	371	174	197	459	229	230
Means (standard deviation)						
RAVENS Raw Score (max = 60)	44.48 (7.97)	46.05 (6.72)	43.10 (8.72)	40.46 (8.96)	39.73 (9.59)	41.19 (8.24)
RAVENS Percentile Score	52.92 (30.64)	58.43 (29.03)	48.06 (31.27)	43.87 (31.41)	41.57 (32.22)	46.19 (30.47)
RAVENS Age Equivalent	18.46 (3.30)	19.03 (2.50)	17.95 (3.80)	16.78 (4.67)	16.23 (5.00)	17.33 (4.25)
Correlation with other measures						
Numeracy		**.26***	**.39*** **		**.52*** **	**.61*** **
SES: Parental Education		.10	.01		**.21***	**.36*** **
SES: Parental Occupation		.15	-.01		.10	**.31*** **
SES: Family Affluence		.04	-.05		.19	.04
EF: Inhibition		.06	.20		.19	**.24** **
EF: Cognitive Flexibility		.15	.14		**.29*** **	**.39*** **
EF: Working Memory		.21	**.36*** **		**.27*** **	**.27*** **
EF: Planning		-.14	.20		.13	.17

Notes – correlations use age-adjusted data (percentile score from Ravens, standardized score from WRAT, age-adjusted EF and SES scores); bolded items are statistically significant, * $p < .05$, ** $p < .01$, *** $p < .001$

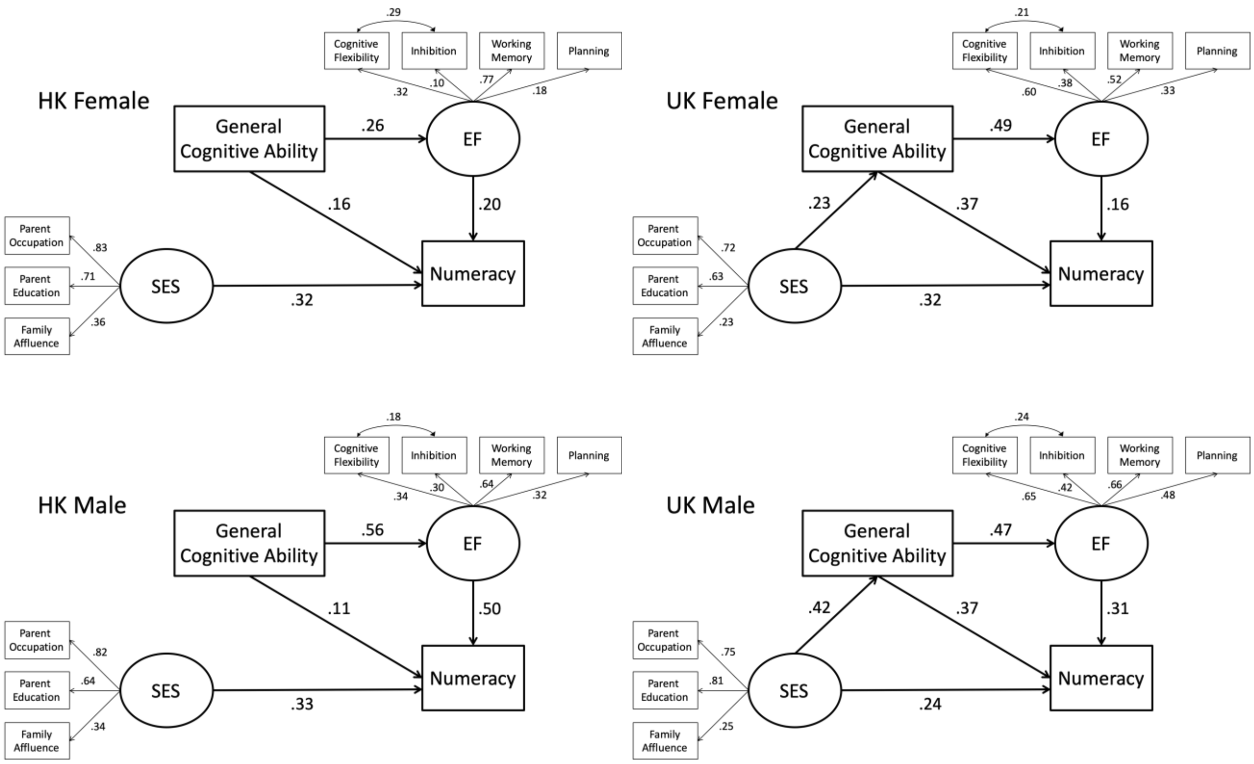

Figure 4. The most parsimonious models resulting for males and females in HK and UK for the follow-up analysis, with the standardized parameter estimates.

Table 8.

Follow-up Analysis: Full and nested models for each site and gender.

	HK Female	HK Male	UK Female	UK Male
Model 1: Full Model				
χ^2 (22)	39.09, *p* =.01	22.32, *p* = .44	30.60, *p* = .11	25.26, *p* = .29
Model 2: Tau Equivalence for SES				
χ^2 (24)	71.14, *p* < .001	47.01, *p* = .003	47.58, *p* = .003	56.59, *p* < .001
$\Delta\chi^2$ (2)	21.06, *p* < .001	15.99, *p* < .001	12.68, *p* = .002	32.17, *p* < .001
Model 3: Tau Equivalence for EF				
χ^2 (25)	47.12, *p* =.01	36.79, *p* = .06	38.53, *p* = .04	37.97, *p* = .047
$\Delta\chi^2$ (3)	7.22, *p* = .07	10.43, *p* = .02	7.43, *p* = .06	14.61, *p* = .002
Model 4: Remove Correlated Residual				
χ^2 (25)	49.00, *p* = .001	22.94, *p* = .22	33.88, *p* = .07	29.95, *p* = .15
$\Delta\chi^2$ (3)	6.27, *p* = .01	6.84, *p* = .01	4.05, *p* = .04	5.33, *p* = .02
Model 5: Remove SES → EF				
χ^2 (23)	40.56, *p* = .01	22.35, *p* = .50	30.90, *p* = .13	27.90, *p* = .27
$\Delta\chi^2$ (1)	1.31, *p* = .25	0.02, *p* = .88	0.45, *p* = .50	2.55, *p* = .11
Model 6: Tau Equivalence for EF, remove SES → EF				
χ^2 (26)	48.07, *p* = .01	36.81, *p* = .08	38.97, *p* = .049	41.23, *p* = .03
$\Delta\chi^2$ (4)	8.77, *p* = .07	11.09, *p* = .03	8.66, *p* = .07	17.75, *p* = .001
Model 7: Remove SES → EF and SES → General Cognitive Ability				
χ^2 (24)	44.26, *p* = .01	22.37, *p* = .56	37.65, *p* =.04	61.31, *p* < .001
$\Delta\chi^2$ (2)	4.10, *p* = .13	0.03, *p* = .98	7.61, *p* = .02	34.30, *p* < .001
Model 8: Tau equivalence for EF, remove SES → EF and SES → General Cognitive Ability				
χ^2 (27)	51.78, *p* = .003	36.82, *p* = .10	45.76, *p* = .01	74.63, *p* < .001
$\Delta\chi^2$ (5)	11.51 *p* = .04	11.32, *p* = .045	15.00, *p* = .01	53.15, *p* < .001
Model 9: Remove SES → EF, SES → General Cognitive Ability, SES → Numeracy and General Cognitive Ability → Numeracy				
χ^2 (26)	70.61, *p* < .001	44.04, *p* = .02	73.84, *p* < .001	104.78, *p* < .001
$\Delta\chi^2$ (4)	25.72, *p* < .001	20.02, *p* < .001	33.03, *p* < .001	94.57, *p* < .001

Table 8 Notes - all χ^2values do not use robust methods to allow for $\Delta\chi^2$ comparisons; all model comparisons ($\Delta\chi^2$) were with Model 1. Removing any other links produced significant $\Delta\chi^2$ results for all sites and genders. A model removing SES → EF, SES → General Cognitive Ability, SES → Numeracy produced similar results to Model 9. Similarly, incorporating tau equivalence for EF with Model 9 produced similar results as Model 9.

Table 9.

Theoretical Model 2: Goodness of Fit Indices (using robust methods) for each site and gender.

	HK Female	HK Male	UK Female	UK Male
Preferred Model	7	7	5	5
χ^2	44.26	22.37	30.90	27.90
CFI	.99	1.00	.97	.99
TLI	.83	1.00	.95	.98
AIC	4078	4770	4732	4827
BIC	4078	4772	4738	4834
RMSEA	.07	.00	.04	.03
SRMR	.07	.04	.05	.05

Notes: CFI: Comparative Fit Indie, TLI: Tucker-Lewis Index; AIC: Akaike; BIC: sample-size adjusted Bayesian; RMSEA: Root Mean Square Error of Approximation; SRMR: Standardized Root Mean Square Residual

Discussion

Our results contribute three key findings: (1) EF consistently predicts numeracy skills across sites and genders; (2) Associations between SES and EF differ by site and gender; and (3) SES-EF associations with numeracy skills also differ by site and gender. Below, we discuss each of these findings in turn.

EF Consistently Predicts Numeracy Skills Across Sites and Genders

Our finding of consistent links between EF and numeracy skills across both sites and genders follows the results of previous studies with older children (e.g., Lan et al., 2011; Ng et al., 2015), even after accounting for general cognitive ability. This consistency of associations between EF and numeracy skills across our two study sites and genders is striking for at least two reasons. First, HK children out-perform their UK counterparts on both EF and numeracy. Ellefson et al. (2017) reported 2-year gap difference in EF skills for this sample. Here, the grade equivalent scores indicate a 6-year gap for numeracy skills (and a 2-year-gap for general cognitive ability). The similarity in associations between individual differences in these two measures suggests that EF skills are salient for numeracy across a wide range of ability levels. Second, it is well recognized that schools in the UK and HK adopt different pedagogical strategies, (e.g., Fan et al., 2017; Forestier & Crossley, 2015). For example, HK teachers tend to adopt a more direct, whole-class instruction followed by arithmetic practice problems (Mok & Morris, 2001). These differences might be predicted to affect the strength of association between EF and numeracy. However, the data reported here suggest this link may be culturally universal.

Associations Between SES and EF Differ by Site and Gender

In contrast to Lawson et al.’s (2018) meta-analytic findings indicating a consistent association between EF and SES, our four best-fitting models indicated that this association is only statistically significant for UK males. Furthermore, that direct link appears to be mediated by general cognitive ability. Here, two points deserve note. First, with regards to the UK dataset, the finding that EF performance varies significantly with SES for males but not females is consistent with the wider view that educational effects of social disadvantage might be especially marked in males (e.g., Cobb-Clark & Moschion, 2017; Ridge et al., 2017). However, the follow up analyses indicate that both males and females from the UK show indirect effect from SES to EF through general cognitive ability. Finally, with regards to the HK sample, the lack of a significant association between SES and EF or between SES and general cognitive ability is somewhat consistent with earlier studies showing small effect size links (e.g., Philipson, 2009; Wang et al., 2016).

The vast majority of published studies (those in the meta-analysis and Lawson and Farah, 2017) use a sample recruited from the USA and those studies have not looked at whether SES effects on EF are consistent for males and females. In the current study, if we did not need to follow the results of the measurement invariance analysis and split the sites by gender, then the SES and EF links would have been statistically significant for the UK. Splitting by gender indicates that the mediation is driven by UK males and might not be universal across gender as well as across site. However, these results could have been influenced by the distribution of SES scores being slightly skewed towards the middle end. Continued efforts to run studies with participants from a diverse set of SES backgrounds is important for understanding how SES influences cognitive development.

An alternative explanation could be that the non-significant links between SES and EF for UK females and HK males and females are driven by the computerized nature of the EF tasks. Lawson and Farah’s (2017) findings of a significant link between SES and EF for a sample from USA also used computerized tasks, but Wang et al. (2016) and Phillipson (2009) used researcher-administered tasks. Lawson et al.’s (2018) meta-analysis of SES and EF includes tasks that use both researcher- and computer-administered EF tasks. A recent survey indicates that male and female adolescents use technology differently in a sample from the USA (Anderson & Jiang, 2018). Possible variation in usage patterns across site and gender could have influenced our results. Further work is needed to determine whether the way EF skills are measured influence how it links to SES.

SES-EF Associations with Numeracy Skills Differ by Site and Gender

Our models indicate that predictors of numeracy skills are consistent across sites for females, but the results for males differ across sites and with females. Specifically, SES and EF are independent predictors of numeracy skills for the HK participants and UK females. In contrast, UK male results indicated that EF mediates the link between SES and numeracy skills even when accounting for general cognitive ability. In contrast, UK females now looks like UK males, with SES influencing EF skills indirectly through general cognitive ability for both groups. Considered together, the findings suggest that the cognitive and academic skills of UK males might be most affected by SES, followed by the other groups.

As detailed elsewhere (Ellefson et al., 2017), the parents in our sample also completed a relatively lengthy set of EF tasks and so, to minimize the burden upon participants, we did not include detailed measures of the home learning environment or other SES-related factors that might have contributed to our findings. Further studies that gather richer information about SES influences on EF and numeracy skills in these or other international sites would be a valuable direction for future cross-cultural studies as they could better inform intervention programs aimed at the most vulnerable across societies.

Conclusion

Ours is the first published study to explore in a large sample whether the links between SES, EF and numeracy skills are consistent across genders in HK and the UK. Our aim was to further the discussions on the cultural universality of EF as a possible mediator of the association between SES and numeracy and to directly investigate whether these links are consistent across genders. Studies comparing additional sites will help to explore this idea further (e.g., comparing HK to mainland China). Together with previous findings, our results suggest that associations between SES, EF and numeracy skills may be culture specific and indicate that cultural insights may enable impactful shifts in public policy to narrow the achievement gap between children from affluent and disadvantaged families.

Acknowledgements

A joint-council award to the authors funded this research (ES/K010225/1: Economic and Social Research Council, Research Grants Council of HK). Thinking Games website development supported by the Institute of Educational Sciences, U.S. Department of Education, through Grant R305A110932 to the University of Cambridge. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education. Electronic access to dataset: http://reshare.ukdataservice.ac.uk/852570/.

Thanks to (1) Geoff Martin for Thinking Games website programming; (2) Annabel Amodia-Bidakowska, Jeff Chan, Emma Chatzispyridou, Yiming Han, Joyce Hoi-Ling Ng, Katherine Parkin, Annie Raff, and Irene Nga-Lam Sze for data collection assistance; (3) Rosie Blunt, Hannah Bush, Ying-Kit Chan, Claudia Chu, Shehnaz Dowlet, Ellie Frank, Anton Evans, Yanning Gu, Nelly Hu-Kwo, Rina Lai, and Tanya Paes for scoring and data entry assistance; and (4) Richard Parkin for proofreading.

References

Adams, J., & Weakleim, D. L. (2011). August B. Hollingshead's "four factor index of social status": From unpublished paper to citation classic. *Yale Journal of Sociology*, *8*, 11-19.

Ahmed, S. F., Tang, S., Waters, N. E., & Davis-Keen, P. (2018). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*, Advance online publication. DOI: dx.doi.org/10.1037/edu0000296

Anderson, M., & Jiang, J. (2019) *Teens social media & technology*. Technical Report for Pew Research Center. Downloaded from https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/

Alves, A. F., Martins, A., & Almeida, L. S., (2016). Interactions between sex, socioeconomic level, and children's cognitive performance. *Psychological Reports*, *118*, 471-486. DOI: 10.1177/0033294116639428

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*, 327-336. DOI:10.1016/j.lindif.2011.01.007

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, *29*, 109-160. DOI: 10.1017/S0140525X06319036

Blair, C., Granger, D. A., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., . . . the FLP Investigators. (2011). Salivary cortisol mediates effects of poverty and parenting on executive functions in early childhood. *Child Development*, *82*, 1970-1984. DOI:10.1111/j.1467-8624.2011.01643.x

Blair, C., & Raver, C. C. (2016). Poverty, stress, and brain development: New directions for prevention and intervention. *Academic Pediatrics*, *16*, S30-S36. DOI: 10.1016/j.acap.2016.01.010

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647-663. DOI: 10.1111/j.1467-8624.2007.01019.x

Blums, A., Belsky, J., Grimm, K., & Chen, Z. (2017). Building links between early socioeconomic status, cognitive ability, and math and science achievement. *Journal of Cognition and Development*, *18*, 16-40. doi: 10.1080/15248372.2016.1228652

Boudreau, B., & Poulin, C. (2009). An examination of the validity of the Family Affluence Scale II (FAS II) in a general adolescent population of Canada. *Social Indicators Research*, *94*, 29-42. http://doi.org/10.1007/s11205-008-9334-4

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*, 205-228. DOI: 10.1080/87565640801982312

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*, 36-41. DOI: 10.1111/cdep.12059

Clark, C. A., Sheffield, T. D., Wiebe, S., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development*, *84*, 662-677. DOI: 10.1111/J.1467-8624.2012.01854.X

Cobb-Clark, D. A., & Moschion, J. (2017). Gender gaps in educational achievement*. Journal of Popular Economics*, *30*, 1093-1134. DOI: 10.1007/s00148-017-0638-z

Cole, D. A., Ceisla, J. A., & Steiger, J H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, *12*, 381-398. DOI: 10.1037/1082-989X.12.4.381

Corsi, P.M. (1972). *Memory and the medial temporal area of the brain*. Doctoral Thesis: McGill University.

Corso, H. B., Cromley, J. G., Sperb, T., & Fumagalli Salles, J. (2016). Modeling the relationship among reading comprehension, intelligence, socioeconomic status, and neuropsychological functions: The mediating role of executive functions. *Psychology & Neuroscience, 9*, 32-45. DOI: 10.1037/pne0000036

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*, 63-68. DOI: 10.1016/j.tine.2013.12.001

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, *162*, 12-26. DOI: 10.1016/j.cognition.2017.01.014

Dell, C. A., Harrold, B., & Dell, T. (2008). Test review: Wide Range Achievement Test (4th ed.). *Rehabilitation Counseling Bulletin*, *52*, 57-60. DOI: 10.1177/0034355208320076.

Devine, R. T., Bignardi, G., & Hughes, C. (2016). Executive function mediates the relations between parental behaviors and children's early academic ability. *Frontiers in Psychology*, *7*, 1902. DOI: 10.3389/fpsyg.2016.01902

Duncan, R. J., McClelland, M. M., & Acock, A. C. (2017). Relations between executive function, behavioral regulation, and achievement: Moderation by family income. *Journal of Applied Developmental Psychology*, *49*, 21-30. DOI: 10.1016/j.appdev.2017.01.004

Ellefson, M.R., Ng, F.F., Wang, Q., & Hughes, C. (2017). Efficiency of executive function: A two-generation cross-cultural comparison of samples from Hong Kong and the United Kingdom. *Psychological Science*, *28*, 555-566. DOI: 10.1177/0956797616687812

Ellefson, M., Shapiro, L., & Chater, N. (2006). Asymmetrical switch costs in children. *Cognitive Development*, *21*, 108-130. DOI: 10.1016/j.cogdev.2006.01.002

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

Engelhardt, L. E., Mann, F. D., Briley, D. A., Church, J. A., Harden, K. P., & Tucker-Drob, E. M. (2016). Strong genetic overlap between executive functions and intelligence. *Journal of Experimental Psychology: General*. *145*, 1141-1159. DOI: 10.1037/xge0000195

Fan, H., Xu, J., Cai, Z., He., J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year meta-analysis, 1986-2015. *Educational Research Review*, *20*, 35-54. DOI: 10.1016/j.edurev.2016.11.003

Farah, M. J. (2018). Socioeconomic status and the brain: prospects for neuroscience-informed policy. *Nature Reviews Neuroscience*, *19*, 428-438. DOI: 10.1038/s41583-018-0023-2

Forestier, K., & Crossley, M. (2015). International education policy transfer – borrowing both ways: the Hong Kong and England experience. *Compare: A Journal of Comparative and International Education*, *45*, 664-685. DOI: 10.1080/03057925.2014.928508

Fox. J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage, Thousand Oaks, CA. http://tinyurl.com/carbook

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172-179. DOI: 10.1111/j.1467-9280.2006.01681.x

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29-44. DOI:10.1016/j.edurev.2013.05.003

Fuhs, M. W., Nesbitt, K. T., Farran, D. C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, *50*, 1698-1709. DOI: 10.1037/a0036633

Grissom, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: Evaluating gender and sex differences in executive function. *Neuropsychopharmacology Reviews*, *44*, 86-96. DOI: 10.1038/s41386-018-0179-5

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, *37*, 161-168. DOI: 10.1016/j.lindif.2015.01.008

Guo, X., Lv, B., Zhou, H., Liu, C., Liu, J., Jiang, K., & Luo, L., (2018). Gender differences in how family income and parental education related to reading achievement in china: The mediating role of parental expectation and parental involvement. *Frontiers in Psychology*, *9*, 783. DOI: 10.3389/fpsyg.2018.00783

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished working paper, Department of Sociology, Yale University. Later published in *Yale Journal of Sociology*, *8*, 21-51.

Howard, S. J., Cook, C. J., Everts, L., Melhuish, E., Scerif, G., Norris, S., Twine, R., Kahn, K., & Draper, C. E. (2019). Challenging socioeconomic status: A cross-cultural comparison of early executive function. *Developmental Science*, e12854. DOI: 10.1111/desc.12854

Hughes, C. & Devine, R. T. (2019). For better or for worse? Positive and negative parental influences on young children's executive function. *Child Development*, *90*, 593-609. DOI: 10.1111/cdev.12915

Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair? *Developmental Neuropsychology*, *28*, 645-668. DOI: 10.1207/s15326942dn2802_5

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2010). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, *35*, 20-36. DOI: 10.1080/87565640903325691

Hyde, J. S. (2016). Sex and cognition: gender and cognitive functions. *Current Opinion in Neurobiology*, *38*, 53-56. DOI: 10.1016/j.conb.2016.02.007

Imada, T., Carlson, S.M., & Itakura, S. (2013). East–West cultural differences in context-sensitivity are evident in early childhood. *Developmental Science*, *16*, 198-208. DOI: 10.1111/desc.12016

Johnson, S. B., Riss, J. L., & Noble, K. G. (2016). State of the art review: Poverty and the developing brain. *Pediatrics*, *137*, e20153075. DOI: 10.1542/peds.2015-3075

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). *semTools: Useful tools for structural equation modeling. R package version 0.5-1*. Retrieved from https://CRAN.R-project.org/package=semTools

Kalaycıoğlua, D. B. (2015). The influence of socioeconomic status, self-efficacy, and anxiety on mathematics achievement in England, Greece, Hong Kong, the Netherlands, Turkey and the USA. *Educational Sciences Theory & Practice*, *15*, 1391-1401. DOI: 10.12738/estp.2015.5.2731

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*, 677-692. DOI: 10.1016/j.jecp.2010.11.001

Lawson, G. M., & Farah, M. J (2017). Executive function as a mediator between SES and academic achievement throughout childhood. *International Journal of Behavioural Development*, *41*, 94-104. DOI: 10.1177/0165025415603489

Lawson, G. M., Hook, C. J.. & Farah, M. J (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*, e12529. DOI: 10.1111/desc.12529

Lee, K., & Bull, R. (2016). Developmental Changes in Working Memory, Updating, and Math Achievement. *Journal of Educational Psychology*, *108*, 869-882. DOI: 10.1037/edu0000090

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101*, 373-387. DOI:10.1037/a0013843

Lewis, C., Koyasu, M., Oh, S., Ogawa, A., Short, B., & Huang, Z. (2009). Culture, executive function, and social understanding. In C. Lewis & J. I. M. Carpendale (Eds.), Social interaction and the development of executive function. *New Directions in Child and Adolescent Development*, *123*, 69-85. DOI: 10.1002/cd.236

Liu, X., & Lu, K. (2008). Student performance and family socioeconomic status: Results from a survey of compulsory education in western China. *Chinese Education & Society*, *41*, 70-83. DOI: 10.2753/CED1061-1932410505

Logan, G. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego, CA US: Academic Press.

Mok, I. A., C., & Morris, P. (2001). The metamorphosis of the &virtuoso': pedagogic patterns in Hong Kong primary mathematics classrooms. *Teacher and Teacher Education*, *17*, 455-468. DOI: 10.1016/S0742-051X(01)00006-3

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/

Navarro, D. J. (2015). *Learning statistics with R: A tutorial for psychology students and other beginners* (Version 0.5). University of Adelaide: Adelaide, Australia.

Nesbitt, K. T., Baker-Ward, L., & Willoughby, M. T., (2013). Executive function mediates socio-economic and racial differences in early academic achievement. *Early Childhood Research Quarterly*, *28*, 774-783. DOI: 10.1016/j.ecresq.2013.07.005

Ng, F. F.-Y., Tamis-Lemonda, C., Yoshikawa, H., & Sze, I. N. L. (2015). Inhibitory control in preschool predicts early math skills in first grade: Evidence from an ethnically diverse sample. *International Journal of Behavioral Development*, *39*, 139-149. DOI: 10.1177/0165025414538558

OECD. (2016). *PISA 2015 results (volume 1): Excellents and equity in education*, PISA, OECD Publishing, Paris. DOI: 10.1787/9789264266490-en

Phillipson, S. (2009). Context of academic achievement: lessons from Hong Kong. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *29*, 447-468. DOI: 10.1080/01443410903059024

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*, 1-48. DOI: 10.1006/cogp.1999.0735

Raven, J. C., Styles, I., & Raven, M. A. (1998). *Raven's Progressive Matrices: SPM plus test booklet*. Oxford, England: Oxford Psychologists Press.

Revelle, W. (2018). *psych: Procedures for personality and psychological research* (version = 1.8.12). Northwestern University, Evanston, IL. https://CRAN.R-project.org/package=psych

Rhoades, B. L., Greenberg, M. T., Lanza, S. T., & Blair, C. (2011). Demographic and familial predictors of early executive function development: Contribution of a person-centered perspective. *Journal of Experimental Child Psychology*, *108*, 638-662. DOI: 10.1016/j.jecp.2010.08.004

Ridge, N., Kippels, S., & Chung, B. J. (2017). The challenges and implications of a global decline in the educational attainment and retention of boys. WISE:World Education Summit for Education. Retrived from: https://www.wise-qatar.org/sites/default/files/rr.2.2017_qasimi.pdf.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 207-231.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1-36. URL http://www.jstatsoft.org/v48/i02/

Royall, D. R., & Palmer, R. F. (2014). “Executive functions” cannot be distinguished from general intelligence: two variations on a single theme within a symphony of latent variance. *Frontiers in Behavioral Neuroscience*, *8*, 369. DOI: 10.3389/fnbeh.2014.00369

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*, 74-81. DOI: 10.1111/j.1467-9280.2005.01667.x

Samuels, W. E., Tournaki, N., Blackman, S., & Zilinski, C. (2016). Executive functioning predicts academic achievement in middle school: A four-year longitudinal study. *The Journal of Educational Research*, *109*, 478-490. DOI: 10.1080/00220671.2014.979913

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinsh, S., & Boyce, W. T. (2011). Family socioeconomic status and child executive functions: The roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*, *17*, 120-132. DOI:10.1017/S1355617710001335

Shen, C., Vasilyeva, M., & Laski, E. V. (2016). Here, but not there: Cross-national variability of gender effects in arithmetic. *Journal of Experimental Child Psychology*, *146*, 50-65. DOI: 10.1016/j.jecp.2016.01.016

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*, 417-453. DOI: 10.3102/00346543075003417

Snelbaker A. J., Wilkinson G. S., Robertson G. J., & Glutting J. J. (2001) Wide Range Achievement Test 3 (WRAT3). In: W. I.Dorfman, & M. Hersen (Eds.) *Understanding psychological assessment: Perspectives on individual differences*. (pp. 259-274) Springer, Boston, MA.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*, 745-759. DOI: 10.1080/17470210500162854

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, *35*, 401-426. DOI: 10.1016/j.intell.2006.09.004

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science*, *27*, 138-149. DOI: 10.1177/0956797615612727

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*, 427-449. DOI:10.1016/j.intell. 2006.09.001

Viterbori, P., Usai, M. C., Traverso, L., & De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, *140*, 38-55. DOI: 10.1016/j.jecp.2015.06.014

von Stumm, S. (2017). Socioeconomic status amplifies the achievement gap throughout compulsory education independent of intelligence. *Intelligence*, *60*, 57-62. DOI: 10.1016/j.intell.2016.11.006

von Stumm, S., & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, *48*, 30-36. DOI: 10.1016/j.intell.2014.10.002

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*, 1174-1204. DOI: 10.1037/a0036620

Wang, Z., Devine, R. T., Wong, K. K., & Hughes, C. (2016). Theory of mind and executive function in middle childhood across cultures. *Journal of Experimental Child Psychology*, *149*, 6-22. DOI: 10.1016/j.jecp.2015.09.028

Wardle, J., Robb, K., & Johnson, F. (2002). Assessing socioeconomic status in adolescents: the validity of a home affluence scale. *Journal of Epidemiology and Community Health*, *56*, 595-599. DOI: /10.1136/jech.56.8.595

Wei, W., Guo, L., Georgiou, G. K., Tavouktsoglou, A., & Deng, C. (2018). Different subcomponents of executive functioning predict different growth parameters in mathematics: Evidence from a 4-year longitudinal study with Chinese children. *Frontiers in Psychology*, *9*, 1037. DOI: 10.3389/fpsyg.2018.01037

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*, 45-53. DOI: 10.1037/a0016738

Welsh, M. (1991). Rule-guided behavior and self-monitoring on the tower of Hanoi disk-transfer task. *Cognitive Development*, *6*, 59-76.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575-587. DOI: 10.1037/0012-1649.44.2.575

Wilkinson, G. S. (1993). *Wide Range Achievement Test 3*. Wilmington, DE: Wide Range, Inc.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test* (4th ed.). Lutz, FL: Psychological Assessment Resources