
The Confluence Model: Birth Order as a Within-Family or Between-Family Dynamic?

R. B. Zajonc

Stanford University

Frank J. Sulloway

University of California–Berkeley

The confluence model explains birth-order differences in intellectual performance by quantifying the changing dynamics within the family. Wichman, Rodgers, and MacCallum (2006) claimed that these differences are a between-family phenomenon—and hence are not directly related to birth order itself. The study design and analyses presented by Wichman et al. nevertheless suffer from crucial shortcomings, including their use of unfocused tests, which cause statistically significant trends to be overlooked. In addition, Wichman et al. treated birth-order effects as a linear phenomenon thereby ignoring the confluence model’s prediction that these two samples may manifest opposing results based on age. This article cites between- and within-family data that demonstrate systematic birth-order effects as predicted by the confluence model. The corpus of evidence invoked here offers strong support for the assumption of the confluence model that birth-order differences in intellectual performance are primarily a within-family phenomenon.

Keywords: *birth order; confluence model; intellectual performance*

A recent article by Wichman, Rodgers, and MacCallum (2006) claims to have decisively shown that the typical birth-order effects on intellectual performance found in hundreds of studies are not a function of within-family factors but are the consequence of between-family influences. They argue that

If the birth order effect is primarily a within-family . . . phenomenon, then accounting for between-family . . . variance should have little influence on the effect.

However, if the birth order effect is primarily due to factors varying between families, then the addition of a between-family variable could reduce or even eliminate the effect. (p. 124)

That between-family variable is mother’s age at birth of the first child. The authors ran three regression analyses on 3,671 children. In two analyses—one without external controls and another with a partial control of participants’ age—Wichman et al. did obtain significant birth-order differences on three measures of cognitive performance. In a third, “critical” (Wichman et al., 2006, p. 124) analysis, mother’s age at birth of the first child enters into the regression equation. That analysis renders birth-order effects not significant on all three measures, although two of these measures remain near significance.

There is nothing wrong with the logic of this argument, and Wichman et al.’s (2006) efforts to devise an innovative within-family methodology represent a useful addition to existing research. Wichman et al., however, are

Authors’ Note: We are grateful to Professor Hazel Markus, Stanford University, for her very valuable suggestions and to Robert Rosenthal, University of California–Riverside, for insightful statistical advice. We also thank Aaron Wichman for providing us with unpublished information about the sample sizes in Wichman, Rodgers, and MacCallum’s (2006) Table 5. Finally, we are grateful for constructive suggestions from several anonymous reviewers. Correspondence concerning this article may be addressed to R. B. Zajonc, Psychology Department, Stanford University, Stanford, CA 94305, e-mail: zajonc@psych.stanford.edu; or to Frank J. Sulloway, Institute of Personality and Social Psychology, University of California, Berkeley, CA 94720; e-mail: sulloway@berkeley.edu.

PSPB, Vol. 33 No. 9, September 2007 1187-1194
DOI: 10.1177/0146167207303017

© 2007 by the Society for Personality and Social Psychology, Inc.

virtually unique in having mitigated birth-order effects by adding a between-family control. Nearly all other large datasets (which are mostly uncited by these authors) retain their typical birth-order patterns even when exogenous between-family factors are introduced. Thus, the pioneering study by Belmont and Marolla (1973), with an N of 386,114, divided its population of Dutch 19-year-old individuals into nonmanual, manual, and farm subgroups. Significant birth-order effects remained, as they also did when these data were stratified by the number of children in the family. Because the number of children in the family is closely related to the age of the mother at first birth, one would expect, according to the Wichman et al. thesis, that birth-order effects would be substantially diluted when controlled for family size. The literature shows that they are not (Belmont and Marolla, 1973; Breland, 1974; Claudy, Gross, & Strause, 1974; Davis, Cahan, & Bashi, 1977; Zajonc, 1983). The same result is true of a large dataset published by the French Institute of Demographic Studies (Gille et al., 1954, pp. 55-58). Their sample of 65,250 children, 6 to 14 years old, consisted of professional and executive levels ($N = 1,325$), small businesses ($N = 7,289$), clerical ($N = 14,527$), skilled workers ($N = 17,160$), farmers ($N = 3,733$), special workers ($N = 8,255$), and farm workers ($N = 12,961$). Overall intellectual performance declined with lower socioeconomic status (SES) levels, but within each SES level, intellectual performance also systematically declined with birth order and family size. Adding a significant between-family control therefore had no effect on birth-order differences. A similar pattern of results was obtained by Claudy et al. (1974). These researchers stratified their project talent sample of American 12-year-old students ($N = 81,175$) into five socioeconomic categories. Data from each SES group showed identical patterns of declining scores by birth order.

There are also 41,482 data points from Scotland (Scottish Council for Research on Education, 1949), 3,418 from England (Douglas, 1964), and 36,000 from Columbia (Velandia, Grandon, & Page, 1978) that conform to these patterns. All of these results are extensively discussed by Zajonc (1983) but are ignored by Wichman et al. (2006). Also ignored by them is the classic within-family study by Tabah and Sutter (1954a, pp. 97-138) that obtained birth-order effects on 1,244 sibling pairs.

A particularly impressive study examined 127,902 Norwegian 18- and 19-year-olds from the same families (Bjerkedal, Kristensen, Skjeret, and Bervik, in press). In this study, IQ declined significantly with increasing birth rank. The reported disparity between firstborns and secondborns was 2.3 IQ points. These findings were nearly identical to those obtained by the same researchers in a large between-family analysis ($N = 112,799$). In another

study based on this same Norwegian sample, Kristensen and Bjerkedal (2007) found that IQ corresponded with how subjects were raised—taking into account the early death of older siblings—rather than with how subjects were born, and these results also appear to rule out a biological (gestational) explanation. Such findings are scarcely compatible with the hypothesis that birth order effects are spurious and derive from uncontrolled differences between families (Sulloway, 2007b).

Now, why did Wichman et al. (2006) fail to find birth-order effects when they introduced mother's age at the birth of her first offspring into their regression equation? Maternal age does influence family size, and it is also correlated with SES. But how exactly would these factors dilute within-family differences in birth order? Wichman et al. asserted that they "expected that maternal age values would be positively associated with child intelligence" (p. 121), given that older mothers are more likely to be career oriented, to value education, and to foster educational achievement in their children. Women in their sample who have a first child at a later age, Wichman et al. implied, are likely to raise the intelligence scores of firstborns relative to children of higher birth orders because such women tend to have smaller families than do other women and perhaps also value education more highly. This argument rests partly on the premise that birth-order patterns vary substantially with family size—a premise that the vast birth-order literature does not support.

Wichman et al.'s (2006) failure to obtain birth-order effects when maternal age is introduced into the equation can also be attributed to their problematic study design. First, the participant population sample in Wichman et al.'s analysis is seriously skewed by family size. Second, the number of mothers (and their within-family offspring) is perfectly correlated with this skew in family size. In the younger cohort, there are 1,329 mothers with one child in the study, 259 mothers with two children, 17 mothers with three children, and, *nota bene*, just 1 mother with four children. Among the older participants, there is also just 1 mother of four offspring. At the level of participants, rather than mothers, 50% of the sample are firstborns but only 14% are thirdborns and 4% are fourthborns. In addition, owing to the peculiar distribution of the sample, only 34% of the participants constitute a within-family comparison.

The more critical methodological issue, however, is why the mean scores by birth order are not statistically significant when Wichman et al. (2006) introduced maternal age at first birth as a control. A closer look at the data indicates that the search for statistical significance is obscuring a much more important conclusion from this study, namely, that the mean differences

between birth ranks in the controlled data are close to what would be expected given the nature of the sample and the specific cognitive tests included in their study. Table 5 of Wichman et al.'s study shows that differences in reading recognition involve a mean decrease of 0.85 points per increase in birth rank, from 103.17 to 100.61. Similarly, the mean decrease in reading comprehension scores by birth rank is 0.72, from 104.35 to 102.19. (To put these differences in perspective, they represent 62% of the highly significant differences observed in the magnitude of the β_{1-4} scores reported in Wichman et al.'s Table 3 before controlling for cohort effects and for mother's age at the birth of her first child.) Only the scores in mathematics exhibit a trend in the opposite direction (there is a negligible increase of 0.12 points in the mean scores by birth rank, from 99.11 to 99.46).

Typically, however, larger differences by birth order and family size are found in language-related tests than in mathematics tests (Ernst & Angst, 1983). These aggregate findings reflect the richer verbal environment that, according to the confluence model, older children are expected to experience within the family. Also important to note is that birth-order differences in most intellectual test scores involve a decline of about one IQ point per birth rank. In the study by Wichman et al. (2006), the mean decrease by birth rank in the two reading scores is 0.79 points, or a total of 2.4 points between the firstborn and the fourthborn after controlling for mother's age at the birth of her first child. In addition, it can be seen from Wichman et al.'s Table 5 that these differences are nearly statistically significant, suggesting that these findings reflect inadequate statistical power to test the expected results rather than a refutation of the confluence model. This lack of statistical power is aggravated, moreover, by the skewed distribution of the data, which underrepresents precisely those birth ranks that are expected to exert the greatest influence on the statistical results. (There are only 158 fourthborns and 507 thirdborns in Wichman et al.'s sample, compared with 1,162 secondborns and 1,844 firstborns.)

In addition, in analyses that are controlled for the critical role of the mother's age at the birth of her first offspring, Wichman et al. (2006) tested the rather uninteresting hypothesis that scores by birth order would differ from one another in some indeterminate way. This test is decidedly not a test of the confluence model, and it is also not the recommended way to test any hypothesis about predicted trends (Rosenthal, Rosnow, & Rubin, 2000). Instead, Wichman et al. should have employed focused comparisons to test the expected results, namely, that scores by birth rank would be successively different. We note, for example, that the ordering of the means in

Wichman et al.'s Table 5 matches the predicted order on each of the two reading tests ($1/4! p = .04$). In addition, from Wichman et al.'s Table 5, one can compute alerting correlations (Rosenthal et al., 2000), which provide a useful guide in such cases. For reading comprehension scores, r_{alerting} is .99. For reading recognition, r_{alerting} is .98. These two alerting correlations indicate that reading scores by birth order are consistently ordered in a linear trend, with firstborns being highest and fourthborns being lowest.

The variances in test scores reported by Wichman et al. (2006) in Table 5 are significantly heterogeneous by birth order, posing a problem for using t tests to conduct contrast analysis of the reading scores. For reading comprehension scores, $F_{\text{max}} = 2.43$ ($df = 4, 431, p < .0001$), and for reading recognition, $F_{\text{max}} = 2.74$ ($df = 4, 463, p < .0001$)—for the relevant dfs , see Tables 1 and 2. To deal with this problem and to preserve statistical independence, we have computed Satterthwaite-adjusted contrasts for each of the two reading tests. There are two different ways of performing such contrasts while also preserving statistical independence: We may contrast firstborns with thirdborns and secondborns with fourthborns. We may also contrast firstborns with fourthborns and secondborns with thirdborns. As shown for reading comprehension scores (Table 1: Entries 1 and 2), the contrast between firstborns and thirdborns yields a significant difference ($p = .035$), whereas the contrast between secondborns and fourthborns yields a nonsignificant difference ($p = .122$). Because these two independent contrasts represent the common effects of birth order in relation to intellectual performance, we may combine them using Stouffer's method (Table 1: Entry 3), which yields a significant overall difference ($p = .018$) and $r_{\text{equivalent}} = .06$ (Rosenthal & Rubin, 2003). Alternatively, we may derive an overall effect size from these two pairwise contrasts by using the method of weighted z_{rs} , which yields the same effect size obtained with $r_{\text{equivalent}}$.

The alternative set of pairwise contrasts—between firstborns and fourthborns, and secondborns and thirdborns—yields similar results (Table 1: Entries 4 through 6). Inasmuch as Satterthwaite's correction exerts only a small influence on results involving large samples, we have also calculated t_{contrast} and r_{contrast} directly from these data as the natural complement to r_{alerting} . These results (Table 1: Entry 7) are closely comparable to the others in Table 1.

As we have noted, the sample sizes in Wichman et al.'s (2006) study diminish in direct proportion to the observed birth-order differences, thereby underestimating the effects that might have been observed in a more balanced sample with the same means and variances. For reading comprehension scores, the resulting power loss is a substantial 66% (Rosenthal et al., 2003, Equation 3.25). As shown in Entry 8 of Table 1,

TABLE 1: Methods of Contrasting Reading Comprehension Scores by Birth Order

<i>Groups Contrasted</i>	<i>Method(s)</i>	<i>t or Z</i>	<i>df</i>	<i>p (one tail)</i>	<i>r</i>
1. Firstborns versus thirdborns	Satterthwaite's method	1.81 (<i>t</i>)	1,068	.035	.06
2. Secondborns versus fourthborns	Satterthwaite's method	1.17 (<i>t</i>)	183	.122	.09
3. Pairwise contrasts combined (groups 1 and 2, above)	Stouffer's method and $r_{equivalent}$; also weighted z_{rs}	2.10 (<i>Z</i>)	1,251	.018	.06 ($r_{equivalent}$), .06 ($r_{contrast}$)
4. Firstborns versus fourthborns	Satterthwaite's method	1.54 (<i>t</i>)	191	.063	.11
5. Secondborns versus thirdborns	Satterthwaite's method	1.24 (<i>t</i>)	937	.107	.04
6. Pairwise contrasts combined (groups 4 and 5, above)	Stouffer's method and $r_{equivalent}$; also weighted z_{rs}	1.97 (<i>Z</i>)	1,128	.024	.06 ($r_{equivalent}$), .05 ($r_{contrast}$)
7. All four birth orders (weighted +3, +1, -1, -3)	Contrast <i>t</i> test	2.80 (<i>t</i>)	3,124	.003	.05 ($r_{contrast}$)
8. All four birth orders (weighted +3, +1, -1, -3), equal- <i>ns</i> (Rosenthal et al., 2000: Equation 3.25)	Contrast <i>t</i> test	4.82 (<i>t</i>)	3,124	.00001	.09 ($r_{contrast}$)

NOTE: The relevant *ns* for these computations are 1,571 firstborns, 990 secondborns, 432 thirdborns, and 135 fourthborns.

TABLE 2: Methods of Contrasting Reading Recognition Scores by Birth Order

<i>Groups Contrasted</i>	<i>Method(s)</i>	<i>t or Z</i>	<i>df</i>	<i>p (one tail)</i>	<i>r</i>
1. Firstborns versus thirdborns	Satterthwaite's method	1.34 (<i>t</i>)	1,252	.092	.04
2. Secondborns versus fourthborns	Satterthwaite's method	1.34 (<i>t</i>)	208	.193	.09
3. Pairwise contrasts combined (groups 1 and 2 above)	Stouffer's method and $r_{equivalent}$; also weighted z_{rs}	1.89 (<i>Z</i>)	1,460	.029	.05 ($r_{equivalent}$), .05 ($r_{contrast}$)
4. Firstborns versus fourthborns	Satterthwaite's method	1.69 (<i>t</i>)	226	.047	.11
5. Secondborns versus thirdborns	Satterthwaite's method	0.80 (<i>t</i>)	1,041	.219	.02
6. Pairwise contrasts combined (groups 4 and 5 above)	Stouffer's method and $r_{equivalent}$; also weighted z_{rs}	1.76 (<i>Z</i>)	1,267	.039	.05 ($r_{equivalent}$), .04 ($r_{contrast}$)
7. All four birth orders (weighted +3, +1, -1, -3)	Contrast <i>t</i> test	2.82 (<i>t</i>)	3,304	.002	.05 ($r_{contrast}$)
8. All four birth orders (weighted +3, +1, -1, -3), equal- <i>ns</i> (Rosenthal et al., 2000: Equation 3.25)	Contrast <i>t</i> test	4.72 (<i>t</i>)	3,304	.00001	.08 ($r_{contrast}$)

NOTE: The relevant *ns* for these computations are 1,626 firstborns, 1,067 secondborns, 464 thirdborns, and 151 fourthborns.

had the sample sizes for each birth-order group been equal (keeping the total *N* unchanged), $t_{contrast}$ would have been 4.82, $p = .00001$, yielding $r_{contrast} = .09$ —nearly twice the effect size obtained from the unbalanced sample ($r_{contrast} = .05$). In short, Wichman et al.'s (2006) unbalanced study design appears to substantially underestimate the effect sizes and hence the statistical sig-

nificance of the birth-order trends inherent in their data.

We have performed the same series of contrast analyses for reading recognition scores (Table 2). To summarize these results, all of the values for *Z* and $t_{contrast}$ are statistically significant, and the effect sizes computed using these differing methods agree closely. In addition, the results derived from a balanced research design would have

yielded a substantially larger effect compared with the unbalanced sample ($r_{contrast} = .08$ versus $r_{contrast} = .05$). Hence, when analyzed by the method of contrasts, the scores for both reading tests yield significant differences by birth order and also exhibit effect sizes similar to those found in other large, between-family studies. Although these findings are consistent with the confluence model's predictions about intellectual performance, they are also consonant with other explanations of how birth order affects intellectual ability, including theories based on sibling differences in family niches and differences in parental investment (Hertwig, Davis, & Sulloway, 2002).

There are at least three other reasons why the Wichman et al. (2006) study fails to document statistically significant birth-order effects. First, their study appears to include an unknown number of only children, who are not expected to exhibit as high scores on intelligence tests as are firstborns with younger siblings, owing to the teaching function that a younger sibling provides for the firstborn in the family (Zajonc, 1976) as well as higher rates of single-parent households among families with only children (Ernst & Angst, 1983). Second, the sample is limited to children up to the age of 14, around the point where the confluence model predicts that the scores of firstborns have begun to outstrip those of laterborns, owing to the dynamics of how the confluence model actually relates to the family's intellectual environment. If children are tested for birth-order effects at the same age, then necessarily the younger child in a two-child family will experience a more favorable intellectual environment at that age than the older child had. Thus, for example, when two siblings born 3 years apart are both tested at 7 years of age, the environment of the younger will encompass a 10-year-old sibling, whereas the environment of the older will encompass a less mature 4-year-old sibling (Zajonc, Markus, & Markus, 1979, pp. 1333-1336). According to the confluence model, it is only as children approach adulthood that the intellectual superiority of higher birth ranks finally manifests itself in a distinct manner as older children finally achieve maximum benefit from their efforts at teaching what they have learned to younger siblings. The confluence model is also consistent with the argument that older siblings act as surrogate parents within the family system and thereby attempt to increase parental investment by occupying a responsible and adult-like family niche, which typically includes efforts to do well in scholastic pursuits (Sulloway, 1996, 2007a, 2007b). These two complementary family dynamics also explain why only children generally score at a lower level than do firstborns of two. The cumulative benefits of teaching and surrogate parenting more generally are presumably why the data analyzed by Belmont and Marolla (1973), which included only 19-

year-old individuals, produced such clear and consistent patterns of birth-order and family-size differences for nearly 400,000 individuals. They reported a significance level at $p < 10^{-15}$, controlled for sibship size and social class. If such differences were merely artifacts of uncontrolled maternal age, one would not have expected the extensive data published by Belmont and Marolla to have produced such systematic birth-order differences, after being controlled for sibship size and socioeconomic status, given that these two variables are consistently correlated with mother's age at first birth.

This prediction about the reversal of birth-order effects with age, which is made by the confluence model, brings up a third important point. A number of studies have failed to obtain the typical birth-order effects (reviewed in Zajonc, 2001; Zajonc et al., 1979). These studies are almost all conducted on children under the age of 12 and are thus congruent with the predictions of the confluence model.

Bearing in mind this information about the ages of the subjects in Wichman et al.'s (2006) study and how age relates to the confluence model, it is now relevant to note that Wichman et al. combined their two samples (of 7- and 8-year-old and 13- and 14-year-old children) in all of their statistical analyses, thereby obscuring the important fact that laterborns are expected to have slightly higher test scores than are firstborns at the younger age. Given the mean age of their combined sample (10.4 years), their results closely match the trends observed in other studies, as indicated in Figure 1 (reproduced, in revised form, from Zajonc, 2001). That is to say, in younger samples laterborns tend to outstrip firstborns, and in older samples firstborns tend to outstrip laterborns. In general, the crossover age is predicted to be 11 ± 2 years (Zajonc & Mullally, 1997). It should be noted, however, that exactly when this crossover point is reached in any given dataset depends on the age spacing between offspring as well as family size—matters that are not addressed by Wichman et al. (Zajonc, 1983; Zajonc et al., 1979).

From Figure 1, which is based on 51 samples involving more than 150,000 observations, it can now be seen that Wichman et al. (2006) could hardly have picked a more inappropriate set of subjects with which to conduct their "critical" tests of the confluence model. For a sample with an average age of 10.4 years, laterborns are actually expected to have slightly higher scores than are firstborns by about 0.07 standard deviation units, although additional data on age spacing and family size would be required to make an accurate prediction for this particular dataset. More to the point, it may be seen in Figure 1 that Wichman et al.'s results are close to the regression line for the overall trend, which is statistically significant for the 51 samples ($r = .63$, $p < .001$).

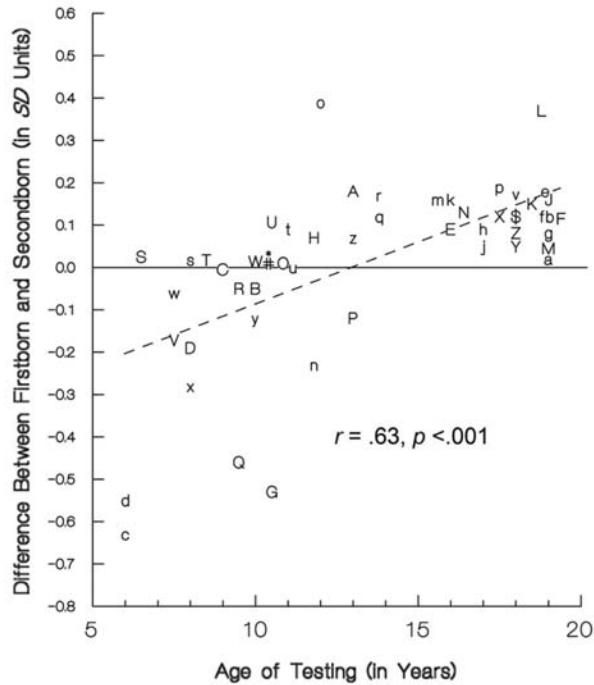


Figure 1 Differences in intellectual performance (in standard scores) between firstborn and secondborn children as a function of the age of testing (51 samples).

NOTE: In any given study, the expected outcome by birth order depends not only on age but also on age spacing between offspring and on family size. To preserve statistical independence, the correlation indicated in the figure ($r = .63$) combines the two plotted results for Wichman, Rodgers, and MacCallum (2006; # = mathematics, * = reading ability) into a single mean-weighted score. Other symbols refer to data from the following studies: a = Altus (1965, male participants); b = Altus (1965, female participants); c = Arthur (1926, Sample 2); d = Arthur (1926, Sample 1); e = Belmont and Marolla (1973, nonmanual condition); f = Belmont and Marolla (1973, manual condition); g = Belmont and Marolla (1973, farm group); h = Breland (1974, male participants); j = Breland (1974, female participants); k = Burton (1968, male participants); m = Burton (1968, female participants); n = Cicirelli (1977, male participants); o = Cicirelli (1977, female participants); p = Claudy (1976); q = Davis, Cahan, and Bashi (1977, Western Israeli participants); r = Davis et al. (1977, Asian Israeli participants); s = Douglas (1964, 8-year-olds); t = Douglas (1964, 11-year-olds); u = Eysenck and Cookson (1970); v = Galbraith (1982); w = Hsiao (1931, 7.5-year-olds); x = Hsiao (1931, 8-year-olds); y = Hsiao (1931, 10-year-olds); z = Hsiao (1931, Sample 1, 13-year-olds); A = Hsiao (1931, Sample 2, 13-year-olds); B = Institut National d'Études Démographiques (1973); C = Jensen (personal communication, December 12, 1979); D = Koch (1954); E = Kunz and Peterson (1977); F = Page and Grandon (1979); G = Richardson (1936, spacing of 3.6 years); H = Richardson (1936, spacing of less than 1 year); J = Rosenberg and Sutton-Smith (1969, male participants, large spacing); K = Rosenberg and Sutton-Smith (1969, female participants, large spacing); L = Rosenberg and Sutton-Smith (1969, male participants, short spacing); M = Rosenberg and Sutton-Smith (1969, female participants, short spacing); N = Schachter (1963); O = Scottish Council for Research on Education (1949); P = Steckel (1930); Q = Steelman and Mercy (1980, below poverty level); R = Steelman and Mercy (1980, above poverty level); S = Svanum and Bringle (1980, 6- to 7-year-olds); T = Svanum and Bringle (1980, 8- to 9-year-olds); U = Svanum and Bringle (1980, 10- to 11-year-olds); V = Tabah and Sutter (1954a, 6- to 9-year-olds); W = Tabah and Sutter (1954a, 9- to 12-year-olds); X = Velandia, Grandon, and Page (1978); Y = Zajonc and Bargh (1980, 1970-1971 cohort); Z = Zajonc and Bargh (1980, 1973-1974 cohort); \$ = Zajonc and Bargh (1980, 1976-1977 cohort).

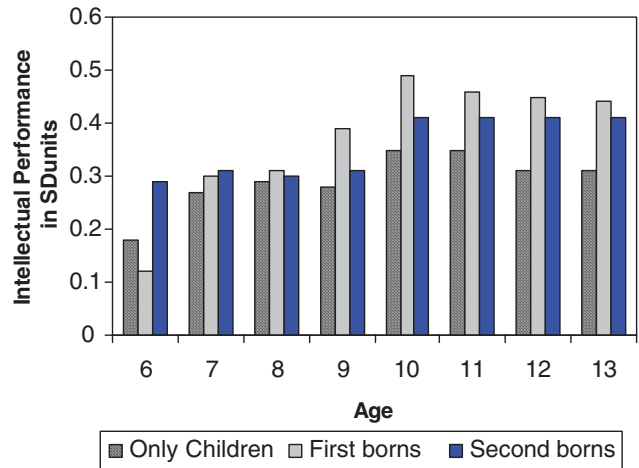


Figure 2 Birth-order data on intellectual performance, in standard deviation units, of one- and two-child families tested between ages 6 and 13 ($N = 2,488$).

SOURCE: Based on data from the Institut National d'Études Démographiques (Zajonc, Markus, & Markus, 1979).

NOTE: All scores are positive because they are relative to other subjects in the same sample, which includes families of up to six children.

Significant also are previously published one- and two-child family data (Zajonc, 2001; Zajonc et al., 1979) for a large sample ($N = 33,339$) collected by the Institut National d'Études Démographiques (1973), shown in Figure 2. This figure is quite clear in demonstrating three important features of family dynamics, predicted by the confluence model, as they change with age. First, at an early age, the secondborn is ahead of the firstborn in intellectual performance. Second, beginning at ages 8 and 9, the firstborn starts surpassing the younger sibling. Third, the only child does not benefit from the teaching function. Hence, he or she has a lower score than either of the children in two-child families except at the very youngest age, when there is less opportunity for instructing siblings. (This only-child disparity may also reflect a higher rate of single-parent households among people with one child.)

Of even greater relevance are within-family sibling data published by Tabah and Sutter (1954b). Using a sample of 1,244 two-child families, they found secondborn children 9 years and younger to score higher than firstborns on the French intellectual performance test (standard scores of $-.175$ vs. $-.003$ for firstborn and secondborn children, respectively). Siblings older than 9 years, however, manifested the opposite trend. The scores of the firstborns were higher than those of the secondborns ($.048$ vs. $.036$). Zajonc et al. (1979) found the above interaction to be significant ($F[1, 2476] = 6.65, p < .01$). This age-dependence of birth-order effects found in within-family data is parallel to the aggregate two-child family data shown in Figure 2. Although

Wichman et al. (2006) alluded to a small number of other within-family studies that show no significant birth-order effects, these other studies, like that by Wichman et al., involve modest sample sizes (especially for children of high birth rank) and also fail to adopt adequate controls for the changing patterns of birth-order results that are expected for children of differing ages.

These patterns of differences replicate the results of the large between-family datasets of Belmont and Marolla (1973), Breland (1974), and Claudy et al. (1974)—each based on many thousands of data points. In addition, the study by Bjerkedal et al. (in press) offers clear evidence that IQ declines with increasing birth rank among 127,902 Norwegian siblings raised in the same families. Such collective findings from both between-family and within-family studies, spanning the ages from 6 to early adulthood, are entirely consistent with the changes in family circumstances that involve the relatively poor early environment of a young first-born with siblings, the tendency for older siblings to benefit from a teaching function over time, and associated niche partitioning among siblings.

One additional methodological point bears emphasizing. Critics of birth-order theories about intellectual ability, and about human behavior more generally, have often dismissed such reported differences as being negligible because they typically explain only 1% or less of the variance in test or survey data (Ernst & Angst, 1983). One must keep in mind that error variance substantially reduces reported effect sizes. Even ignoring this important methodological fact, a relationship that explains just 1% of the variance in any test outcome (which is equivalent to an odds ratio of 1.38) is much more substantial than most people realize. For example, one well-known drug study whose results explained substantially less than 1% of the variance in therapeutic outcomes was terminated early because researchers considered it unethical to further withhold such lifesaving medications either from the untreated control population or from the public at large. Even more notably, the much acclaimed results of the 1954 Salk vaccine trial for polio explained only 1/100th of 1% of the variance in treatment outcomes (Rosnow & Rosenthal, 2003). In a world in which intelligence and human behavior are influenced by innumerable factors, it is noteworthy whenever one can isolate a single influence that explains just 1% of the variance in test scores. We are not, therefore, arguing that birth order is a major determinant of intellectual ability—only that it makes a modest and meaningful contribution after being controlled for other influences.

In sum, the study by Wichman et al. (2006) blends two different datasets with two different birth-order trends that are expected to cancel each other out. Given

other problems in their study—such as their skewed sample distributions, the presumed mixing of only children with firstborns having younger siblings, and age distributions that are not in fact expected to produce large birth-order effects (especially after being blended together, in violation of the confluence model)—it is noteworthy that their controlled results were nearly statistically significant for two of the three tests (and are significant when tests entailing the proper contrasts are conducted). It is even more noteworthy that these results produced almost the same effect sizes that have been reported for birth order and verbal ability in other studies conducted on participants of similar ages. The problem of Wichman et al.'s study is thus one of inadequate statistical power compounded by multiple methodological limitations all tending to reduce the reported effect sizes as well as to deprive them of statistical significance.

Despite its methodological limitations in design, sampling, and analysis, the study by Wichman et al. (2006) actually adds to rather than subtracts from the extensive and compelling published record of birth-order effects that support the confluence model. Indeed, given the age of their subjects, Wichman et al.'s findings are almost as predicted by the confluence model. The fact that the authors themselves fail to assert this conclusion needs to be understood in terms of the various methodological problems associated with their study and their failure to specify clearly what the confluence model actually predicts about birth order and intellectual ability.

REFERENCES

- Altus, W. D. (1965). Birth order and scholastic aptitude. *Journal of Consulting Psychology, 29*, 202-205.
- Arthur, W. D. (1926). The relation of IQ to position in family. *Journal of Educational Psychology, 17*, 541-550.
- Belmont, L., & Marolla, F. A. (1973). Birth order, family size, and intelligence. *Science, 182*, 1096-1101.
- Bjerkedal, T., Kristensen, P., Skjeret, G. A., & Brevik, J. I. (in press). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence*, DOI:10.1016/j.intell.2007.01.004.
- Breland, H. M. (1974). Birth order, family configuration, and verbal achievement. *Child Development, 45*, 1011-1019.
- Burton, D. (1968). Birth order and intelligence. *Journal of Social Psychology, 76*, 199-206.
- Cicirelli, V. G. (1977). Children's school grades and sibling structure. *Psychological Reports, 41*, 1055-1058.
- Claudy, J. G., Gross, D., & Strause, R. (1974). *Family size, birth order, and characteristics of young adults* [Final Report AIR-41900-10/74-FR]. Palo Alto, CA: American Institute for Research.
- Davis, D., Cahan, S., & Bashi, J. (1977). Birth order and intellectual development: The confluence model in the light of cross-cultural evidence. *Science, 196*, 1470-1472.
- Douglas, J. W. B. (1964). *The home and the school: A study of ability and attainment in the primary school*. London: Macgibbon & Kee.
- Ernst, C., & Angst, J. (1983). *Birth order: Its influence on personality*. Berlin: Springer-Verlag.

- Eysenck, H. J., & Cookson, D. (1970). Personality in primary school children: 3. Family background. *British Journal of Educational Psychology*, 40, 117-131.
- Galbraith, R. C. (1982). Sibling spacing and intellectual development: A closer look at the confluence models. *Developmental Psychology*, 18, 181-191.
- Gille, R., Henry, L., Tabah, L., Sutter, J., Bergues, H., Girard, A., & Bastide, H. (1954). *Le niveau intellectuel des enfant d'age scolaire: Cahier No. 23. La détermination des aptitudes, l'influence des facteurs constitutionnels, familiaux et sociaux*. Institut National d'Études Démographiques. Paris: Presses Universitaire de France.
- Hertwig, R., Davis, J. N., & Sulloway, F. J. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin*, 128, 728-745.
- Hsiao, H. H. (1931). The status of the first-born with special reference to intelligence. *Genetic Psychology Monographs*, 9, 43-115.
- Institut National d'Études Démographiques. (1973). *Enquête nationale sur le niveau intellectuel des enfants d'âge scolaire* [National survey of intellectual level of school-age children]. Paris: Presses Universitaire de France.
- Koch, H. L. (1954). The relation of "primary mental abilities" in five- and six-year-olds to sex of child and characteristics of his sibling. *Child Development*, 25, 209-223.
- Kristensen, P., & Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, 317, 1717.
- Kunz, P. R., & Peterson, E. T. (1977). Family size, birth order, and academic achievement. *Social Biology*, 24, 144-148.
- Page, E. B., & Grandon, G. M. (1979). Family configuration and mental ability: Two theories contrasted with U.S. data. *American Educational Research Journal*, 16, 257-272.
- Richardson, S. K. (1936). The correlation of intelligence quotients of siblings at the same chronological age levels. *Journal of Juvenile Research*, 20, 186-197.
- Rosenberg, B. G., & Sutton-Smith, B. (1969). Sibling age spacing effects upon cognition. *Journal of Developmental Psychology*, 1, 661-668.
- Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge, UK: Cambridge University Press.
- Rosenthal, R., & Rubin, D. 2003. $r_{equivalent}$: A simple effect size indicator. *Psychological Methods*, 8, 492-496.
- Rosnow, R. L., & Rosenthal, R. 2003. Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, 57, 221-237.
- Schachter, S. (1963). Birth order, eminence and higher education. *American Sociological Review*, 28, 757-768.
- Scottish Council for Research on Education. (1949). *The trend of Scottish intelligence*. London: University of London Press.
- Steckel, M. L. (1930). Intelligence and birth order. *Journal of Social Psychology*, 1, 329-344.
- Steelman, L. C., & Mercy, J. A. (1980). Unconfounding the confluence model: A test of sibship size and birth order effects on intelligence. *American Sociological Review*, 45, 571-582.
- Sulloway, F. J. (1996). *Born to rebel: Birth order, family dynamics, and creative lives*. New York: Pantheon.
- Sulloway, F. J. (2007a). Birth order and sibling competition. In R. Dunbar & L. Barrett (Eds.), *Handbook of evolutionary psychology* (pp. 297-311). Oxford, UK: Oxford University Press.
- Sulloway, F. J. (2007b). Birth order and intelligence. *Science*, 317, 1711-1712.
- Svanum, S., & Bringle, R. G. (1980). Evaluation of confluence model variables on IQ and achievement test scores in a sample of 6- to 11-year-old children. *Journal of Educational Psychology*, 72, 427-436.
- Tabah, L., & Sutter, L. (1954a). Le niveau intellectuel des enfants d'une même famille [The intellectual level of children of the same family]. In R. Gille, L. Henry, L. Tabah, J. Sutter, H. Bergues, A. Girard, & H. Bastide, *Le niveau intellectuel des enfants d'âge scolaire: Cahier No. 23. La détermination des aptitudes, l'influence des facteurs constitutionnels, familiaux et sociaux*. Institut National d'Études Démographiques (pp. 97-138). Paris: Presses Universitaire de France.
- Tabah, L., & Sutter, J. (1954b). Le niveau intellectuel des enfants d'une meme famille. *Annals of Human Genetics*, 19, 120-150.
- Velandia, W., Grandon, G. M., & Page, E. B. (1978). Family size, birth order, and intelligence in a large South American sample. *American Educational Research Journal*, 15, 399-416.
- Wichman, A. L., Rodgers, J. L., & MacCallum, R. C. (2006). A multilevel approach to the relationship between birth order and intelligence. *Personality and Social Psychology Bulletin*, 32, 117-127.
- Zajonc, R. B. (1976). Family configuration and intelligence. *Science*, 192, 227-236.
- Zajonc, R. B. (1983). Validating the confluence model. *Psychological Bulletin*, 93, 457-480.
- Zajonc, R. B. (2001). The family dynamics of intellectual development. *American Psychologist*, 56, 490-496.
- Zajonc, R. B., & Bargh, J. (1980). Birth order, family size, and decline of SAT scores. *American Psychologist*, 35, 662-668.
- Zajonc, R. B., Markus, H., & Markus, G. B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology*, 37, 1325-1341.
- Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist*, 52, 685-699.

Received March 7, 2006

Revision accepted January 18, 2007