Analyzing Effects of Birth Order on Intelligence, Educational Attainment, Big Five, and Risk Aversion in an Indonesian Sample

Laura J. Botzet*1, Julia M. Rohrerb,c,d, and Ruben C. Arslane

*Department for Biological Personality Psychology, University of Goettingen, 37073 Goettingen, Germany; bInternational Max Planck Research School on the Life Course (LIFE), Max Planck Institute for Human Development, 14195 Berlin, Germany; cGerman Institute for Economic Research (DIW Berlin), 10117 Berlin, Germany; dDepartment of Psychology, University of Leipzig, 04109 Leipzig, Germany; eCenter for Adaptive Rationality, Max Planck Institute for Human Development, 14195 Berlin, Germany

R. C. Arslan developed the study concept. L. J. Botzet and R. C. Arslan contributed to the data analysis. All authors contributed to the interpretation of the results. L. J. Botzet drafted the manuscript and J. M. Rohrer and R. C. Arslan provided critical revisions. All authors approved the final version of the manuscript for submission.

1 To whom correspondence should be addressed.

Laura J. Botzet
Institute of Psychology, Biological Personality Psychology, Goßlerstraße 14, 37073 Göttingen, Germany
E-mail: laura.botzet@stud.uni-goettingen.de
Phone: +49 17654745686
Abstract

Few studies have examined birth order effects on personality in countries that are not Western, educated, industrialized, rich, and democratic (WEIRD). However, theories have generally suggested interculturally universal family dynamics as the mechanism behind birth order effects, and prominent theories such as resource dilution would even predict stronger effects in poorer countries. Here, we investigate a subset of up to 11,188 participants of the Indonesian Family Life Survey, an ongoing representative panel study, to investigate whether later-born siblings differ from earlier-borns in intelligence, educational attainment, Big Five personality traits, and risk aversion. Analyses were performed using within-family designs in mixed-effects models. In model comparisons we tested for linear and non-linear birth order effects as well as for possible interactions of birth order and sibship size. Our estimated effect sizes are consistent with the emerging account of birth order as having relatively little impact on intelligence, Big Five, and risk aversion; and they exclude recent estimates from WEIRD populations based on large sample sizes. Considering educational attainment, we found a nonlinear pattern that was not robust to the imputation of missing data and in any case not aligned with trends in WEIRD countries. Overall, the small effects of birth order reported in other studies appear to be culturally specific.

Keywords: Birth Order Research, Intelligence, Educational Attainment, Personality, Risk Aversion

This article contains supporting information online at https://osf.io/v2n6q/ and https://laurabotzet.github.io/birth_order_ifls/.
Introduction

Balinese names immediately reveal a person's birth order: First-borns are called Wayan, second-borns Made, and so on. Given the everyday salience of sibling ranks, one might expect particularly pronounced birth order effects in Bali. Previous birth order research on outcomes such as intelligence, educational attainment, and personality, has almost exclusively focused on “WEIRD” populations – populations from Western, educated, industrialized, rich, democratic countries (Heinrich, Heine, & Norenzayan, 2010). In helping to understand human universality and variability we need to move our focus from WEIRD samples to more diverse populations (Rad, Martingano, & Ginges, 2018).

Prominent theoretical accounts of birth order effects like resource dilution (Blake, 1981), the confluence model (Zajonc & Marcus, 1975), and the family niche model (Sulloway, 1996) have taken no explicit stance on the potential for cultural specificity. The resource dilution theory focuses on the fact that with each additional child, parental resources need to be shared among more offspring. While the first-born child can enjoy undiluted parental resources until the first sibling arrives, later-borns have to share from the very start - thus receiving less support for their intellectual development. This leads to a decrease in intelligence by birth order position (Blake, 1981). The confluence model argues that earlier-born children grow up in a more stimulating intellectual environment than their younger siblings because first-borns interact mostly with adults in their early development phase, leading to a decrease in intelligence by birth order position (Zajonc & Marcus, 1975). The family niche model assumes that siblings compete for parental investment (Trivers, 1985) and therefore develop strategies to increase parental attention by trying to fill different niches in one family (Sulloway, 1996). The first-born takes the traditional niche resulting in higher values in the dominance aspect of extraversion, in neuroticism, in conscientiousness, and in the intellectual aspects of openness. The second-born takes the rebellious niche, resulting in
higher values in the sociability aspect of extraversion, in agreeableness, and in unconventional aspects of openness (Sulloway, 2001).

These theories do not discuss potential influences by culture and instead seem to imply universal family dynamics. This suggests that birth order effects should not be specific to WEIRD populations; otherwise, we would have to conclude that family dynamics are less universal than assumed. In fact, explanations like resource dilution would predict even stronger birth order effects when families are large and resources are few. Hence, non-WEIRD populations would offer the most favorable conditions for detecting birth order effects due to resource dilution.

The Republic of Indonesia – the world’s largest island country – is located in Southeast Asia. It is an interesting example to study birth order effects not only because it is the world’s 4th most populous country and the most populous Muslim-majority country, but also because it is home to a very diverse population that differs from WEIRD samples on which theoretical accounts of birth order effects were based. In 2015 the estimated population was about 258 million people (median age: 28.4, 49.65% female), with a total fertility rate of 2.5 children per woman and a life expectancy of 68.6 years (United Nations, 2015). In 2010 87.18% of the total population was Muslim, 9.87% was Christian, 1.69% was Hindu, 0.72% was Buddhist, and 0.54% believed in another religion or did not believe in any religion (Badan Pusat Statistik, 2010). According to the census, in 2010 there were over 300 ethnic groups in Indonesia. 40.22% of the total population was Javanese, 15.50% was Sundanese, and 44.28% belonged to one of the other ethnic groups (each less than 5%; Badan Pusat Statistik, 2010). Based on data from the UNESCO Institute of Statistics, the literacy rate in 2015 for people aged 15 years and older was 95.40% and the mean number of years in school was 7.9 (United Nations Educational, Scientific and Cultural Organization, 2015).
Previous Findings Regarding Birth Order Effects

*Intelligence and Educational Attainment.*

Effects of birth order on intelligence, educational attainment, Big Five personality traits, and risk aversion have been studied extensively. The clearest evidence has emerged regarding intelligence and educational attainment. Several studies showed a slight decline in intelligence from earlier-borns to later-borns (first study to show these results: Netherlands: Belmont & Marolla, 1973; additional but not complete list of recent studies: Germany, Great Britain und United States of America: Rohrer, Egloff, & Schmukle, 2015; Norway: Bjerkedal, Kristensen, Skjeret, & Brevik, 2007; Black, Devereux, & Salvanes, 2011; Sweden: Barclay, 2015a; United States of America: Heiland, 2009; but see also Wichman, Rodgers, & MacCallum, 2006, and Damian & Roberts, 2015 for a critical assessment of the relevance of these effects). Likewise, several studies suggest a corresponding decline in educational attainment from earlier-borns to later-borns (e.g., Germany: Härkönen, 2014; Great Britain: Booth & Kee, 2009; Norway: Black, Devereux, & Salvanes, 2005; Kristensen & Bjerkedal, 2010; Sweden: Barclay, 2015b; United States of America: Behrman & Taubman, 1986; de Haan, 2010; Kantarevic & Mechoulan, 2006).

However, all of the previously cited studies used samples from WEIRD countries. For non-WEIRD countries, the evidence is more mixed. Regarding birth order and intelligence, studies have reported negative effects (Indonesia: Calimeris & Peters, 2017; Kenya: Munroe & Munroe, 1983; Zimbabwe: Wilson, Mundy-Castle, & Panditji, 1990), a parabolic relationship (Colombia: Velandia, Grandon, & Pagem, 1978; Israel: Davis, Cahan, & Bashi, 1976), or no effects at all (Kuwait: Abdel-Khalek & Lynn, 2008; see Table S1 in the supplemental material for a more detailed summary of these studies). Regarding birth order and educational attainment, studies have reported...
achievement, studies have reported both parabolic (Turkey: Dayioğlu, Kirdar, & Tansel, 2009) and positive relationships (Bangladesh: Park & Chung, 2012; Brazil: Emerson & Souza, 2008; Philippines: Ejrnæs & Pörtner, 2004; 12 African countries: Tenikue & Verheyden, 2010; for more information see Table S1 in the supplemental material).

**Big Five**

Considering the effects of birth order on the Big Five, even though there is empirical support for some of the effects predicted by the family niche model on the basis of WEIRD samples (e.g., Belgium: Saroglou & Fiasse, 2003; Canada: Paulhus, Trapnell, & Chen, 1999; New Zealand: Healey & Ellis, 2007; United States of America: Michalski & Shackelford, 2002) a reconstruction of Sulloway’s meta-analysis from the year 1996 could not replicate the empirical patterns found earlier (Townsend, 2000). In line with this failure to replicate, more recent studies with large sample sizes resulted in at best weak support for the family niche model. For example, in a sample of 377,000 U.S. high school students, Damian and Roberts (2015) found that first-borns tended to be higher in the dominance aspect of extraversion, more conscientious, and less agreeable. However, the correlations found in this study were very small, which led the authors to the conclusion that effects of birth order on Big Five personality traits were negligible. In a study combining data from Germany, the US, and Great Britain (n = 20,186) Rohrer et al. (2015) found that birth order did not have a significant effect on extraversion, neuroticism, conscientiousness, or agreeableness, but reported a negative effect of birth order on intellect which is in line with findings regarding intelligence and educational attainment. Thus, the overall evidence for Big Five in WEIRD countries suggests that there are at best weak birth order effects. In contrast to this vast body of literature, there is little research on the effects on Big Five in non-WEIRD countries. Previous studies were limited to India, only considered very specific outcomes, and used comparably small
samples (Begum, Banu, Jahan, & Begum, 1981; Kaur & Dheer, 1982; Sethi & Gupta, 1973; Sharma 1987; see Table S1).

**Risk Taking**

Lastly, moving beyond the Big Five, there is some support for a positive effect of birth order on risk taking based on WEIRD samples (United States of America: Argys, Rees, Averett, & Witoonchart, 2006; Averett, Argys, & Rees, 2011; Meta-analysis: Sulloway & Zweigenhaft, 2010). However, Eisenmann (1987) stated that first-born males tended to choose the riskier option, and Wang, Kruger, & Wilke (2009) noted that the relationship between birth order and risk preference differed for varying risk domains - both studies used samples from the United States of America. A recent study based on three large samples (two German samples and an international database of explorers and revolutionaries) found no robust effect of birth order on risk taking (Lejarraga, Frey, Schnitzlein, & Hertwig, 2019). To our knowledge, no evidence for or against birth order effects on risk taking or risk aversion in non-WEIRD populations exists.

**The Current Study**

Taken together, birth order effects in WEIRD countries have been investigated extensively, but little is known whether similar patterns (i.e., effects on intelligence and educational attainment, lack of strong effects on Big Five) occur in the rest of the world. Because recent birth order research suggests that at least some of the confusion about the effects of birth order on Big Five had to do with suboptimal methods (risk of overfitting owing to small samples, flexible model specification, and post-hoc theorising; see Rohrer, Egloff, & Schmukle, 2017), it seems wise to implement best practices to avoid these potential pitfalls. In our study, we aim to fill the research gap regarding birth order outside of the WEIRD world while adhering to best practices in birth order research. In particular, we kept all analyses
straightforward and comparable to recent work on WEIRD populations, applied appropriate control for sibship size and conducted extensive robustness checks. On the basis of the assumption that previously reported birth order effects generalize, we predicted the following: Intelligence, educational attainment and intellect decrease with higher birth order while extraversion, neuroticism, conscientiousness, agreeableness, and risk aversion remain unaffected. Analyses were not pre-registered; however, they are fully in line with earlier studies on the topic (Rohrer et al., 2015), and extensive robustness checks are provided.

Method

Data

Our data comes from RAND’s Indonesian Family Life Survey (IFLS), an ongoing longitudinal study with 50,148 individuals living in Indonesia. Since 1993 five waves have been administered (Frankenberg & Thomas, 2000; Strauss, Beegle, Sikoki, Dwiyanto, Herawati, & Witoelar, 2000; Strauss, Witoelar, Sikoki, 2016; Strauss, Witoelar, Sikoki, & Wattie, 2009). For the first wave, a sample of households that represented about 83% of the Indonesian population was approached. In the following waves every household and all split-off households were contacted. All analyses reported in this study were run on data based on this representative national panel study. Hence, we had no control over the exact sample size, but with \( N = 11,188 \), the sample size is comparable to or even larger than samples from recent literature on birth order effects.

In each wave, women aged 15 to 49 answered questions about their pregnancy and marriage history. These questions included information about the number and order of pregnancies as well as the birth date and gender of each child. Overall, 15,983 women reported 49,868 pregnancies. Marriage history allowed us to approximately infer the identity of the father. Based on this data we were able to construct full sibling order (based on the same mother and father) for 42,682 individuals.
In keeping with earlier studies we excluded families with multiple births or only children. The designs suitable to investigate birth order effects - comparison among children with the same number of siblings, or within-family comparisons - are not suitable to investigate whether only children differ from other children. In particular, when investigating only children, great care needs to be taken to control for systematic differences between families of different sizes. For analyses of birth order effects, we could only include individuals who participated in the fifth wave of the IFLS. A summary of the inclusion process for the full sibling birth order is shown in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Data Sampling and Exclusion Criteria</th>
<th>Number of individuals remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>42,682</td>
</tr>
<tr>
<td>No multiple births</td>
<td>38,589</td>
</tr>
<tr>
<td>No only children</td>
<td>34,786</td>
</tr>
<tr>
<td>Still alive</td>
<td>32,116</td>
</tr>
<tr>
<td>Any outcome data</td>
<td>11,188</td>
</tr>
<tr>
<td>Intelligence</td>
<td>5,698</td>
</tr>
<tr>
<td>Interviewed for IFLS 5</td>
<td></td>
</tr>
<tr>
<td>Educational attainment</td>
<td>6,035</td>
</tr>
<tr>
<td>Big Five</td>
<td>5,805</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>5,232 / 5,492</td>
</tr>
</tbody>
</table>

Notes. For the intelligence sample we only reported individuals who completed all intelligence tests. For risk aversion we report two numbers (risk A/risk B) because the sample size for the measurements differed slightly. IFLS 5: Indonesian Family Life Survey Wave 5 (2014)

Survey materials and data are openly available on the IFLS website www.rand.org/labor/FLS/IFLS.html (Frankenberg & Thomas, 2000; Strauss et al., 2000; Strauss et al., 2016; Strauss et al., 2009). The necessary datasets to reproduce our
analyses are described on the Open Science Framework https://osf.io/v2n6g/. All our analyses and code are documented on our website https://laurabotzet.github.io/birth_order_ifls/ and available to download on the Open Science Framework. A codebook generated using the codebook package (Arslan, 2019) can be found at https://laurabotzet.github.io/birth_order_ifls/2_codebook.html.

Outcomes

We included intelligence, educational attainment, personality and risk aversion as outcomes in our main analyses. Additional analyses (detailed results reported on the supplementary website) included income, self employment, working category (e.g., unpaid family worker), working sector, and smoking behavior as outcomes. All outcomes are based on the fifth wave of the IFLS. Continuous outcomes were z-standardized ($M = 0, SD = 1$) to make effect sizes easier to compare. For a detailed description of all outcomes see Strauss et al. (2016).

Intelligence

In the fifth wave of IFLS five intelligence subtests were conducted. All respondents aged 15 or older were asked to take part in these tests. a) They answered a shortened version of a Raven’s matrices test that consisted of eight items. For each item, respondents were asked to identify the missing element out of six possible elements to complete a pattern. b) They were asked to count backwards from 100 in steps of 7s. c) They were given a delayed word recall test. Participants heard a list of 10 nouns and had to recall as many words as possible four to five minutes later. d) They were given an adaptive number series test in which each participant answered six out of 15 items. Each item showed a pattern of numbers with one missing value. Participants had to name the missing number (e.g., “7 – 8 – ? – 10”). The first three items were given to all participants. Based on the accuracy of the first
three responses a subsequent set of three items was chosen. A Rasch scoring model was used to identify a person’s ability for a given set of response patterns with varying difficulties. Thus a composite score was calculated for each participant (Strauss et al., 2016). e) All respondents aged 15 to 59 participated in a math test. Each respondent answered five multiple choice questions measuring mathematical abilities (three mathematical calculations and two math text problems).

**Years of Education**

Participants reported the highest educational level they had attended (elementary, junior high, senior high, university) and the highest grade they completed at this educational level. Based on these items we were able to reconstruct years of education.

**Big Five Personality Traits**

Personality was assessed with the Big Five Index 15 (BFI 15) for all respondents aged 15 or older. The BFI 15 is based on the Big Five Inventory-SOEP (Schupp & Gerlitz, 2008), which in turn is a very short version of the Big Five Inventory and precludes examining single personality facets (e.g., intellect). It was translated into Indonesian. All 15 items started with the phrase “I see myself as someone who ...”. Three items were asked for each of the Big Five personality dimensions: extraversion (e.g., “… is talkative.”), neuroticism (e.g., “… worries a lot.”), conscientiousness (e.g., “… does a thorough job.”), agreeableness (e.g., “… is considerate and kind to almost everyone.”), and openness (e.g. “… is original, comes up with new ideas.”). Participants expressed their agreement on a 5-point Likert scale (from 1 = “Disagree strongly” to 5 = “Agree strongly”). The BFI 15 is used in many large-scale surveys, the items were simply translated to Indonesian (details provided in Strauss et al., 2016). An earlier study by Wibowo, Yudiana, Reswara, and Jatmiko (2017) using the 44-item Big Five Inventory showed sufficient reliability (Cronbach’s alpha ranged between .69 and .85) but
limited validity (exploratory factor analysis revealed eight instead of five factors) in an Indonesian sample.

**Risk Aversion**

Risk aversion was assessed with an adaptive hypothetical lottery choice task for all respondents aged 15 or older. Analyses of the Mexican Family Life Survey had suggested that hypothetical lotteries yield similar results to lotteries that are paid out (Hamoudi, 2006). Two different sets of questions, A and B, were asked (randomized order across participants). The sets differed in the amount of the payoffs and the variance of their expected payoffs. Set B’s certain payoffs were higher than A’s. The uncertain payoffs in Set B had higher coefficients of variation than those in A, reflecting a higher risk-reward ratio (for a more detailed description see Ng, 2013). Many participants gave inconsistent responses across the two tasks and current research suggests that lottery tasks may be poor measures of individual differences in risk preferences compared to self-reports (Frey, Pedroni, Mata, Rieskamp, & Hertwig, 2017).

**Additional Outcomes**

To further investigate the effects of birth order on additional outcomes we included income in the last year, self employment (0 = no, 1 = yes), smoking behavior in the last year (0 = no, 1 = yes), category of work (six categories: casual worker in agriculture; casual worker not in agriculture; government worker; private worker; self employment; unpaid family worker; 0 = no, 1 = yes), and sector of work (eight sectors: agriculture, forestry, fishing, and hunting; construction; electricity, gas, and water; finance, insurance, real estate, and business service; manufacturing; mining and quarrying; social services; transportation, storage, and communication; 0 = no, 1 = yes). All of these outcomes were assessed in the fifth wave of the IFLS.
Analysis

Birth order and sibship size were calculated for full sibships (same father and mother) based on the maternal reports of pregnancy and marriage history of all women aged 15–49 who participated in an IFLS wave. In all models, we adjusted for the categorical effect of sibship size (effects of sibship size 2, 3, 4, 5, and over 5), self-reported gender, a third-order polynomial for age, and a family random effect to account for dependencies within sibships. We then tested for a linear birth order effect, for nonlinear effects (by testing categorical effects of birth order 1, 2, 3, 4, 5, and over 5), and for potential interactions with number of siblings, and went on to iteratively compare models to see whether each model improved upon the preceding one. Additionally, we reported estimates and confidence intervals of the birth order effects to further examine the influence of birth order on the outcomes and compare effect sizes to the existing literature.

We performed additional robustness analyses based on maternal sibships and maternal pregnancy order (including e.g., stillbirths) and we tested the effect of excluding all individuals in sibships bigger than five. Furthermore, we repeated all analyses after multiple imputation (for details see section about handling missing data). We state where results changed depending on the analysis approach and report all robustness analyses on the website.

We reported the results of our model comparisons for each outcome based on full sibling order. Because of the number of outcomes and owing to calls for more stringent significance cutoffs in empirical science, we set the significance threshold to .005 (Benjamin et al., 2017). We summarise results here briefly, and report them in full online (https://laurabotzet.github.io/birth_order_ifls/4_analyses.html).

All analyses were conducted in R, version 3.3.3 (R Core Team, 2013). We conducted within-family analyses by running mixed effects models using the lme4 package version 1.1–12 (Bates, Mächler, Bolker, & Walker, 2015).
Handling Missing Data

As in other large panel studies, the IFLS had systematic missing data. Some questions were only answered by a subset of participants (depending on age), some participants were absent for one of the waves, some mothers did not fill out the pregnancy questionnaire which we used to ascertain birth order, and some missing data were due to panel mortality. To impute birth order for individuals whose mothers had not filled out the pregnancy questionnaire, we computed variables which we called “naive birth order” and “naive sibling count.” This was simply the order according to birth year by people reporting the same mother. Pre-imputation analyses showed high agreement between the naive birth order and the full sibling order ($r = .91$, 99.5% CI: [.90, .91]), although systematic missing data are likely.

Following Grund, Lüdtke and Robitzsch (2017), we performed multiple imputation for multilevel data. We included the identity of the mother as a grouping variable, a third-order polynomial for age and the categorical interaction between sibling order and birth order. We let all focal variables predict each other and any variables which were correlated $r > .1$ (see Table S2 in the supplemental material for a list of all variables included). To impute full birth order and sibling order from the "naive" birth order and sibling order, we used the linear variables to regenerate the categorical interaction. We used the R packages MICE package version 3.7.0 (Buuren & Groothuis-Oudshoorn, 2010) and pan package version 1.6 (Schäfer, 1997) to impute the data. We generated 10 imputed datasets, ran all models 10 times, and aggregated estimates and standard errors using the mitml package version 0.3-7 (Grund, Robitzsch, & Lüdtke, 2018). To evaluate the quality of imputations we examined intraclass correlations, density plots, and trace plots. For further information see https://laurabotzet.github.io/birth_order_ifls/3_imputation.html.
Note on the Usage of the Term “Effect”

To remain consistent with previous studies, and to avoid clumsy language, we will talk about birth order effects. However, it should be noted that the causal identification of birth order effects is a non-trivial issue that has not been addressed in the literature to date. While adjustment for family size can rule out certain obvious confounders such as socio-economic status, other issues remain. For example, first-borns’ personality may affect parents’ decision to have more children (Jokela, 2010); hence, personality may actually causally affect birth order position. Such effects could result in birth order differences in personality which do not reflect birth order effects but rather the effect on children’s personality on family size. Furthermore parental age at birth and birth order position within-families are per definition confounded. Thus, birth order effects might be entangled with effects of parental age (Arslan, Penke, Iacono, McGue, & Johnson, 2014).

Results

Sample Analysis

We found that our main sample of people for whom birth order could be computed systematically differed from those for whom the required information was missing. Our main sample was 25.31 years younger and the percentage of female individuals was 2 percentage points higher. The main sample was also more intelligent ($d = 0.82$) and had more years of educational attainment ($d = 0.82$); it was more extraverted ($d = 0.09$), more neurotic ($d = 0.08$), less conscientious ($d = -0.19$), and less agreeable ($d = -0.13$); it also scored higher on openness ($d = 0.23$) and showed differences in risk aversion with inconsistent signs across the two measures (decreased risk aversion for risk A: $d = -0.10$; increased for risk B: $d = 0.08$). These differences were all significant (all $ps < .001$).
**Intelligence g-factor**

Based on the results of the Raven’s matrices test, the math test, the backwards counting task, the delayed word recall, and the adaptive number series from all individuals who took part in wave 5 of the IFLS, we computed a g-factor of intelligence. Based on a sample of participants who completed all five of the intelligence tests in wave 5 regardless of whether birth order information was available ($n = 27,526$), we ran a confirmatory factor analysis expecting one factor. The g-factor explained 30% of variance on average in the five intelligence measurements (Raven's matrices test: 42%, math test: 25%, backwards counting task: 20%, delayed word recall: 23%, adaptive number series: 40%).

**Descriptive Statistics**

Table 2 shows raw means, standard deviations, internal consistency, and a correlation matrix for age, gender, intelligence, educational attainment, extraversion, neuroticism, conscientiousness, agreeableness, and both risk aversion measurements based on our main sample with birth order information ($n = 11,188$). Note that the internal consistency measures for short-form scales, as used here, probably underestimate reliability (Eisenbarth, Lilienfeld, & Yarkoni, 2015). The correlation between age and intelligence ($r = -.10, 99.5\%\ CI: [-.13, -.06]$) and age and Big Five (extraversion: $r = .00 [-.03, .03]$, neuroticism: $r = -.13 [-.17, -.10]$, conscientiousness: $r = .24 [.20, .27]$, agreeableness: $r = .10 [.06, .13]$, and openness: $r = .00 [-.04, .04]$) were consistent with trends found in WEIRD samples (Nisbett, Aronson, Blair, Dickensen, Flynn, Halpern & Turkheimer, 2012; Roberts, Walton, & Viechtbauer, 2006). Being male was negatively correlated with extraversion ($r = -.13 [-.16, -.10]$) and risk aversion (risk A: $r = -.13 [-.17, -.10]$; risk B: $r = -.12 [-.15, -.08]$). Intelligence was positively correlated with extraversion ($r = .06 [.02, .10]$), and openness ($r = .08 [.05, .12]$), and slightly negatively correlated with neuroticism ($r = -.05 [-.07, -.02]$) and agreeableness ($r = -.04 [-.08, -.01]$). No consistent correlation pattern was visible for
intelligence and risk aversion (risk A: \( r = -.15 [-.18, -.11] \); risk B: \( r = .05 [.01, .08] \)). The correlations between the five dimensions of personality matched those found in a German sample (Hahn, Gottschling, & Spinath, 2012) with positive correlations between all personality dimensions except for neuroticism. The two risk aversion tasks correlated moderately with each other (\( r = .30 [.27, .34] \)).
Table 2

Means, Standard Deviations, Internal Consistency, and Correlation Matrix for Age, Gender, Intelligence, Educational Attainment, Big Five, and Risk Aversion.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>α</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Age</td>
<td>13.75</td>
<td>10.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Gender: Female</td>
<td></td>
<td></td>
<td>.51</td>
<td>-.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Intelligence</td>
<td>0.45</td>
<td>0.68</td>
<td>.61</td>
<td>-.10</td>
<td>-.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Years of education</td>
<td>11.40</td>
<td>3.49</td>
<td></td>
<td>.24</td>
<td>-.05</td>
<td>.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) E</td>
<td>3.49</td>
<td>0.68</td>
<td>.43</td>
<td>-.13</td>
<td>.06</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) N</td>
<td>2.72</td>
<td>0.66</td>
<td>.33</td>
<td>-.13</td>
<td>-.13</td>
<td>-.05</td>
<td>-.08</td>
<td>-.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) C</td>
<td>3.73</td>
<td>0.58</td>
<td>.39</td>
<td>.24</td>
<td>.02</td>
<td>-.03</td>
<td>.10</td>
<td>.07</td>
<td>-.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) A</td>
<td>3.85</td>
<td>0.51</td>
<td>.36</td>
<td>.10</td>
<td>.03</td>
<td>-.04</td>
<td>.03</td>
<td>.07</td>
<td>-.17</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) O</td>
<td>3.82</td>
<td>0.60</td>
<td>.46</td>
<td>.00</td>
<td>.08</td>
<td>.08</td>
<td>.15</td>
<td>.17</td>
<td>-.07</td>
<td>.27</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Risk A</td>
<td>3.29</td>
<td>1.45</td>
<td></td>
<td>-.06</td>
<td>-.13</td>
<td>-.14</td>
<td>-.19</td>
<td>-.01</td>
<td>.04</td>
<td>-.04</td>
<td>-.02</td>
<td>-.08</td>
<td></td>
</tr>
<tr>
<td>(11) Risk B</td>
<td>4.29</td>
<td>1.31</td>
<td></td>
<td>-.01</td>
<td>-.12</td>
<td>.05</td>
<td>-.02</td>
<td>.00</td>
<td>.02</td>
<td>-.01</td>
<td>-.03</td>
<td>.30</td>
<td></td>
</tr>
</tbody>
</table>

Notes. Data is based on the main sample ($n = 11,188$). $\alpha$: Cronbach’s alpha as a measure of internal consistency (Age, gender, and years of education were only measured with one item and risk aversion is based on an adaptive hypothetical lottery choice task. Therefore Cronbach’s alpha is missing for these measures.). E: Extraversion, N: Neuroticism, C: Conscientiousness, A: Agreeableness, O: Openness. Bold numbers are significant at $p < .005$. 
Birth Order Effects

Analyses showed no significant improvement of model fit when including birth order as a predictor for intelligence, Big Five, and risk aversion (all $p > .02$). All results for model comparisons are summarized in Table 3.

Figure 1 shows linear effects of birth order on all main outcomes, with effect sizes based on z-standardized outcomes. The dotted line shows an estimate of the linear birth order effect on intelligence ($d = -0.14$) and the grey area shows the 99.5% confidence interval [-0.20, -0.07] based on a within-family analysis in a Western sample (sample reported in Rohrer et al., 2015; an additional within-family analysis was run to estimate the linear effect to ensure that the comparison was meaningful). Not only do the confidence intervals of our estimates include zero, they also exclude the estimate for intelligence based on a large WEIRD sample.

We found evidence for a nonlinear effect of birth order on educational attainment ($X^2(4, N = 6035) = 21.48, p < .001$). A closer look at the categorical effect showed a checkmark-shaped pattern, indicating the same amount of educational attainment for the first-borns compared to the second-borns (estimation of effect: -0.06, 99.5%-CI: [-0.12, -0.005]), to the third-borns (0.02 [-0.06, 0.09]), and to the fourth-borns (0.09 [-0.01, 0.19]). The fifth-borns had more educational attainment compared to the first-borns (0.14 [0.01, 0.26]). Differences between sixth-and later-borns compared to first-borns were not significant (estimation of effect: 0.08 [-0.05, 0.21]).

Including the interaction of sibship size and birth order as a predictor did not improve any of the models significantly (all $p > 0.06$). In Figure 2, we show the interaction of sibship size and birth order (adjusted for gender, a third-order polynomial effect of age and maternal identity).

Additional analyses showed no significant effects of birth order on income (linear compared to covariates-only model: $p = .51$, categorical compared to linear model: $p = .30$,
interaction compared to categorical model: $p = .56$), self employment (linear: $p = .80$, categorical: $p = .86$, interaction: $p = .62$), and smoking behavior (linear: $p = .9996$, categorical: $p = .54$, interaction: $p = .65$). We found no birth order effects in the seven analyses on working category (linear: all $p > .26$, categorical: all $p > .32$, interaction: all $p > .17$) or in the eight analyses on working sector (linear: all $p > .14$, categorical: all $p > .06$, interaction: all $p > .40$). Sample sizes for some of these analyses might have been too small to detect birth order effects (income: $n = 2,477$, self employment: $n = 3,763$, smoking behavior: $n = 6,104$, working category: $n = 3,763$, working sector: $n = 3,610$). For more details on additional outcomes see Table S3 in the supplemental material and 

Additional robustness analyses based on maternal sibship and maternal pregnancy order did not differ from the analyses reported here. For all details on the robustness analyses see https://laurabotzet.github.io/birth_order_ifls/4_analyses_robust.html. All analyses based on the imputed dataset yielded non-significant results (all $p < .27$). Contrary to our main analyses, we found no evidence for a linear or a nonlinear effect of birth order on educational attainment (linear compared to covariates-only model: $p = .45$, categorical compared to linear model: $p = .96$). All details for the analyses based on the imputed dataset can be found online:

Table 3

*Model Comparisons for Birth Order Effects on Intelligence, Educational Attainment, Big Five, and Risk Aversion.*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
<th>$X^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence</td>
<td>Covariates</td>
<td>14247</td>
<td>14320</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,698$</td>
<td>Linear birth order</td>
<td>14247</td>
<td>14327</td>
<td>1.79</td>
<td>1</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>14253</td>
<td>14359</td>
<td>2.11</td>
<td>4</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>14253</td>
<td>14440</td>
<td>5.62</td>
<td>10</td>
<td>.85</td>
</tr>
<tr>
<td>Educational Attainment</td>
<td>Covariates</td>
<td>13356</td>
<td>13430</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 6,035$</td>
<td>Linear birth order</td>
<td>13352</td>
<td>13432</td>
<td>6.34</td>
<td>1</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>13338</td>
<td>13446</td>
<td>21.48</td>
<td>4</td>
<td>&lt; .001</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>13353</td>
<td>13528</td>
<td>4.93</td>
<td>10</td>
<td>.90</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Covariates</td>
<td>16605</td>
<td>16679</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,805$</td>
<td>Linear birth order</td>
<td>16603</td>
<td>16683</td>
<td>4.06</td>
<td>1</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>16600</td>
<td>16706</td>
<td>11.43</td>
<td>4</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>16617</td>
<td>16790</td>
<td>3.21</td>
<td>10</td>
<td>.98</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>Covariates</td>
<td>16235</td>
<td>16308</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,805$</td>
<td>Linear birth order</td>
<td>16237</td>
<td>16317</td>
<td>0.11</td>
<td>1</td>
<td>.74</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>16234</td>
<td>16340</td>
<td>11.14</td>
<td>4</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>16234</td>
<td>16416</td>
<td>10.72</td>
<td>10</td>
<td>.38</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Covariates</td>
<td>16366</td>
<td>16440</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,805$</td>
<td>Linear birth order</td>
<td>16364</td>
<td>16444</td>
<td>3.79</td>
<td>1</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>16372</td>
<td>16478</td>
<td>0.80</td>
<td>4</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>16375</td>
<td>16548</td>
<td>16.83</td>
<td>10</td>
<td>.08</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>Covariates</td>
<td>16314</td>
<td>16387</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,805$</td>
<td>Linear birth order</td>
<td>16316</td>
<td>16396</td>
<td>0.05</td>
<td>1</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>16321</td>
<td>16428</td>
<td>2.83</td>
<td>4</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>16334</td>
<td>16508</td>
<td>6.65</td>
<td>10</td>
<td>.76</td>
</tr>
<tr>
<td>Openness</td>
<td>Covariates</td>
<td>15673</td>
<td>15746</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,805$</td>
<td>Linear birth order</td>
<td>15675</td>
<td>15755</td>
<td>0.34</td>
<td>1</td>
<td>.56</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>15678</td>
<td>15785</td>
<td>4.72</td>
<td>4</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>15692</td>
<td>15865</td>
<td>6.55</td>
<td>10</td>
<td>.77</td>
</tr>
<tr>
<td>Risk A</td>
<td>Covariates</td>
<td>14385</td>
<td>14457</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n = 5,232$</td>
<td>Linear birth order</td>
<td>14386</td>
<td>14465</td>
<td>0.05</td>
<td>1</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Categorical birth order</td>
<td>14392</td>
<td>14497</td>
<td>2.06</td>
<td>4</td>
<td>.73</td>
</tr>
<tr>
<td></td>
<td>Interaction</td>
<td>14395</td>
<td>14565</td>
<td>17.69</td>
<td>10</td>
<td>.06</td>
</tr>
<tr>
<td>Risk B</td>
<td>Covariates</td>
<td>14888</td>
<td>14961</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n = 5,492</td>
<td>Linear birth order</td>
<td>14890</td>
<td>14970</td>
<td>0.002</td>
<td>1</td>
<td>.96</td>
</tr>
<tr>
<td>Categorical birth order</td>
<td>14896</td>
<td>15001</td>
<td>2.66</td>
<td>4</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>14901</td>
<td>15073</td>
<td>14.83</td>
<td>10</td>
<td>.14</td>
<td></td>
</tr>
</tbody>
</table>

Notes. The covariates model included the categorical effect of sibship size (effects of sibship size 2, 3, 4, 5, and over 5), self-reported gender, a third-order polynomial for age, and a family random effect. The linear birth order model added birth order as a linear predictor, the categorical model added birth order as a categorical predictor (effects of birth order 1, 2, 3, 4, 5, and over 5), the interaction model included the interaction of the categorical birth order and the categorical sibship size. The linear birth order model was compared to the covariates model, the categorical birth order model was compared to the linear birth order model and the interaction model was compared to the categorical birth order model. Sample sizes differed slightly because not all individuals completed all outcome measurements. $AIC =$ Akaike information criterion, $BIC =$ Bayesian information criterion, $df =$ degrees of freedom.
Figure 1. Effect size estimates in standard deviations for the linear effect of birth order. Effect size and 99.5% confidence intervals from linear mixed effects models with the categorical effect of sibship size (effects of sibship size 2, 3, 4, 5, and over 5), birth order, gender, and a third-order polynomial effect of age as fixed effects and maternal identity as random effect are shown. The dotted line shows an estimate of the linear birth order effect on intelligence $(d = -0.14)$ and the grey area shows the 99.5% confidence interval $[-0.20, -0.07]$ based on a within-family analysis in a Western sample (Rohrer et al., 2015).
Figure 2. Interaction effects of sibship size and birth order position on intelligence, Big Five, and risk aversion. Predicted mean scores and 99.5% confidence intervals from linear mixed effects models with sibship size, birth order, gender, and a third-order polynomial effect of age as fixed effects and maternal identity as random effect are displayed for (a.) intelligence, (b.) educational attainment, (c.) extraversion, (d.) neuroticism, (e.) conscientiousness, (f.) agreeableness, (g.) openness, and (h. – i.) risk aversion. All outcome measurements are z-standardized based on the full sample. Numbers in parentheses show sample size.
Discussion

We found no birth order effects on intelligence, agreeableness, conscientiousness, extraversion, neuroticism, openness, or risk aversion, regardless of whether we included birth order as a continuous or categorical predictor, or whether we considered its interaction with sibship size. Model comparisons supported a small nonlinear effect of birth order on educational attainment in form of a checkmark-shaped pattern. However, this effect did not emerge when missing values were imputed.

Our results were consistent with null effects on agreeableness, conscientiousness, extraversion, neuroticism, and risk aversion found in WEIRD populations (Damian & Roberts, 2015; Lejarraga et al., 2019; Rohrer et al., 2015). Yet, we found no effect of birth order on intelligence nor openness in contrast to the small negative estimates reported for WEIRD populations (Barclay, 2015b; Damian & Roberts, 2015; Rohrer et al., 2015; Rohrer et al., 2017). Considering educational attainment, our results were sensitive to the imputation of missing data and one should thus be careful in their interpretation. In WEIRD samples higher birth order is related to lower educational attainment (Black et al., 2005; Booth & Kee, 2009; Härkönen, 2013; Kristensen & Bjerkedal, 2010), even in fully adopted sibling groups (Barclay, 2015). However, in this Indonesian sample, higher birth order was related to higher educational attainment, if there was any association at all.

Thus, our results are inconsistent with predictions from Blake’s resource dilution model (Blake, 1981), Zajonc’s confluence model (Zajonc & Markus, 1975), and Sulloway’s sibling roles (Sulloway, 1996). These theories made no allowance for fundamental cultural differences in family dynamics. Given that Indonesia is a poorer country with larger families than most previously studied countries, resource dilution would have predicted larger effects on average. Observed effects fell short of initial predictions from theories on family dynamics that were based on the largest, best evidence available in the WEIRD world. It is therefore appropriate to look for other explanations. Perhaps birth order influences the social and
parental expectations for first-born children in some countries, through remnants of Western
cultural norms like primogeniture, or through policies such as parental leave (Barclay, 2015b).
Children’s traits may then adapt due to external influences on their educational and
occupational choices, such as a father expecting a first-born to take over the family business
(Barclay, Hällsten, & Myrskylä, 2017). Such an indirect effect would be consistent with the
generally observed small average birth order effects, as well as with the absence of those
effects in a culturally different country.

It is important to note that our main sample of people for whom birth order could be
computed systematically differed from those for whom the required information was missing. The large differences in intelligence and educational attainment are in part due to age
differences. The main sample for whom birth order could be computed was 25 years younger
than the rest of the sample, likely because birth order could only be computed for individuals
whose mother took part in one of the waves of the IFLS, whereas outcome measurements
were available for all participants. In the full sample, age correlated negatively with intelligence
\( r = -0.39, p < 0.001 \) and educational attainment \( r = -0.39, p < 0.001 \), thus, differences in
intelligence and educational attainment might actually reflect age differences. The main
sample was 0.59 [0.58, 0.60] standard deviations more intelligent and attained 3.11 [3.06, 3.17] more years of educational attainment on average. Controlling for a third-order
polynomial effect of age and a linear effect of our naive sibship size measurement reduced
the adjusted mean differences to 0.33 [0.32, 0.33] standard deviations for intelligence and
1.94 [1.93, 1.96] years of educational attainment. Since the effects of interest for this study -
namely birth order effects - are within-family effects and since we controlled for potential age
effects, the potential issues introduced by this discrepancy might not be grave. Nevertheless,
generalizability of our findings might be limited to the younger, more intelligent and more
educated Indonesian generation. Especially the finding of a check mark pattern for education
should be interpreted cautiously, because it was not reproduced in the analyses of the
multiple imputed datasets. This difference may reflect model error in the imputation, or a lack of generalizability to the full sample.

In our data, we found no evidence for birth order effects on various outcomes related to type of employment and work sector, corroborating the emerging narrative that birth order is generally not a very important predictor of life outcomes.

Limitations

Because of limitations of the available data, there are several alternative explanations in favor of the existence of birth order effects that we could not rule out.

First, the study population was limited. Our analyses of intelligence, Big Five, and risk aversion do not include individuals younger than 15 years - but according to Sulloway (2010), effects of birth order on personality should be especially visible during childhood and adolescence. Thus, we cannot rule out that we missed substantial birth order effects among younger Indonesians; we can only say that if they existed, they had dissipated with age.

Furthermore, the conclusions of our study are of course limited to present-day Indonesia and need not generalize to other (WEIRD or non-WEIRD) countries or across time. While the absence of birth order effects in present-day Indonesia casts doubt on broad theories which claim that such effects emerge universally, it does of course not rule out that birth order effects emerge under different societal conditions.

Second, the outcome measures were limited. Intelligence was measured with a g-factor based on different subtests with no particular theoretical background. On the other hand, Johnson, Bouchard, Krueger, McGue, and Gottesman (2004) showed that g-factor batteries correlated highly with each other, even though the batteries measured different mental abilities. Thus, our g-factor is still a valid measurement for general intelligence. The brief Big Five measure used made it impossible to test for effects on narrower facets. While
the hunt for birth order effects on facets in WEIRD samples has not brought up consistent patterns, it is of course possible that such patterns exist in Indonesia.

Taking the critique of measurement even further, one could argue that self-reports are generally not suitable for detecting birth order effects. This criticism certainly fully applies to our analyses of the Big Five personality traits, but not necessarily to our measures of risk aversion and other reported outcomes (unless respondents systematically lied), and not to the assessment of intelligence. Sulloway (1999) suggested that the socially desirable responding of firstborns could cancel out existing birth order effects. This hypothesis could be tested using either other-reports of personality or behavioral outcomes that are not as easily affected by social desirability.

Again, it should be noted that there is not much evidence for this alternative explanation in WEIRD countries - older studies using personality comparisons made by other family members or even comparisons made by the targets themselves (e.g., ranking themselves among their siblings) are hard to interpret as they might be affected by stereotypes; the only study that we are aware of which uses other-reports by third parties does not provide much evidence for the expected effects (Jefferson, Herbst, & McCrae, 1998); and the few studies using alternative outcomes (i.e., behavior in economic games, Courtioll, Raymond, & Faurie, 2009; Salmon, Cuthbertson, & Figueredo, 2016) suffer from various quality issues. However, recent population-based studies of college major choices in Sweden found that earlier-borns were more likely to study engineering and medicine, while later-borns were more likely to study journalism, art, and business (Barclay et al., 2017). In addition, rationalist blogger Scott Alexander reported a sizable overrepresentation of first-borns among his readership which is heavily biased towards computer scientists. Such strong patterns regarding life choices in the absence of strong effects on intelligence and personality could perhaps be better explained through varying parental expectations, investment, and specific social norms. Perhaps, these major life choices occasionally have
hanger-on effects on the traits of intelligence and personality studied here. Importantly, the hanger-on effects would then not be found across cultures as expected if they were due to universal family dynamics, but rather fairly specific to choice environments where university education is common and choices are affected by parents.

Of course, the overall lack of good evidence is not the same as evidence for the lack of effects. As already stated in Rohrer et al. (2017), researchers who aim to venture down this path should ensure that they follow best practices to avoid wrong conclusions. In particular, given the high effort involved in collecting behavioral data or observer reports for a sufficiently large sample to detect the potentially subtle effects of birth order, we would like to encourage researchers to consider using a Registered Reports format, lest their efforts result in a negative finding that might be hard to publish.

**Birth Order Effect: Stereotypes vs. Reality?**

The lack of consistent birth order effects on personality in more recent empirical studies is often contrasted with the prevalence of strong stereotypes, in particular in popular media reporting about birth order research. In the spirit of stereotype accuracy, strong stereotypes might be interpreted as evidence for the existence of actual birth order effects. For example, Paulhus (2008) has argued that such stereotypes flourish because they have (at least) a kernel of truth. With respect to that notion, we would like to raise two questions. First, to which extent do such stereotypes indeed “flourish”? Second, what could such a “kernel of truth” be?

With respect to the prevalence of stereotypes, Herrera, Zajonc, Wieczorkowska, and Cichomski (2003) provided some evidence for stereotypes in line with Sulloway’s Family Niche Theory, with average ratings made by Stanford undergraduate students suggesting that last-borns are perceived as boldest, first borns as most intelligent, and so on. However, these average ratings do not tell us how many percent of the respondents actually held a
stereotypical view of birth order positions in line with any particular theory or how high inter-rater agreement was, making it hard to assess the underlying variability. And, once again, there is a lack of studies on these patterns in non-WEIRD countries. It would be interesting to have a better understanding of how reliable these stereotypes actually are (e.g., in comparison to stereotypes that have been researched more extensively, such as gender stereotypes) and to which extent they are idiosyncratic (e.g., people may have strong but divergent views about the personality of different birth order positions, possibly informed by their own family).

More knowledge about the strength and the consistency of these stereotypes could in turn inform reasoning about potential “kernels of truth.” If the stereotypes are held indeed almost universally, a more powerful cause seems plausible. If they are inconsistent and only on average slightly tilted in favor of the picture painted by models such as the Family Niche Theory, even a very small systematic factor could give rise to them, including past media reports about scientific findings.

For example, one could imagine that the birth order stereotypes reflect real birth order effects which occur in younger age, or used to occur in the past and are kept alive as part of cultural memory. It would also be possible that some of the stereotypes with respect to traits are a result of birth order effects in other domains - as mentioned above, there is some evidence for birth order effects on choice of college major (Barclay et al., 2017). They could also reflect “spurious” birth order associations resulting from within-family comparison. Such comparisons are confounded by age - for example, when comparing siblings within families, the firstborns will tend to be more conscientious as they are older and thus tend to be more mature (the supplementary material of Rohrer et al., 2017 provides empirical evidence for this pattern: https://osf.io/wu7h3/). Such comparisons, if made by parents, could also reflect changing interpretations and expectations regarding children’s behavior as parenting expertise increases.
It might be worth highlighting that a similar (if not stronger) discrepancy between stereotypes and actual personality may exist for only children. Dufner, Back, Oehme, and Schmukle (2019) contrasted laypeople’s expectations regarding narcissism in only children with actual data from a representative German panel study. While raters ascribed only children both higher narcissistic admiration and rivalry, no such trend was found when actually comparing narcissism between only children and children with siblings.

Taken together, the findings regarding birth order position and only children might indicate that laypeople (as well as psychological researchers) are prone to overestimate the role of “coarse” features of the childhood family environment (the absence or presence of siblings, one’s position among them) for long-term personality development. Similarly, it goes against many people’s expectation that the shared environment has repeatedly been shown to have little influence on most behavioral traits. Birth order is an environmental factor that is not shared within the family, but there is a similar pattern regarding the shared environment: Many people intuit that it must matter a great deal for adult personality, yet behavioral genetic studies have repeatedly shown that it contributes little (Turkheimer, 2000). Thus, it seems like our broad intuitions about how the environment affects our personality in the long run are often miscalibrated or too simplistic.

**Conclusion**

From a broader perspective, our study highlights the need to go beyond Western countries to test whether theoretical accounts implicitly or explicitly assumed to hold universally actually do operate across cultures. Robust discrepancies between findings from different cultural backgrounds can inform theory and elucidate mechanisms that drive effects – but only if researchers do not limit themselves to WEIRD populations.
Acknowledgements

We are grateful to the RAND corporation that conducts the Indonesian Family Life Survey and to the survey participants. Additionally, we would like to thank Deb Ain for her careful scientific editing. All remaining errors are ours.
References


Barclay, K. J. (2015b). Birth order and educational attainment: Evidence from fully adopted
sibling groups. *Intelligence, 48*, 109–122. doi:10.1016/j.intell.2014.10.009


doi:10.1017/S0730938400014702


doi:10.1080/00224545.1990.9924599