



**About Those Baby Brainwaves: Why “Policy Relevant”
Social Science is Mostly a Fraud**

Jordan Lasker*

Center for the Study of Partisanship and Ideology

Executive Summary

4/4/2022

To read the full report, [click here](#).

Code is available for all original analyses.¹

* PhD Student at Texas Tech University

CSPI

Table of Contents

Table of Contents	2
Summary.....	3
Introduction.....	3
Problems With the Literature Linking EEGs to Psychological Outcomes	4
Multiple Comparisons	4
Small Samples and Low Power.....	5
Unrepresentative Samples	9
Problems With the Study Itself.....	9
Gaming Preregistration	9
Similar Magnitudes Don't Matter	10
Long-Term Treatment Effects and Parental Responsibility	10
Most Interventions Have Small to Nil Effects	11
Policy Conclusions Were Unwarranted	13
Conclusion	14

CSPI

Summary

1. A recent study claimed small cash transfers to the parents of newborns improved their babies' brain activity. The study was lauded in the media and by D.C. policymakers who argued its results supported redistributive policies, most notably the child tax credit.
2. Closer inspection of the baby brainwaves study reveals its claims to be wildly overstated and its methodology to be suspect, as is common for policy relevant social science.
3. Authors Troller-Renfree et al. committed numerous bad research practices, including:
 - Deviating from their analysis plan, and justifying it by referencing studies that did not support and even contradicted their results.
 - Highlighting results obtained through typically unimportant methodological decisions that weren't preregistered.
 - Ignoring that larger and more prolonged interventions typically generate smaller effects than they found, meaning their results were likely due to chance or error.
4. Looking beyond the study in question, the theoretical and empirical basis it claims to build upon is largely a house of sand. The evidence that is purported to show a link between brain waves and cognitive outcomes is extremely weak, and most interventions seeking to improve cognitive ability in children have small to nil effects.
5. Despite its inconclusiveness, the study was portrayed as relevant to child tax credit policy, vigorously promoted in media outlets like *Vox* and *The New York Times*, and championed by think tanks and policymakers.
6. These issues are not unique to the baby brainwaves study. Social scientists frequently engage in questionable research practices and exaggerate the strength and implications of their findings. Policymakers capitalize on low-quality social science research to justify their own agendas. This is particularly true when it comes to research that claims to improve cognitive and behavioral outcomes.
7. Thus, we ought to be skeptical of social scientists' ability to reliably inform public policy, and of policymakers' ability to evaluate social-scientific research objectively, particularly when it is "policy relevant."

Introduction

On January 24th, 2022, *The New York Times* published a piece that went out as a news alert titled "Cash Aid to Poor Mothers Increases Brain Activity in Babies, Study Finds."² The article discussed Troller-Renfree et al. (2022), a study published in the *Proceedings of the National Academy of Science* (PNAS) that purported to show a monthly subsidy of \$333 yielded considerable improvements in children's brainwave activity.³ *Vox* also covered the story, opining that "we can be reasonably confident the cash [parents received] is a primary cause of these changes in babies' brains. And we can be reasonably confident it will be a causative factor in whatever future outcomes the... researchers find."⁴ They weren't alone: think tanks, medical news aggregators, and other mainstream outlets like *NBC* and *Forbes* all praised the study and the alleged psychological benefits of brainwave improvements.⁵

Laudatory coverage was accompanied by claims that the study had special relevance to expanding child tax credits. The study authors even put out a pro-child tax credit press release on their website: "This study's findings on infant brain activity... really speak to how anti-poverty policies – including the types of expanded child tax credits being debated in the U.S. – can and should be viewed as investments in children."⁶

High praise and lofty implications aside, a closer examination of the study reveals its results to be dubious and its methods severely flawed. While the authors' public statements portrayed the study as strong, reliable, and policy relevant, none of these claims hold up to scrutiny. The study and its portrayal stand as excellent examples of what can go wrong when political desire supersedes scientific reasoning, and why we ought to be skeptical of social scientists' ability to reliably inform public policy with their research. A deeper dive into the problems with this study and the literatures it purports to build upon are therefore warranted.

Several academics and public intellectuals pointed out flaws in Troller-Renfree et al.'s (2022) study shortly after it was released. Psychologist Stuart Ritchie noted the article's potentially-dubious peer review status, the disconnect between observations made by the study's authors and outcomes that matter, the potential for fade-out effects, deviations from preregistered methodology, and the nonsignificance of the reported effects of cash transfers on any of the study's main outcomes.⁷ Statistician Andrew Gelman found that the significance of results was not robust and that the graphs produced by the authors could be replicated even if the children in the sample were randomized into artificial treatment and control groups.⁸ Finally, psychiatrist Scott Alexander offered further commentary on the article to the effect that the results seemed unlikely for various reasons, alongside a summary of what other people had said.⁹

In this article, I show that Troller-Renfree et al.'s results were even less robust than critical commentators said they were. The study's results not only reveal sloppy methodological choices, but indicate that the authors took extra steps to portray their findings as more policy relevant than was warranted. And even if their findings were wholly accurate, which is highly unlikely, their policy suggestions would not follow from their results. While this essay assumes only a passing familiarity with the study in question, I have also written a full-length report that contains a longer and more technical discussion of the study and expands upon some of the work from Ritchie, Gelman, and Alexander.¹⁰

Problems With the Literature Linking EEGs to Psychological Outcomes

For the results of Troller-Renfree et al.'s study to be important, changes in electroencephalography (EEG) power¹¹ would need to be causally linked to changes in its psychological correlates, such as IQ or language skills. The study's authors cited several pieces of research that supposedly linked EEG power to their desired, preregistered cognitive outcomes, and studies relating EEG power to socioeconomic status (SES). Cumulatively, Troller-Renfree et al. argued that these studies linked absolute power in mid-to-high frequency bands to linguistic, cognitive, and social-emotional development, while simultaneously arguing that low-frequency band power was linked to behavioral and learning issues.

However, these studies suffered from many major problems: failing to deal with the issue of multiple comparisons, small sample sizes and low power, inconsistent results, confounding with genes and family environments, representativeness problems, various reporting errors, and nonsense models. The relevant descriptive statistics for their cited studies relating EEG power to cognition are provided in Table 1, and the studies relating it to SES are in Table 2. I cover the problems with these studies in greater depth in the full-length report, but here I give a short overview of some of the basic flaws in their literature:

Multiple Comparisons

CSPI

When too many tests are run, the likelihood of getting a false-positive significant effect increases. This is the problem of multiple comparisons, and it is amenable to a graphical explanation. Figure 1 illustrates the relationship between the number of statistical tests and the probability that at least one of the results is significant if there is no true effect, based on a significance level of 0.05.

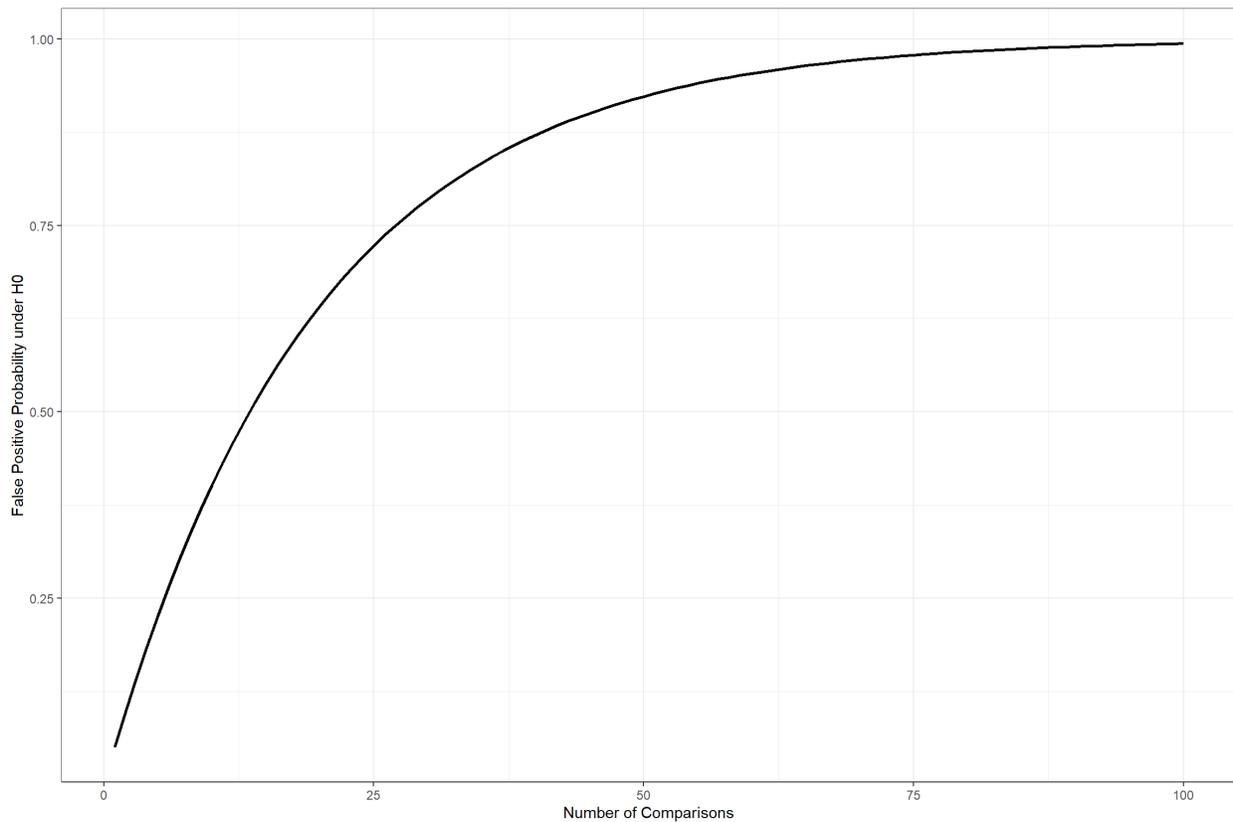


Figure 1

The typical study in Tables 1 or 2 below had well over ten tests, but most also did not correct for multiple comparisons, so the risk of false positives was high in the literature cited by Troller-Renfree et al.

Small Samples and Low Power

Small sample sizes are a serious problem in scientific research because studies that rely on them lack statistical power: the probability that a statistical test detects a real effect of a given size at some level of significance. If I have a sample size of 20 and I need to find a correlation between X and Y that has a p -value below 0.05, the lowest significant correlation would be Pearson's $r = 0.44$. If I quintuple my sample size, the lowest significant correlation is 0.20. This does not mean that a small sample cannot yield small correlations, only that, when it does, they will not be significant.

The issue with small sample sizes is also open to graphical explanation. Figure 2 illustrates the relationship between sample sizes and the average sample correlation conditioned on reaching a 5% significance threshold (the mean significant correlation) in simulated data with

CSPI

different true underlying correlations. The mean sample sizes in Tables 1 and 2 were 89 and 67, respectively. If the true correlation is typical of the correlations observed in large, reliable studies in neuroscience (closer to 0.10), we can infer that the correlations in the studies Trolley-Renfree et al. cited were exaggerated by 2.6 to 3 times.¹²

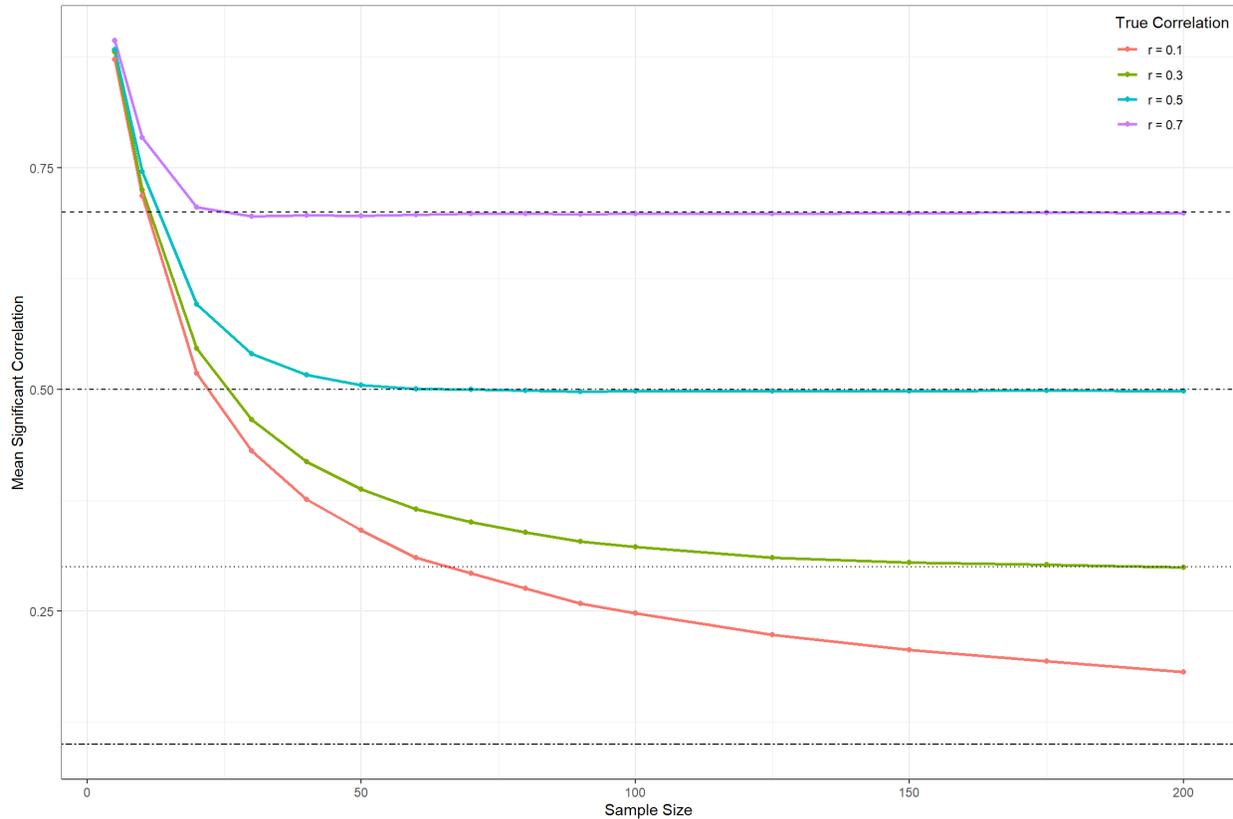


Figure 2

There is some evidence in Tables 1 and 2 that suggests these effect sizes were exaggerated, because smaller studies had larger effects. Moreover, there was a significant negative correlation between the size of correlations and their precision,¹³ an indication that researchers were probably trying to find significant results and, as such, had to find and publish exaggerated effects because their samples were too small.¹⁴

CSPI

Table 1. Study Descriptives for Cited Cognitive Ability-EEG Studies¹⁵

Study	Outcome	Multiple Comparisons	Sample Size *	Effect Size **	Critical r ***	Power †
Benasich et al. (2008)	Associations with Absolute Gamma Power	Not corrected for	At most, 63	0.52 - 0.75	0.25	0.12
Harmony et al. (1990)	Associations with Absolute and Relative Power (Various)	Not corrected for	At most, 81	0.52 - 0.93	0.22	0.14
Gou, Choudhury & Benasich (2011)	Associations with Absolute Gamma Power	Not corrected for	At most, 17	0.51 - 0.72	0.49	0.06
Williams et al. (2012)	Associations with Absolute Beta Power	p -value cutoff set to 0.006	13	0.77‡	0.72<	0.01
Brito et al. (2016)	Associations with Absolute Gamma Power	Not corrected for	At most, 64	0.28 - 0.32	0.25	0.12
Brito et al. (2019)	Associations with Absolute Power (Various)	Vague: "running FDR to correct"	At most, 129	ns	0.17	0.20
Cantiani et al. (2019)	Associations with Gamma Power	Not corrected for	At most, 60	0.26 - 0.41	0.25	0.12
Troller-Renfree et al. (2020)	Associations with Relative Theta, Alpha, and Gamma Power	Not corrected for	79	ns	0.22	0.14
Maguire & Schneider (2019)	Associations with Absolute Alpha and Theta Power	Monte Carlo cluster-corrected permutation analysis	At most, 90	0.32	0.21	0.16

* This is the sample size for their correlations, which is sometimes not the maximum size of their samples. ** Significant correlations, in terms of r , and for Harmony et al. (1990), canonical r . *** This is the smallest significant correlation at their maximum sample size. † This is the power to detect an effect of $r = 0.1$ with their maximum sample size, 80% power, no imbalance, and a p -value cutoff of 0.05. ‡ The correlation was reported as two different values, 0.796 and 0.77, but based on the reported p -value of 0.002, it could not be 0.796. < The critical r for this study was computed at their alpha value. N 's are chosen for maximum generosity to the studies, which exaggerates their power in every case.

CSPI

Table 2. Study Descriptives for Cited SES-EEG Studies¹⁶

Study	Outcome	Multiple Comparisons	Sample Size	Effect Sizes *	Critical r/F **	Power †
Otero (1994)	Associations with Absolute and Relative Delta, Alpha, and Beta Power	Not corrected for	50	6.02 – 19.82	F = 4.04	0.11
Otero et al. (2003)	Associations with Absolute and Relative Delta, Theta, and Alpha Power	Bonferroni without mention of number of tests	42	? ‡	F = 4.08	0.10
McLaughlin et al. (2010)	Associations with Theta, Alpha, and Beta Power	Not corrected for	135	4.61 – 9.42	F = 2.67	0.14
Tomalski et al. (2013)	Associations with Gamma Power	Not corrected for in main analyses	45	3.23 – 4.75	F = 4.07	0.10
Pierce et al. (2019)	Associations with Delta, Theta, Beta, and Gamma Power	Not corrected for	At most, 70	Unclear because of overadjustment, but mostly nonsignificant.	$r = 0.24$	0.13
Cantini et al. (2019) †	Associations with Gamma Power	Not corrected for	At most, 81 for r and 78 for F	0.23 – 0.30; 6.67	$r = 0.22$; F = 3.96	0.14/0.15
Debnath et al. (2019) †	Associated with Absolute and Relative Theta, Alpha, and Beta Power	Not corrected for in main analyses	45; 138	-0.329 – 0.303; 5.051 – 5.950	$r = 0.29$; F = 3.06	0.10/0.16
Troller-Renfree et al. (2020)	Associations with Absolute and Relative Theta, Alpha, Beta, and Gamma Power	Not corrected for	79	Ns #	0.22	0.14
Brito et al. (2020)	Associations with Beta and Gamma Power	Vague: “FDR correction”	At most, 60	Ns	0.25	0.12
Jensen et al. (2021)	Associations with Theta, Alpha, Beta, and Gamma Power	Not corrected for	At most, 187	-0.204 – -0.176	0.14	0.28

* Significant correlations, or a significant F-value (not to be confused with f). Some studies did many different types of tests, so their main ones are highlighted. ** This is the smallest significant correlation or lowest F-value at their maximum sample size. † This is the power to detect an effect of $r = 0.1/f = 0.1$ with their maximum sample size, 80% power, and a p -value cutoff of 0.05. ‡ Many of their values were left unreported. † These authors also reported a saturated mediation model and, for indiscernible reasons, provided its fit measures. † Marshall et al. 2008 and Vanderwert et al. 2010 did substantially the same analyses with the same sample and were therefore not analyzed separately. # This was confusing. All relative power relationships with parental education and income were nonsignificant, and the same was true for all but absolute theta’s relationship with parental education, which had impossible values. A beta of 0.033 cannot be significant with an SE of 0.166, but 0.33 could, although marginally, in the ‘wrong’ direction, and not after correction for multiple comparisons. Harmony et al. (1988) could have had usable data, but they did not provide summary statistics for their samples, so SES-ability relationships could not be investigated.

Unrepresentative Samples

The studies cited by Troller-Renfree et al. had problematic and unusual sampling. Some of the studies featured samples of people with congenital heart defects,¹⁷ some of the samples had very high average IQs, in another 85% of the sample's parents belonged to the highest socioeconomic strata and a third of the involved families had histories of language impairment, and so on. The samples were often decidedly unrepresentative of the general population, so it is difficult to generalize from them.

Problems With the Study Itself

Gaming Preregistration

Small studies lead to inconsistent and overstated effects. Combined with an editorial bias towards papers with statistically significant results, they encourage scouring datasets for significant relationships (commonly known as *p*-hacking). One way to limit the pernicious effects of searching for significance is for researchers to preregister the hypotheses they intend to check, forcing them to report specific outcomes they may have decided not to publish due to a lack of significance.

Troller-Renfree et al. preregistered that they were going to examine effects on theta, alpha, and gamma power. Respectively, these are brainwave frequency bands ranging from 4 – 8, 8 – 13, and >35 Hz. Curiously, their largest effect was in the beta band (13 – 35 Hz). They preregistered their intention to investigate only alpha, gamma, and theta because, when they formulated the study, the literature showed effects on those bands – not beta. However, “between 2018 and the present investigation,” they wrote, “evidence has emerged linking income to beta activity, including from the first and senior authors’ lab.”¹⁸ Two studies were cited to support that claim.

The first study they cited linking beta to income was by Brito et al. (2020). It was entirely cross-sectional and featured, at most, correlations in a study of 60 kids. It yielded no significant relationships between SES and beta or gamma power, either by region or in total, and it also showcased no relationship between SES and a child's auditory or expressive communication abilities. The second study by Jensen et al. (2021) was the largest study among those cited, with a sample size of at most 187. If this study is taken seriously, it contradicts Troller-Renfree et al. because it found no relationships between SES and EEG power for their sample of 160 6-month-olds, and *negative* relationships between SES and both beta and gamma power among their 187 36-month-olds.

Troller-Renfree et al. found a positive effect of their intervention on both beta and gamma, despite their cited beta evidence being consistent with no effect or a negative one. If Troller-Renfree et al.'s reasoning was truly informed by these two studies, they should have been *alarmed* to find that two high-frequency bands, gamma and beta, within which power may have been negatively related to SES were increased by their experiment, indicating kids may have been *harmed* by the intervention.

Moreover, Troller-Renfree et al. did not preregister their choice of robust estimator or treatment effect size. They used regression to estimate the effect of treatment, and to avoid their estimates being affected by violations of any of the several assumptions involved in regression, they utilized robust estimation, a way of running their regressions while obviating issues with its assumptions. There are many types of robust estimators and different statistical software uses

CSPI

different default estimators. In Stata, the analysis program they used, the default is HC1 – a robust estimator that is common by virtue of its default status rather than its performance. More advanced robust estimators exist, and they ought to be used to deliver accurate results when, for example, samples are not very large, like theirs. Using HC5, a more modern alternative, I found that their significant results became wholly nonsignificant (i.e., $p > 0.05$ before correction).

The effect size of their treatment was computed by dividing their adjusted treatment effect by the standard deviation (SD) of their control group; a common method for computing standardized treatment effects. However, they could have also divided by the pooled SD, either weighted by each group's sample size or not. That they did not use the pooled SD would not normally be an issue, but out of all the effect size computation methods they could have chosen, they picked the one that delivered them two significant results prior to multiple comparison correction.

Similar Magnitudes Don't Matter

Troller-Renfree et al. wrote that, “[d]espite the limitations in statistical power, the pattern of impacts, which resulted from a rigorous random assignment study design, were consistent with hypotheses, were similar in magnitude to effects on cognitive outcomes from other scalable interventions, and were largely robust to various tests, leads us to conclude that these findings are important and unlikely to be spurious.” But they never explained why these similarities mattered. If several studies show that giving children iron improves growth by five centimeters ($d = 0.25$), I wouldn't be able to use that finding to show that teaching children chess and finding that they had more theta wave activity ($d = 0.25$) was more plausible just because we can draw vague links between them and the effect sizes are the same. The studies have no necessary relationship with one another. They deal with different interventions and outcomes, so the effect sizes in each domain probably *should* differ. The mere fact that it's possible to generate a likely chance finding about one does not vindicate its comparison with another.

Long-Term Treatment Effects and Parental Responsibility

Despite its weak evidentiary basis, effusive support for Troller-Renfree et al.'s article has been common. Representative Suzan DelBene (D-Wa.) was quoted by *The New York Times* saying it showed “investing in our children has incredible long-term benefits.”¹⁹ In the same venue, Duke economist Lisa Gennetian – one of the study's co-authors – claimed the study showed parents could be trusted to spend money well. Neither piece of commentary finds any support in the study. The former is nonsensical because the study only covered a single year, not the long term, and the latter is simply not viable because, for one, the study did not examine how well money was spent, and our only way to assess this was to look at the money's effect on children.

Moreover, because the authors of the study controlled for variables that indicated parental misbehavior – smoking and alcohol consumption during pregnancy, marital status, and maternal mental health – they removed the ability to draw inferences about how well parents used the money in the same way that controlling for the quality of the venue, service, preparation, and ingredients makes it difficult to accurately compare McDonald's and a five-star restaurant. It would have been reasonable to adjust the data if there were no systematic differences in attrition, and randomization between the treatment and control groups was satisfactory, but their controls remove our ability to draw certain conclusions.

Most Interventions Have Small to Nil Effects

Troller-Renfree et al. exaggerated the sizes of the effects from other studies that they used to validate their own. The citation they provided for the similarity of their experimental effect with other interventions was Kraft (2019), who performed a large meta-analysis of educational interventions and found a meta-analytic effect size of 0.16 *d*.²⁰ However, Kraft's data was typical of meta-analyses in that it suffered from a bias towards studies with significant effects that were typically small. The *weighted* effect size was 0.04 *d* - one-fourth as large, and not significant. Lortie-Forgues and Inglis (2019) conducted a similar study, but theirs did not suffer from the same problems as Kraft because they analyzed studies funded by organizations that required the trials they supported to be published in a standardized fashion.²¹ The effect size they found was 0.06 *d*, and the median study in their analysis was consistent with anecdotal evidence *for no effect*; that is, the quality of the evidence for the median study should not have been regarded as substantial enough to support change in belief about intervention effectiveness.²²

Other interventions have also had similarly minor effects. The 2010 Head Start Impact Study conducted by the Office of Planning, Research and Evaluation looked at effects elicited by the Head Start program for three- and four-year-old children.²³ Using a *p*-value cutoff of 0.10, it found significant positive effects on eight of fourteen cognitive tests given in the year of the program for the three-year-old children, and this reduced to two of 14 by the next year, at age four. In kindergarten, there were 19 tests administered, with one significant positive effect, and one significant negative one, while, in the first grade, only one effect of twenty-two remained significant. Because this was not significantly affected *during* Head Start and its effect only became significant amongst a pile of nonsignificant ones, it was probably noise. For the four-year-old group, seven of 14 scores were significantly positively affected in the Head Start year, but effects disappeared in kindergarten, and one effect became significant again in first grade. Relatedly, Duncan and Magnuson (2013) found that the effect sizes associated with childcare programs declined over time.²⁴ When I repurposed Duncan and Magnuson's data, there was a negative relationship between a study's precision and how large an effect it produced – publication bias!²⁵

Studies of nutritional interventions have fared similarly. Dulal et al. (2018) ran a double-blind RCT in Nepal where groups of mothers were given prenatal multiple micronutrient supplementation in the treatment group, and iron and folic acid only in the control group.²⁶ At 12 years of age, there were no effects on children's cognitive ability. Behrens et al. (2020) reviewed whether vitamin B affected rates of cognitive decline and found no effects, despite well-defined mechanisms and plausible effects at the time of writing.²⁷ Similarly, a large-scale follow-up of two double-blinded, placebo-controlled, cluster-randomized trials of vitamin A effects on prenatal and newborn children by Ali et al. (2017) yielded nonsignificant effects across the board.²⁸

Even when the intervention is medical in nature, effects are often minute. Welch et al. (2018) reported the results of a systematic review of deworming effects on a variety of outcomes including cognitive ability, height, and weight.²⁹ Despite parasitic worms often being debilitating and the mechanisms through which they should impede development being abundantly clear, there were no significant meta-analytic effects on weight, height, or cognitive ability. Welch et al. (2018) compared their results to earlier systematic reviews by Taylor-Robinson et al. (2015)

CSPI

and Welch et al. (2016), which, altogether, found one marginally significant meta-analytic effect on weight and none for height or cognitive ability.³⁰

Analyses of cash transfers specifically have not yielded large effects either. Zimmerman et al. (2021) evaluated the effects of cash transfers on psychological distress and depression and found evidence consistent with no effect.³¹ There may have been effects on variables like an individual's outlook, hope, and feelings of power in their sexual relationships, but these would have nonetheless been modest, imprecise, and insufficiently replicated. Manley et al. (2020) examined cash transfer effects on more plausible outcomes – height, weight, dietary diversity, the consumption of animal-based foods, etc.³² Among published studies, smaller studies had larger effects, indicating a small study bias, but there was a significant effect (p -value between 0.05 and 0.01) on stunting. Among unpublished studies, there was more concentration of effects around zero, and the aggregate effect on growth stunting was nonsignificant. With some evidence for publication bias, the strongest interpretation of that meta-analysis is a marginal effect on meaningful growth parameters. The full-length report contains additional details on several other relevant studies.

Perhaps if the cash transfers were massively scaled up, they might have an effect. Certain studies *do* seem to look like extremely extensive and large-scaled interventions. For example, the Moving to Opportunity program that relocated people from low- to high-quality neighborhoods had an extraordinary economic value, but it failed to elicit an effect on measured mathematical and reading skills.³³ Perhaps that was because family *inside* the home has a much greater effect than the location of the home. Adoption puts people into new homes and families that are generally much wealthier and better educated than their birth families or the general population, amounting to a natural experiment in giving people comprehensively high SES. The effects from good adoption studies generally span between nil (Ericsson et al., 2017) and *just* over the effect size Troller-Renfree et al. observed (Kendler et al., 2015).³⁴

Some people object to adoption studies on the grounds that we cannot properly control for preadoption circumstances and their persistent effects on children. But when Korean adoptees in Sweden were examined, there was no relationship between age at adoption and subsequent cognitive ability.³⁵ The reason for this appears to be that Korea would not allow people to select which baby they wanted, so in the absence of parental selection, the relationship between age at adoption and subsequent performance breaks down.³⁶ With the size of adoption effects noted, we can plot how they compare to the cross-sectional association between parental SES and cognitive performance. Figure 3 does this and shows that despite adoption affecting how well people score (their performance), the effect is not strong enough to generate the same relationship between SES and performance that we see in the general population, where SES is tied not only to people's environments, but also to their genes.³⁷

CSPI

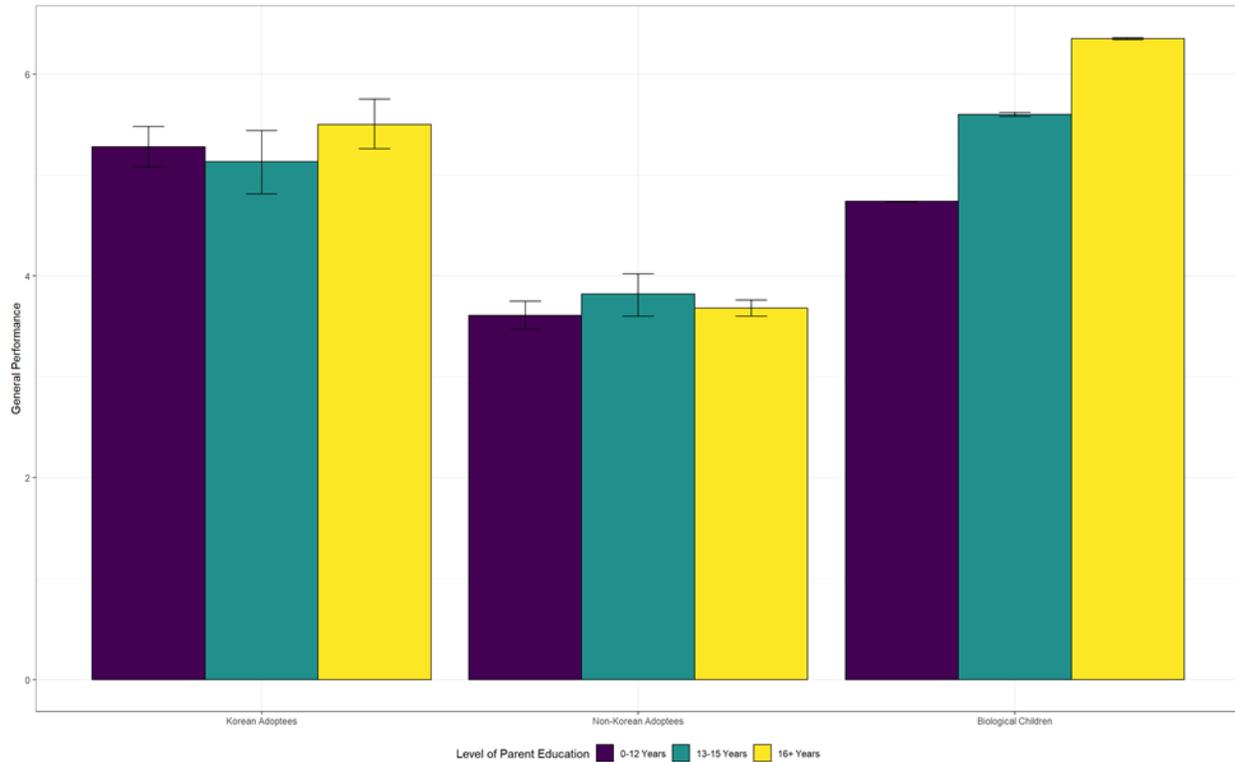


Figure 3

The issue of genetic confounding is extremely important. Cross-sectional studies like the ones discussed here are generally not informative for Troller-Renfree et al.'s purposes because associations between variables like SES and specific brainwave patterns may occur due to a common cause or correlated causes. In the case of comparisons between members of different families, the issue may be genetic confounding – where genes explain the relationships among different outcomes, like EEG power and family income. The relationship between intelligence and brain white and gray matter volumes is accounted for by shared genes.³⁸ Similarly, the relationship between socioeconomic status and intelligence has been found to have a substantial genetic component, and the link between the two is virtually severed when genetics are accounted for.³⁹ Even EEG power is heritable, with most relationships among EEG power at different frequencies attributable to shared genes.⁴⁰

As a result of substantial genetic influence and other potential forms of confounding, how are we to generalize from experimental effects on EEG parameters to cross-sectional samples or vice versa? Troller-Renfree et al. were at least somewhat aware of these facts, as evidenced by their citation of Wax (2017), who argued that “the so-called neuroscience of deprivation has no unique practical payoff... Because this research does not, and generally cannot, distinguish between innate versus environmental causes of brain characteristics, it cannot predict whether neurological and behavioral deficits can be addressed by reducing social deprivation.”⁴¹

Policy Conclusions Were Unwarranted

Troller-Renfree et al. have attempted to sell their results as relevant to the ongoing debate over child tax credits. Their press release for the study said as much, and mainstream sources

CSPI

have used their results to justify child tax credits on the grounds that they will make kids' brains work better.⁴² Since the strong interpretation of their results is unwarranted, so is the conclusion that the child tax credits are justified by brain improvements. Insofar as the results were basically uncertain, it was wrong to frame them as policy relevant.

But child tax credits do present an interesting opportunity for study, albeit outside of the context of Troller-Renfree et al.'s flawed work. If their conclusions were totally correct and we could generalize the largest of their effects on EEG power to some psychological trait like intelligence, we should be able to pick that up in children's achievement *now*. Children whose parents received child tax credits can be identified, and we should see a major boost in their achievement down the line. Those who believe interventions timed earlier in life have larger effects have an appropriate test in the form of credits to families with kids of varying ages, and people who believe effects should persist should expect the ability distribution of those whose families received cash when they were younger to be raised. If the case for child tax credits based on cognitive improvement is as strong as Troller-Renfree et al. and others suggested, the effects will be extremely noticeable. To get an idea of how big this could be, if the IQ of the population increased by 0.26 SDs, we would have almost three times as many people with IQs of 160 or greater, and almost half as many with IQs of 70 or less.⁴³

After their study was published, Troller-Renfree et al.'s work received immediate pushback. Many who were quick to give the study plaudits recalibrated their judgements upon learning about its flaws. The UBI Center removed their posts about the study, the Niskanen Center put up a disclaimer, and even *Vox* updated its coverage to feature critical commentary.⁴⁴ Unfortunately, *Vox* also published a dubious response from the study's lead author, Kimberly Noble.⁴⁵

She confronted criticisms with a curious argument, arguing that the results were robust because the authors observed predicted region-specific relationships. This was obviously untrue, as the prior literature does not and could not adequately suggest region-specific associations unless she failed to cite it in the article. Moreover, the associations did not survive when I corrected for multiple comparisons (i.e., p became greater than 0.05 universally), which were done incorrectly for this section of the paper. As discussed, work claiming to find region-specific associations requires a lot of power that studies in this literature lack. Noting that you can perform a bunch of tests and recapitulate your initial results does not mean they were supported, it means you think transforming your data and finding the same things you did prior to transformation bolsters your initial conclusions. Anyone could do this and claim any finding is robust, regardless of topic. All it signals is a desire for the study to be considered important rather than its quality.

Conclusion

Troller-Renfree et al. conducted an intervention where poor families were given \$333/month; income subsidies equal to around 20% of their prior incomes. Their results were nonsignificant impacts on infants' EEG parameters. They inferred that if there were effects, they might have been important because other literature linked some of the ostensibly affected EEG parameters to important cognitive outcomes and did so with magnitudes apparently comparable to educational interventions.

Troller-Renfree et al. needed two things to be true for their study to be important. First, they needed a treatment effect on EEG power. Second, they needed specific changes in EEG power to causally impact important psychological traits like cognitive ability, depression,

CSPI

ADHD, etc. The evidence for a treatment effect was absent from their study and the evidence in favor of any link, much less a causal one, between specific bands' EEG power and desirable psychological outcomes was tenuous at best. The groundwork needed for their results to be treated as if they were meaningful or to even treat their inferences as more than a passing flight of fancy was not provided and may not exist. For results based on a chain of reasoning about previous findings to be taken seriously requires that chain be strong, and this one was not.

The authors had the opportunity to test if manipulating an aspect of SES impacted cognitive outcomes, and they did so via proxy. Per their theory, SES impacts the brain, which impacts cognition, but the links between SES and EEG parameters – much less EEG and psychological ones – were so tenuous as to render the theory not only unevidenced, but wholly useless. Given that their paper was systematically tilted towards a positive interpretation of essentially null results, it was, in essence, an explanation of their theory, rather than a test of it.

Using weak science to advocate for policy is a frequent occurrence. Most social science is weak, and journalists and activists are often looking for evidence that confirms what they want to believe. Troller-Renfree et al. provided a case study in many of the most typical mistakes made by people who want to support a cause without doing the legwork. They asked us to not only accept their methods, but their interpretations and theories relying on a chain of strong relationships that have yet to be shown to exist. Those who believe this study has any policy relevance should start placing bets that there will be a massive uptick in rates of genius and plummeting rates of mental disability in the U.S. because of recent child tax credits. There won't be.

The public arena is at least somewhat self-correcting, in that the findings in Troller-Renfree et al. were quickly scoured in the popular press. The bigger problem, however, is that if it had not gotten a write up in the *New York Times* accompanied by a news alert, and if the study's topic and author's framing had not been so palpably absurd, the paper may have never been carefully checked and the results may have turned into conventional wisdom. Bad science is the norm; correcting it is not.

Cases like this ought to make us reconsider the role social science plays in our public policy debates. Political bias and confirmation bias are heuristics that plague us all, and we have no reason to think that social scientists or policymakers are immune. Flashy results that support a popular policy are usually untrustworthy, and large effects are usually exaggerated, *p*-hacked, or due to chance. The reality is that most social interventions and policies have a negligible impact when it comes to improving cognitive ability or behavior. Until researchers and the educated public come to grips with this fact, we should be skeptical of policymakers' ability to evaluate research objectively and social scientists' ability to reliably inform public policy with their work.

¹ R code for original analyses: <https://rpubs.com/JLLJ/TTAN> includes several power analyses and various tests related to this report; <https://rpubs.com/JLLJ/MSRY> includes methods to assess effect size overestimation and plots to aid in understanding the relationship between statistical power and overestimation; and <https://rpubs.com/JLLJ/ODEN> includes various adoption study-related plots.

² DeParle, Jason. 2022. "Cash Aid to Poor Mothers Increases Brain Activity in Babies, Study Finds." *The New York Times*. Available at <https://www.nytimes.com/2022/01/24/us/politics/child-tax-credit-brain-function.html>.

³ Troller-Renfree, Sonya V., Molly A. Costanzo, Greg J. Duncan, Katherine Magnuson, Lisa A. Gennetian, Hirokazu Yoshikawa, Sarah Halpern-Meekin, Nathan A. Fox and Kimberly G. Noble. 2022. "The Impact of a Poverty Reduction Intervention on Infant Brain Activity." *PNAS* 199(5).

⁴ Matthews, Dylan. 2022. “Can Giving Parents Cash Help with Babies’ Brain Development?” *Vox*. Available at <https://www.vox.com/future-perfect/22893313/cash-babies-brain-development>.

⁵ Jirari, Tahra and Ed Prester. 2022. “Cash Benefits to Low-Income Families May Aid Babies’ Cognitive Development.” *Niskanen Center*. Available at <https://www.niskanencenter.org/cash-benefits-to-low-income-families-aids-babies-cognitive-development/>; Columbia University. 2022. “Cash Support for Low-Income Families Impacts Infant Brain Activity.” *Medical Xpress*. Available at <https://medicalxpress.com/news/2022-01-cash-low-income-families-impacts-infant.html>; Sullivan, Kaitlin. 2022. “Giving Low-Income Families Cash Can Help Babies’ Brain Activity.” *NBC News*. Available at <https://www.nbcnews.com/health/health-news/poverty-hurts-early-brain-development-giving-families-cash-can-help-rcna13321>; Smith, Zachary Snowden. 2022. “Giving Moms Money can Boost Babies’ Brainwaves, Study Finds.” *Forbes*. Available at <https://www.forbes.com/sites/zacharysmith/2022/01/24/giving-moms-money-can-boost-babies-brain-activity-study-finds/?sh=2d2a456d20c7>; For a full list of mentions and media appearances, see Troller-Renfree et al.’s press page, available at <https://www.babysfirstyears.com/press>.

⁶ “Press Release: Cash Support for Low-Income Families Impacts Infant Brain Activity.” 2022. *Baby’s First Years*. Available at <https://www.babysfirstyears.com/press-release>.

⁷ Ritchie, Stuart [@StuartJRitchie]. 2022. *Great to see, part 2: [Tweet]*. Twitter. <https://twitter.com/StuartJRitchie/status/1486814686125375499>; Ritchie, Stuart. 2022. “The Real Lesson of that Cash-for-Babies Study.” *The Atlantic*. Available at <https://www.theatlantic.com/science/archive/2022/02/cash-transfer-babies-study-neuroscience-hype/621488/>.

⁸ Gelman, Andrew. 2022. “I’m Skeptical of That Claim That ‘Cash Aid to Poor Mothers Increases Brain Activity in Babies.’” *Statistical Modeling, Causal Inference, and Social Science*. Available at <https://statmodeling.stat.columbia.edu/2022/01/25/im-skeptical-of-that-claim-that-cash-aid-to-poor-mothers-increases-brain-activity-in-babies/>.

⁹ Alexander, Scott. 2022. “Against That Poverty and Infant EEGs Study.” *Astral Codex Ten*. Available at <https://astralcodexten.substack.com/p/against-that-poverty-and-infant-eegs>.

¹⁰ The full length version of the report is available here: <https://cspicenter.org/reports/about-those-baby-brainwaves-why-policy-relevant-social-science-is-mostly-a-fraud-full-report/>.

¹¹ Squared brainwave amplitude. Power is often subdivided into absolute and relative power, with the difference being the amount of activity at a given frequency versus the absolute power as a fraction of absolute power across a certain range of frequencies. These frequencies are subdivided into ranges or bands which are given Greek-lettered names, like delta (0.5 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 13 Hz), beta (13 – 35 Hz), and gamma (>35 Hz), giving us terms like “alpha power” or “delta power” to denote power within those ranges. The ranges for these specific brainwaves vary slightly by source, but the order I’ve listed them in places them in their consistent relative positions from lowest to highest frequency.

¹² This assumes balanced samples, a lack of measurement error, and no heteroskedasticity, high-leverage outliers, or nonnormality, and therefore overestimates power while underestimating the degree of exaggeration.

¹³ This is a statistical term meaning the reciprocal of the variance, where more precise studies have smaller variances and thus we are more certain about the magnitude of their estimated effects.

¹⁴ Trim-and-fill, a technique for publication bias correction, applied to reported correlations, rendered the aggregate relationships between EEG power (regardless of band) and either cognitive ability or SES nonsignificant.

¹⁵ Benasich, April A., Zhenkun Gou, Naseem Choudhury and Kenneth D. Harris. 2008. “Early Cognitive and Language Skills Are Linked to Resting Frontal Gamma Power Across the First Three Years.” *Behavioural Brain Research* 195(2): 215-222; Harmony, Thalía, Erzsébet Marosi, Ana E. Díaz de León, Jacqueline Becker and Thalía Fernández. 1990. “Effect of Sex, Psychosocial Disadvantages and Biological Risk Factors on EEG Maturation.” *Electroencephalography and Clinical Neurophysiology* 75(6): 482-91; Gou, Zhenkun, Naseem Choudhury and April A. Benasich. 2011. “Resting Frontal Gamma Power at 16, 24 and 36 Months Predicts Individual Differences in Language and Cognition at 4 and 5 Years.” *Behavioural Brain Research* 220(2): 263-70; Williams, I. A., A. R. Tarullo, P. G. Grieve, A. Wilpers, E. F. Vignola, M. M. Myers and W. P. Fifer. 2012. “Fetal Cerebrovascular Resistance and Neonatal EEG Predict 18-Month Neurodevelopmental Outcome in Infants with Congenital Heart Disease.” *Ultrasound in Obstetrics & Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology* 40(3): 304-309; Brito, Natalie H., William P. Fifer, Michael M. Myers, Amy J. Elliott

- and Kimberly G. Noble. 2016. "Associations among Family Socioeconomic Status, EEG Power at Birth, and Cognitive Skills during Infancy." *Developmental Cognitive Neuroscience* 19: 144; Brito, Natalie H., Amy J. Elliott, Joseph R. Isler, Cynthia Rodriguez, Christa Friedrich, Lauren C. Shuffrey and William P. Fifer. 2019. "Neonatal EEG Linked to Individual Differences in Socioemotional Outcomes and Autism Risk in Toddlers." *Developmental Psychobiology* 61(8): 1110-19; Cantiani, Chiara, Caterina Piazza, Giulia Mornati, Massimo Molteni and Valentina Riva. 2019. "Oscillatory Gamma Activity Mediates the Pathway from Socioeconomic Status to Language Acquisition in Infancy." *Infant Behavior and Development* 57; Troller-Renfree, Sonya V., Natalie H. Brito, Pooja M. Desai, Ana G. Leon-Santos, Cynthia A. Wiltshire, Summer N. Motton, Jerrold S. Meyer, Joseph Isler, William P. Fifer and Kimberly G. Noble. 2020. "Infants of Mothers with Higher Physiological Stress Show Alterations in Brain Function." *Developmental Science* 23(6); Maguire, Mandy J. and Julie M. Schneider. 2019. "Socioeconomic Status Related Differences in Resting State EEG Activity Correspond to Differences in Vocabulary and Working Memory in Grade School." *Brain and Cognition* 137.
- ¹⁶ Otero, Gloria. 1994. "EEG Spectral Analysis in Children with Sociocultural Handicaps." *The International Journal of Neuroscience* 79(3-4): 213-220; Otero, G. A., F. B. Pliego-Rivero, T. Fernández and J. Ricardo. 2003. "EEG Development in Children with Sociocultural Disadvantages: A Follow-up Study." *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 114(10): 1918-1925; McLaughlin, Katie A., Nathan A. Fox, Charles H. Zeanah, Margaret A. Sheridan, Peter Marshall and Charles A. Nelson. 2010. "Delayed Maturation in Brain Electrical Activity Partially Explains the Association Between Early Environmental Deprivation and Symptoms of Attention-Deficit/Hyperactivity Disorder." *Biological Psychiatry* 68(4): 329-336; Tomalski, Przemyslaw, Derek G. Moore, Helena Ribeiro, Emma L. Axelsson, Elizabeth Murphy, Annette Karmiloff-Smith, Mark H. Johnson and Elena Kushnerenko. 2013. "Socioeconomic Status and Functional Brain Development - Associations in Early Infancy." *Developmental Science* 16(5): 676-687; Cantiani et al. 2019; Debnath, Ranjan, Alva Tang, Charles H. Zeanah, Charles A. Nelson and Nathan A. Fox. 2020. "The Long-Term Effects of Institutional Rearing, Foster Care Intervention and Disruptions in Care on Brain Electrical Activity in Adolescence." *Developmental Science* 23(1); Troller-Renfree et al. 2020; Brito, Natalie H., Sonya V. Troller-Renfree, Ana Leon-Santos, Joseph R. Isler, William P. Fifer and Kimberly G. Noble. 2020. "Associations Among the Home Language Environment and Neural Activity during Infancy." *Developmental Cognitive Neuroscience* 43; Jensen, Sarah K. G., Wanze Xie, Swapna Kumar, Rashidul Haque, William A. Petri and Charles A. Nelson. 2021. "Associations of Socioeconomic and Other Environmental Factors with Early Brain Development in Bangladeshi Infants and Children." *Developmental Cognitive Neuroscience* 50.
- ¹⁷ Individuals with congenital heart disease frequently suffer from cognitive deficits. Phenomena such as enhanced mosaicism in trisomy 21 and greater numbers of copy-number variants have been related to congenital heart disease and cognitive deficits. See Pierpont, Mary Ella, Martina Brueckner, Wendy K. Chung, Vidu Garg, Ronald V. Lacro, Amy L. McGuire, Seema Mital, James R. Priest, William T. Pu, Amy Roberts, Stephanie M. Ware, Bruce D. Gelb, Mark W. Russell and On behalf of the American Heart Association Council on Cardiovascular Disease in the Young; Council on Cardiovascular and Stroke Nursing; and Council on Genomic and Precision Medicine. 2018. "Genetic Basis for Congenital Heart Disease: Revisited: A Scientific Statement from the American Heart Association." *Circulation* 138(21).
- ¹⁸ From the supplement of Troller-Renfree et al. (2022). Available at <https://www.pnas.org/content/suppl/2022/01/20/2115649119.DCSupplemental>.
- ¹⁹ DeParle 2022.
- ²⁰ Kraft, Matthew A. 2019. "Interpreting Effect Sizes of Education Interventions." *Annenberg Institute at Brown University*. Available at https://scholar.harvard.edu/files/mkraft/files/kraft_2019_effect_sizes.pdf.
- ²¹ Lortie-Forgues, Hugues and Matthew Inglis. 2019. "Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned?" *Educational Researcher* 48(3).
- ²² For Bayesians, evidence is thought of as a continuous quantity where you can have different levels of evidence for something. Evidence is thought of in terms of cutoffs with names like "anecdotal", "modest", or "extreme". When evidence is anecdotal, it should barely move the ticker.
- ²³ "Head Start Impact Study: Final Report, Executive Summary." 2010. *Office of Planning, Research & Evaluation*. Available at <https://www.acf.hhs.gov/opre/report/head-start-impact-study-final-report-executive-summary>.

²⁴ Duncan, Greg J. and Katherine Magnuson. 2013. “Investing in Preschool Programs.” *Journal of Economic Perspectives* 27(2): 109-132.

²⁵ Publication bias is the phenomenon whereby significant or otherwise favorable results are more likely to be published. It is often evidenced, for example, by correlating the standard errors of studies with their point estimates or by assessing the symmetry of point estimates around the meta-analytic mean. The first is informative because at some level, smaller studies must yield larger effects, and the pattern also suggests searching for significance because researchers want significant results, but their sample sizes are constrained. This latter suggestion brings us to the second method of checking for asymmetry of points around the meta-analytic mean. Asymmetry means that there are omitted studies with estimates on one side of the meta-analytic mean but can also indicate that larger studies yield smaller effects. This pattern could be argued to emerge because of, for example, smaller studies in a meta-analysis of experiments being more intensive because they can dedicate their limited resources to a smaller group, whereas larger studies must spread out their resources over more people, diluting possible effects. While tempting, the fact that this pattern frequently emerges in studies of experimental and nonexperimental (e.g., correlational) research suggests it is not due to differences in the intensiveness of programs. Suggestions to that effect as explanations in particular cases need to be investigated and evidenced rigorously. Moreover, it is often larger programs that are more intensive rather than the reverse. Because of the ubiquity of patterns indicative of publication bias regardless of the experimental vs nonexperimental nature of studies, many reasonably take it for granted that it is evidence of bias.

²⁶ Dulal, Sophiya, Frédérique Liégeois, David Osrin, Adam Kuczynski, Dharma S. Manandhar, Bhim P. Shrestha, Aman Sen, Naomi Saville, Delan Devakumar and Audrey Prost. 2018. “Does Antenatal Micronutrient Supplementation Improve Children’s Cognitive Function? Evidence from the Follow-Up of a Double-Blind Randomised Controlled Trial in Nepal.” *BMJ Global Health* 3.

²⁷ Behrens, Annika, Elmar Graessel, Anna Pendergrass, and Carolin Donath. 2020. “Vitamin B—Can it Prevent Cognitive Decline? A Systematic Review and Meta-Analysis.” *Systematic Reviews* 9.

²⁸ Ali, Hasmot, Jena Hamadani, Sucheta Mehra, Fahmida Tofail, Md Imrul Hasan, Saijuddin Shaikh, Abu Ahmed Shamim, Lee S-F Wu, Keith P. West, Jr. and Parul Christian. 2017. “Effect of Maternal Antenatal and Newborn Supplementation with Vitamin A on Cognitive Development of School-Aged Children in Rural Bangladesh: A Follow-Up of a Placebo-Controlled, Randomized Trial.” *The American Journal of Clinical Nutrition* 106(1): 77-87.

²⁹ Welch, Vivian A., Elizabeth Ghogomu, Alomgir Hossain, Alison Riddle, Michelle Gaffey, Paul Arora, Omar Dewidar, Rehana Salam, Simon Cousens, Robert Black, T. Déirdre Hollingsworth, Sue Horton, Peter Tugwell, Donald Bundy, Mary Christine Castro, Alison Elliott, Henrik Friis, Huong T. Le, Chengfang Liu, Emily K. Rousham, Fabian Rohner, Charles King, Erliyani Sartono, Taniawati Supali, Peter Steinmann, Emily Webb, Franck Wieringa, Pattanee Winnichagoon, Maria Yazdanbakhsh, Zulfiqar A. Bhutta and George Wells. 2019. “Mass Deworming for Improving Health and Cognition of Children in Endemic Helminth Areas: A Systematic Review and Individual Participant Data Network Meta-Analysis.” *Systematic Review* 15(4).

³⁰ Taylor-Robinson, David C., Nicola Maayan, Karla Soares-Weiser, Sarah Donegan and Paul Garner. 2015. “Deworming Drugs for Soil-Transmitted Intestinal Worms in Children: Effects on Nutritional Indicators, Haemoglobin, and School Performance.” *Cochrane Database of Systematic Reviews* 7; Welch, Vivian A., Elizabeth Ghogomu, Alomgir Hossain, Shally Awasthi, Zulfi Bhutta, Chisa Cumberbatch, Robert Fletcher, Jessie McGowan, Shari Krishnaratne, Elizabeth Kristjansson, Salim Sohani, Shalini Suresh, Peter Tugwell, Howard White, and George Wells. 2016. “Deworming and Adjuvant Interventions for Improving the Developmental Health and Well-Being of Children in Low- and Middle-Income Countries: A Systematic Review and Network Meta-Analysis.” *Campbell Systematic Reviews* 7.

³¹ Zimmerman, Annie, Emily Garman, Mauricio Avendano-Pabon, Ricardo Araya, Sara Evans-Lacko, David McDaid, A-La Park, Philipp Hessel, Yadira Diaz, Alicia Matijasevich, Carola Ziebold, Annette Bauer, Cristiane Silvestre Paula and Crick Lund. 2020. “The Impact of Cash Transfers on Mental Health in Children and Young People in Low-Income and Middle-Income Countries: A Systematic Review and Meta-Analysis.” *BMJ Global Health* 6.

³² Manley, James, Yarlina Balarajan, Shahira Malm, Luke Harman, Jessica Owens, Sheila Murthy, David Stewart, Natalia Elena Winder-Rossi and Atif Khurshid. 2020. “Cash Transfers and Child Nutritional Outcomes: A Systematic Review and Meta-Analysis.” *BMJ Global Health* 5.

- ³³ Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling and Lisa Sanbonmatsu. 2013. “Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity.” *American Economic Review* 103(3): 226-231.
- ³⁴ Ericsson, Malin, Cecilia Lundholm, Stefan Fors, Anna K. Dahl Aslan, Catalina Zavala, Chandra A. Reynolds and Nancy L. Pedersen. 2017. “Childhood Social Class and Cognitive Aging in the Swedish Adoption/Twin Study of Aging.” *PNAS* 114(27): 7001-7006; Kendler, Kenneth S., Eric Turkheimer, Henrik Ohlsson, Jan Sundquist and Kristina Sundquist. 2015. “Family Environment and the Malleability of Cognitive Ability: A Swedish National Home-Reared and Adopted-Away Cosibling Control Study.” *PNAS* 112(15): 4612-4617.
- ³⁵ Odenstad, A., A. Hjern, F. Lindblad, F. Rasmussen, B. Vinnerljung and M. Dalen. 2008. “Does Age at Adoption and Geographic Origin Matter? A National Cohort Study of Cognitive Test Performance in Adult Inter-Country Adoptees.” *Psychological Medicine* 38(12).
- ³⁶ It might be argued that this is due to Korea’s extensive social welfare system for orphaned children, but this would suggest that government care is at least on par with adoption, and that other countries whose adoptees experience similar social support without removing adoptee selection are not doing as much as Korea. Another possibility is that the quality of environments yields diminishing returns *very* rapidly. None of these possibilities are particularly believable.
- ³⁷ See also McGue, Matt, Margaret Keyes, Anu Sharma, Irene Elkins, Lisa Legrand, Wendy Johnson and William G. Lacono. 2007. “The Environments of Adopted and Non-Adopted Youth: Evidence on Range Restriction from the Sibling Interaction and Behavior Study (SIBS).” *Behavior Genetics* 37: 449-462 and Halpern-Manners, Andrew, Helge Marahrens, Jenae M. Neiderhiser, Misaki N. Natsuaki, Daniel S. Shaw, David Reiss and Leslie D. Leve. 2020. “The Intergenerational Transmission of Early Educational Advantages: New Results Based on an Adoption Design.” *Research in Social Stratification and Mobility* 67.
- ³⁸ Posthuma, Daniëlle, Eco J. C. De Geus, Wim F. C. Baaré, Hilleke E. Hulshoff Pol, René S. Kahn and Dorret I. Boomsma. 2002. “The Association Between Brain Volume and Intelligence is Genetic in Origin.” *Nature Neuroscience* 5: 83-84; See also Posthuma, Daniëlle, Wim F. C. Baaré, Hilleke E. Hulshoff Pol, René S. Kahn, Dorret I. Boomsma and Eco J. C. De Geus. 2012. “Genetic Correlations Between Brain Volumes and the WAIS-III Dimensions of Verbal Comprehension, Working Memory, Perceptual Organization, and Processing Speed.” *Twin Research and Human Genetics* 6(2), Dreary, Ian J., Lars Penke and Wendy Johnson. 2010. “The Neuroscience of Human Intelligence Differences.” *Nature Reviews Neuroscience* 11: 201-211 and Jansen, Philip R., Mats Nagel, Kyoko Watanabe, Yongbin Wei, Jeanne E. Savage, Christiaan A. de Leeuw, Martijn P. van den Heuvel, Sophie van der Sluis and Daniëlle Posthuma. 2020. “Genome-Wide Meta-Analysis of Brain Volume Identifies Genomic Loci and Genes Shared with Intelligence.” *Nature Communications* 11.
- ³⁹ Plomin, Robert. 2014. “Genotype-Environment Correlation in the Era of DNA.” *Behavior Genetics* 44: 629-638; Ericsson et al. 2017.
- ⁴⁰ Smit, D. J. A., D. Posthuma, D. I. Boomsma and E. J. C. De Geus. 2005. “Heritability of Background EEG Across the Power Spectrum.” *Psychophysiology* 42(6): 691-697.
- ⁴¹ Wax, Amy. 2017. “The Poverty of the Neuroscience of Poverty: Policy Payoff or False Promise?” *Faculty Scholarship at Penn Law*. Available at https://scholarship.law.upenn.edu/faculty_scholarship/1711/.
- ⁴² “Press Release: Cash Support for Low-Income Families Impacts Infant Brain Activity” 2022.
- ⁴³ Assuming a population mean IQ of 100 with an SD of 15.
- ⁴⁴ UBI Center [@TheUBICenter]. 2022. *We’ve removed all posts on this study... [Tweet]*. Twitter. <https://twitter.com/TheUBICenter/status/1486383741161279490>; Jirari and Presteria 2022; Matthews 2022.
- ⁴⁵ Matthews 2022.