BIRTH ORDER AND INTELLIGENCE: FURTHER TESTS OF THE CONFLUENCE MODEL*

ROBERT D. RETHERFORD
East-West Center
and
University of Hawaii at Manoa

WILLIAM H. SEWELL
University of Wisconsin

Confluence theory was originally developed to explain the negative effect of birth order on intelligence, as well as some peculiar effects of birth order on the intelligence of last-born children, in a large set of Dutch data. Subsequently the theory was elaborated to explain positive effects and nonlinear relationships between birth order and intelligence in other data sets. We test the mathematical form of the theory using aggregate data, between-family data, and within-family data from the Wisconsin Longitudinal Study and find no support for the theory. Fundamental flaws are identified in the statistical methods used to fit the model to the Dutch data. When the analysis is done correctly, the fit is very poor. An alternative analysis based on a within-family study design that uses sibling pairs indexed by birth order in the Wisconsin Longitudinal Study finds no evidence of birth order effects on intelligence.

At least since the time of Galton (1874), scholars have studied the effects of birth order on cognitive achievement, including eminence, educational attainment, score on educational achievement tests, and measured intelligence. Hundreds of studies have examined this question. The earlier studies were reviewed most recently by Adams (1972) and Schooler (1972); Steelman (1985) reviewed the subsequent literature on the confluence model. These reviewers have concluded that previous studies have been so seriously flawed conceptually or methodologically that no reliable conclusions can be drawn about the influence of birth order on cognitive achievement. The major faults are nonprobability sampling, samples that are not representative of any known population, selection bias, inadequate measurement of key variables, and failure to control for socioeconomic background and family structure.

The introduction of confluence theory in recent years has rekindled interest in the effect of birth order on cognitive development. Confluence theory and the confluence model, which is the theory's mathematical form, have been developed by Zajonc and his colleagues (Zajonc 1975, 1976, 1983, 1986a, 1986b; Zajonc and Markus 1975; Markus and Zajonc 1977; Zajonc, Markus, and Markus 1979; Zajonc and Bargh 1980a, 1980b; Berbaum and Moreland 1980, 1985). The theory was inspired by a Dutch data set previously analyzed by Belmont and Marolla (1973), involving some 386,000 Dutch males who reached the age of 19 between 1963 and 1966. Information was collected by the Dutch military on the measured intelligence, family size (number of siblings), and birth order of these men. Belmont and Marolla calculated an average intelligence score for each of the 45 cells in a table cross-classifying family size by birth order. These scores, transformed to mental ages for later convenience, are graphed in Figure 1.

* Joseph L. Rodgers provided a SAS program (Rodgers 1981, Appendix C) for nonlinear least-squares fitting of the confluence model, which Victoria Ho adapted for use on a microcomputer. John Bargh, Michael L. Berbaum, and Robert B. Zajonc graciously responded to requests to provide information about unpublished details on how they fitted the confluence model to data. We thank these colleagues for their help and cooperation, without which our research would have been more difficult and less conclusive. Minja Choe, Richard C. Galbraith, Arthur S. Goldberger, Robert M. Hauser, Norman Y. Luther, and Joseph L. Rodgers provided helpful comments and suggestions. The editorial staff of the *ASR* made suggestions that improved clarity and readability. We also thank Robin Loomis for research assistance. Support for the research reported in this paper was provided by a grant from the National Institutes of Health (SSP 1 R01 HD20049). Support for the Wisconsin Longitudinal Study of Social and Psychological Factors in Educational and Occupational Aspirations and Achievements, which provided the new data used in this paper, came from grants from the National Institutes of Health (MH-06275), the National Science Foundation (SES 83-20140), the Spencer Foundation, and the University of Wisconsin Research Committee.

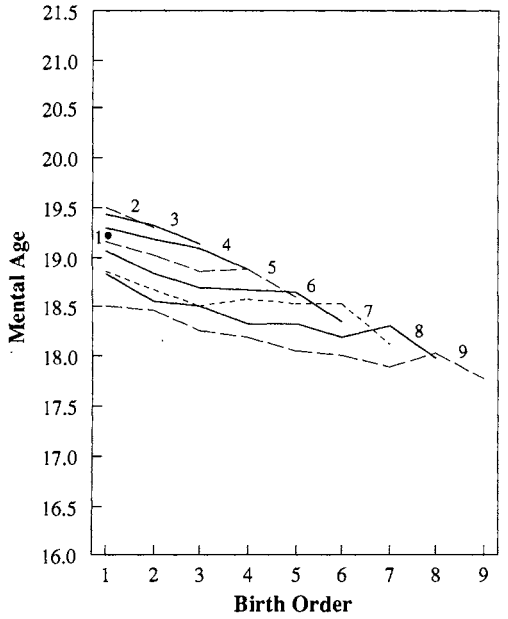


Figure 1: Observed Mental Age by Family Size and Birth Order: Dutch Males

Note: Based on values presented in the upper panel of Table 1. Each piecewise-linear plot (or point, in the case of a one-child family) represents a particular family size.

Mean mental age tends to decline as family size increases and as birth order increases within each family size category. Singletons, however, have lower mean mental age than first-borns in two-, three-, and four-child families. Within each family-size category, mean mental age declines especially rapidly for last-born children.

Zajonc and Markus (1975) formulated their ingenious confluence theory to explain the relationships between birth order, sibship size, and intellectual development. They theorized that the observed pattern in the Dutch data was the result of the confluence of two factors: family intellectual environment and a teaching function. Family intellectual environment, conceptualized as the average mental age of parents and children within the family, tends to decline with each successive birth, due to the low mental ages of young children. Thus, the relationships between intelligence and sibship size and between intelligence and birth order are both negative; the strength of these negative relationships depends partly on how closely births are spaced. A teaching-function effect arises because last-borns (including singletons) lack the opportunity to teach younger siblings. Teaching a younger sibling stimulates the intellectual development of the older child. Since last-born children have no one to teach, they suffer from a "last-born handicap."

Zajonc and Bargh (1980a) applied the confluence model (the mathematical form of the theory) to mean intelligence scores by family size and birth order from several large and divergent data sets and reported that the model fit very well. They then asserted that confluence theory explains the effects of birth order on cognitive development in these data sets.

Studies assessing the theory in relation to cognitive achievement have generally failed to confirm distinctive propositions based on it (Blake 1981, 1989; Ernst and Angst 1983; Galbraith 1982a, 1983; Grotevant, Scarr, and Weinberg 1977; Hauser and Sewell 1985; Lindert 1977; Melican and Feldt 1980; Mercy and Steelman 1982; Olneck and Bills 1979; Page and Grandon 1979; Steelman and Doby 1983; Steelman and Mercy 1980, 1981; Velandia, Grandon, and Page 1978). This has led to a lively controversy between Zajonc and his associates and those who criticize the theory on the basis of their negative findings (Berbaum 1985; Berbaum, Markus, and Zajonc 1982; Berbaum, Moreland, and Zajonc 1986; Blake 1981, 1989; Galbraith 1982a, 1982b, 1982c, 1983; McCall 1985; Rodgers 1984, 1988; Steelman 1985, 1986; Zajonc 1986a).

Although there has been dispute about theoretical and substantive matters, the debate has focused on methodological issues. Zajonc and his associates flatly reject studies using regression techniques, insisting that regression techniques fail to take into account the complexities of the dynamic theory on which the model is based, particularly the changing intellectual environment of the family as children are born and develop.

Critics respond that regression techniques produce comparable results in those instances where both methods have been used. Moreover, regression techniques permit the control of confounding variables that may account for confluence model findings. They further assert that since the theory is designed to explain the cognitive development of individuals, tests of the theory must be based on data for individuals, not on group means, since the use of group means exaggerates relationships inherent in the data (Robinson, 1950). Both sides agree only that a conclusive test of the model must use longitudinal data from large representative samples of families.

Three studies applied the confluence model to within-family data. Berbaum and Moreland (1980) used the nonlinear least-squares method to fit the confluence model to Outhit's (1933) sample of 51 upper-middle-class Canadian and

American families. The model accounted for 51 percent of the variation in the observed mental ages of the children in these families. (This figure was obtained by squaring the correlation between observed and predicted mental ages.) However, others were quick to point out that this estimate of explained variation is spuriously high because of autocorrelation between mental age and chronological age when children are tested at different ages (Galbraith 1982b; Price, Walsh, and Vilberg 1984; Rodgers 1984).

Berbaum and Moreland (1985) examined Grotevant, Scarr, and Weinberg's (1977) unusual sample of 101 upper-middle class transracial families and their natural and adopted children. They found that the confluence model accounted for up to 50 percent of the variation in observed mental ages. In response to previous criticism, however, they also calculated the partial correlation between observed and predicted mental ages while controlling for chronological age at testing, and found that the percentage of variation explained, measured by the square of the partial correlation coefficient, was 3 percent. This is slightly higher than the results originally reported by Grotevant, Scarr, and Weinberg (1977) using regression analysis on the same sample. Thus, after partialling out the effects of autocorrelation, the fit of the confluence model, as measured by the square of the correlation between observed and predicted mental ages, was very poor. However, Berbaum (1985) interpreted it as a confirmation of the theory.

A third study (Rodgers 1984) also used nonlinear least squares to fit the confluence model to individual-level data from another nonrepresentative sample, consisting of 311 families and their children from the files of the Fels Longitudinal Study. Because each child was tested at several different ages, Rodgers applied the confluence model separately for each tested age to avoid problems of autocorrelation. For a subsample of 78 children for whom an IQ score was available for at least one parent, Rodgers found that the confluence model accounted for a median value of 9 percent of the variation in observed mental age. When parental education was used as a proxy for parental IQ for those parents for whom an IQ score was not available, the confluence model accounted for a median value of 16 percent of the variation in observed mental age. Interestingly, Rodgers' simple linear regression models outperformed the confluence model in terms of percentage of variation explained. On the basis of his results, Rodgers argued that the confluence model fit the data better than no model at all, but that a simpler regression model fit the data better.

We believe that the evidence from these studies does little to resolve the controversy over the validity of confluence theory. Zajonc and his associates find the evidence confirmatory and continue to believe that the model is useful for explaining cognitive development (Zajonc 1986a; Berbaum, Moreland, and Zajonc 1986). We conclude that the model needs further testing.

We undertake various tests of the confluence model using data from a large, representative sample — the Wisconsin Longitudinal Study (WLS). We apply the mathematical form of the model to aggregate data, to between-family data on individuals, and to within-family data on individuals. Our results lead us to reanalyze the Dutch data (Belmont and Marolla 1973) that the confluence model was orginally designed to explain. Finally, we reanalyze the Wisconsin within-family data using a much simpler approach involving IQ differences within sibling pairs as a better test of the influence of birth order on measured intelligence.

THE CONFLUENCE MODEL

The confluence model is a simulation model. In the equations that constitute the model, children are portrayed as moving forward in age, one year at a time, from birth to the year of testing. Each child's mental age is recalculated at the start of each year. Although the equations have been published, the method of fitting the model to data has never been adequately described in the literature. Therefore, we present the details here.

The first step in fitting the confluence model is to transform the observed test scores into mental ages, denoted by M. M is related to IQ and chronological age, C, by the formula

$$\mathrm{IQ} = \frac{M}{C} \times 100 \tag{1}$$

which can be solved for M:

$$M = \frac{\mathrm{IQ}}{100} \times C\,. \tag{2}$$

The basic form of the confluence model then is

$$M_{ij}(t) = M_{ij}(t\text{-}1) + \alpha_{ij}(t) + \lambda_{ij}(t) \tag{3}$$

where $M_{ij}(t)$ is mental age at time t of the ith child in a family of j children (t = age of ith child and j

= family size at time t), $\alpha_{ij}(t)$ is the contribution of family intellectual environment to change in mental age between t-1 and t, and $\lambda_{ij}(t)$ is the contribution of the teaching function to change in mental age between t-1 and t.

As an intermediate step toward specifying functional forms for α and λ, intellectual growth between t-1 and t is considered proportional to the change between t-1 and t in a baseline sigmoid function

$$f(t) = 1 - e^{-k^2t^2} \tag{4}$$

where k is one of the parameters to be fitted. The coefficient of t^2 is specified as $-k^2$ instead of $-k$ to ensure that this coefficient will always be negative. This function, which resembles an elongated S, ranges from zero at $t = 0$ to one as $t \rightarrow \infty$. The size of the k parameter governs how fast $f(t)$ approaches one as t increases.[1]

A one-year change in the baseline sigmoid function is denoted $\Delta f(t)$:

$$\Delta f(t) = f(t) - f(t\text{-}1) = e^{-k^2(t-1)^2} - e^{-k^2t^2}. \tag{5}$$

Model 1: Functional forms for α and λ have been specified in the literature in two ways. The simpler specification, which we call Model 1, comes from Berbaum and Moreland (1980). The specification for α is

$$\alpha_{ij}(t) = \omega_1 \Delta f(t) \frac{\sum_{k=1}^{n} M_{kn}(t\text{-}1)}{n} \tag{6}$$

where n denotes family size, including both children and parents; n varies over the simulation and is therefore a function of t. To reduce notational clutter, however, we indicate it simply as n instead of $n(t)$. The index k (not to be confused with the parameter k) varies first over the children and then over the parents. Thus, M refers to either children or parents in equation 6, whereas in equation 3 it referred only to children. We see from equation 6 that $\alpha_{ij}(t)$ is proportional both to a one-year increase in the baseline sigmoid function, $\Delta f(t)$, and to mean mental age. The quantity ω_1 is a constant of proportionality and is one of the parameters to be fitted.

[1] It makes no sense to add an intercept term on the right side of equation 3, as some reviewers have suggested, because a nonzero intercept would make it impossible for mental age to level off in early adulthood.

Also, as part of Model 1,

$$\lambda_{ij}(t) = \omega_2 \, \Delta f(t) \, \Delta f(\tau) \, L(t) \tag{7}$$

where $L(t)$ is a dummy variable equal to 1 for children with a younger sibling and 0 otherwise, and τ is the age of the index child's adjacent younger sibling if there is one. (For example, if the adjacent younger sibling is three years younger than the index child, then $\tau = t$-3.) Thus, $\lambda_{ij}(t)$ is proportional to a one-year change in the baseline sigmoid function for the adjacent younger sibling as well as to a one-year change in the baseline sigmoid function for the index child; it is also proportional to $L(t)$. The constant of proportionality, ω_2, is one of the parameters to be fitted.

Model 2: Model 2, a more complicated specification, is the original specification proposed by Zajonc, Markus, and Markus (1979), and is the specification used by Zajonc and Bargh (1980a). It uses a root mean square instead of a simple average for the family intellectual environment term in the expression for a, and it divides by $(n\text{-}1)^2$ in the expression for λ:

$$\alpha_{ij}(t) = \omega_1 \Delta f(t) \sqrt{\frac{\sum_{k=1}^{n} M_{kn}^2(t\text{-}1)}{n+1}} \tag{8}$$

$$\lambda_{ij}(t) = \frac{\omega_2 \Delta f(t) \Delta f(\tau) L(t)}{(n\text{-}1)^2} \tag{9}$$

where, for reasons not clear to us, n pertains to time t-1 in equation 8 but to time t in equation 9. Zajonc, Markus, and Markus (1979) explain that n+1 is used instead of n in the denominator of the mean family environment term in equation 8 to give added weight to the mental ages of the more mature members. The root mean square specification also tends to give more weight to the more mature members. According to a personal communication from Zajonc, the rationale for the divisor $(n\text{-}1)^2$ in equation 9 is, first, that the index child can teach others in the family but not him- or herself and, second, that the contribution of the teaching function declines rapidly as family size increases, at least in the Dutch data.

In both models, three parameters are estimated: k, ω_1, and ω_2. The presence of the factor $\Delta f(t)$ in both $\alpha_{ij}(t)$ and $\lambda_{ij}(t)$ guarantees that $M_{ij}(t)$ levels off as t increases, since $\Delta f(t)$ approaches zero as t increases.

Methods for Fitting the Confluence Model

In fitting the confluence model to data, the general idea is to pick best-guess starting values of k, ω_1, and ω_2; enter these values into the model's equations to simulate the trajectory of mental age, one year at a time, up to the age of testing; and then compare predicted mental age with observed mental age for each individual in the sample. The analyst (or computer algorithm) then tries other values of k, ω_1, and ω_2 to see if the agreement between predicted and observed mental ages can be improved. The parameters that give the best possible agreement are ultimately chosen.

Two different methods of fitting the model to data have been used:

Method 1. The first method, used by Berbaum and Moreland (1980), is nonlinear least squares regression. The basic idea is to choose values of k, ω_1, and ω_2 that minimize the sum of squared residuals between observed and predicted mental ages over all individuals in the sample, $\Sigma(M-\hat{M})^2$. Nonlinear least squares is widely used and is appropriate in this context; packaged computer programs for this method are available, for example, from BMDP and SAS (Dixon, Brown, Engelman, Hill, and Jennrich 1988; SAS Institute, Inc. 1987).[2]

Method 2. The second method, used by Zajonc and Bargh (1980a), chooses values of k, ω_1, and ω_2 that maximize the ordinary correlation between observed and predicted mental age. We refer to this method as the "method of maximizing the correlation." Surprisingly, it has never been pointed out in the confluence model literature that Method 2 can give results very different from those obtained by Method 1 when the model is nonlinear in the parameters to be estimated (in this case k, ω_1 and ω_2).

In the case of linear models of the form

$$Y = b_0 + b_1X_1 + \ldots + b_kX_k + e \tag{10}$$

(which can also be written $\hat{Y} = b_0 + b_1X_1 + \ldots + b_kX_k$, where $\hat{Y}$ is the value of Y predicted by the model), the method of least squares and the method of maximizing the correlation are equivalent. An appropriate measure of goodness of fit for this model is the multiple correlation coefficient, R. A general property of linear models of this form is that $R = r_{\hat{Y}Y}$, where $r_{\hat{Y}Y}$ is the simple bivariate correlation between Y and $\hat{Y}$. The correlation $r_{\hat{Y}Y}$ can be written

$$r_{\hat{Y}Y} = \sqrt{1 - \frac{\sum(Y-\hat{Y})^2}{\sum(Y-\bar{Y})^2}} \tag{11}$$

where $\Sigma(Y-\hat{Y})^2$ is the sum of squared residuals. It is immediately evident from equation 11 that maximizing $r_{\hat{Y}Y}$ is equivalent to minimizing the sum of squared residuals $\Sigma(Y-\hat{Y})^2$.

If $\hat{Y}$ is obtained from a nonlinear model, however, equation 11 is generally invalid. Maximizing the ordinary correlation $r_{\hat{Y}Y}$ is no longer equivalent to minimizing $\Sigma(Y-\hat{Y})^2$, since the two methods generally yield different estimates of the model parameters (k, ω_1, and ω_2 in the case of the confluence model).

To see more clearly where Zajonc and Bargh's method of maximizing $r_{\hat{Y}Y}$ goes wrong, consider the following formula for r_{YX}:

$$r_{YX} = \sqrt{1 - \frac{\sum[Y - (u + vX)]^2}{\sum(Y-\bar{Y})^2}} \tag{12}$$

where r_{YX} measures goodness of fit to a linear model of the form $Y = u+vX+e$, and u and v are estimated by ordinary least squares. Replacement of X in equation 12 by $\hat{Y}$ yields

$$r_{Y\hat{Y}} = \sqrt{1 - \frac{\sum[Y - (u + v\hat{Y})]^2}{\sum(Y-\bar{Y})^2}}. \tag{13}$$

For underlying linear models of the form $\hat{Y} = b_0 + b_1X_1 + \ldots + b_kX_k$, it is always true that $u = 0$ and $v = 1$, so that equation 13 reduces to equation 11. But for nonlinear models such as the confluence model, u and v can take on other values.

[2] We used the DUD option of the PC version of SAS PROC NLIN, which was used also by Joseph Rodgers, whose program we borrowed. We also tried NLIN's other three options, which use partial derivatives with respect to the parameters, but we found that these options did not always converge. In the confluence model, the partial derivatives, which are derived numerically, apparently do not behave properly, but we were not able to pinpoint the problem. On the other hand, the DUD option, which is a grid-based method that does not use derivatives, always converged. We tested the DUD results by starting with alternative sets of initial values of k, ω_1, and ω_2 and always obtained the same solutions. Michael Berbaum, who requested copies of our programs and data, replicated our results, as did Robert Hauser and Linda Jordan at the University of Wisconsin. We also made some runs on a mainframe computer and obtained the same solutions as we did using an IBM PS/2 Model 80 microcomputer.

In the case of the confluence model, equation 13 becomes

$$r_{M\hat{M}} = \sqrt{1 - \frac{\sum[M - (u + v\hat{M})]^2}{\sum(M-\bar{M})^2}}. \qquad (14)$$

The right side of this equation depends not only on u and v, but also on k, ω_1, and ω_2, that are implicit in $\hat{M}$. What does it mean to maximize $r_{M\hat{M}}$ (or equivalently $r_{\hat{M}M}$) when u and v are free parameters? It means that a meaningless linear model, $M = u+v\hat{M}$, is superimposed on the original nonlinear confluence model, resulting in a hybrid model with five parameters, k, ω_1, ω_2, u, and v. Zajonc and Bargh (1980a) provide only a brief verbal description of their fitting method, and they make no mention of the two extra parameters, u and v, that are implicit in their fitting procedure. It appears that the two extra parameters were added unintentionally and have been overlooked not only by Zajonc and Bargh, but also by their critics.

It is evident from equation 14 that Zajonc and Bargh's fitting procedure minimizes the wrong set of residuals: It minimizes $\sum[M- (u+v\hat{M})]^2$ instead of $\sum(M-\hat{M})^2$. In other words, by means of a least-squares criterion, Zajonc and Bargh choose values of k, ω_1, ω_2, and (implicitly) u and v to produce the best possible agreement between M and $u+v\hat{M}$, not between M and $\hat{M}$.

Given this mistake, it is not surprising that the confluence model, when fitted by Zajonc and Bargh's procedure, yields incongruous results. For example, in Zajonc and Bargh's application of the confluence model to Belmont and Marolla's Dutch data, the procedure of maximizing $r_{\hat{M}M}$ yields predicted values of M such that $M = 16.23+.56\hat{M}$. Thus $u = 16.23$ and $v = .56$. The average difference between M and $\hat{M}$ turns out to be 14.2. Because M generally exceeds $\hat{M}$ by about 14 years, $\sum(M-\hat{M})^2$ is large, so that the model fits poorly by the least squares criterion that should have been used. Yet $r_{\hat{M}M}$ is .95. In effect, the procedure of maximizing $r_{\hat{M}M}$ produces excellent agreement between M and $16.23+.56\hat{M}$, and that is what the correlation of .95 indicates. The trouble is that values of u and v, 16.23 and .56 instead of 0 and 1, indicate a poor fit, because what is really wanted is agreement between M and $\hat{M}$, not between M and $16.23+.56\hat{M}$.

In sum, the method of maximizing $r_{\hat{M}M}$ is a flawed method of fitting the confluence model and can yield absurd results. The results obtained by this method are invalid and must be discarded. The other side of the coin is that $r_{\hat{M}M}$ is a flawed measure of goodness of fit of the confluence model. This means that the studies by Berbaum and Moreland (1980, 1985) and Rodgers (1984), which correctly used nonlinear least squares to fit the confluence model, incorrectly used $r_{\hat{M}M}$ as a measure of goodness of fit. In this paper, we use $\sum(M-\hat{M})^2$ as a measure of goodness of fit in conjunction with nonlinear least squares as the fitting method.[3]

Step-by-step procedure for fitting the confluence model to individual-level data. The confluence model, in the form of either Model 1 or Model 2, can be fitted to individual-level data or aggregate-level data. Because the way this is done is not spelled out adequately in the literature, we present it briefly here. For purposes of illustration, we consider first the fit of either Model 1 or Model 2 to individual-level data using Method 1, nonlinear least squares.

Suppose that we have a sample of 17-year-olds, each of whose intelligence was tested at age 16. For each of these primary respondents, we know the number of siblings and the date of birth of each sibling. Suppose also, as in the typical application, that we do not know the intelligence of the siblings or the parents. We proceed as follows:

(1) Pick arbitrary (i.e., best-guess) starting values of k, ω_1, and ω_2.

(2) In the absence of other information, assume that there is no child mortality; that children leave the home at age 19; that parents have a mental age of 18 (the choice of 19 and 18 for these two ages is clearly somewhat arbitrary); and that parental marriages remain intact throughout the simulation.

(3) Consider the first respondent. Using the model's equations, start the simulation for this respondent's family (respondent, sibs, and their parents) at the time of the birth of the first-born sibling.

(4) Move the simulation forward one year at a time, using the model's equations.

(5) Bring in the birth of each subsequent sibling at the appropriate time. Use the model's equations to move the new sibling forward one

[3] Nonlinear least squares is not the only alternative method that could be used. For example, the method of maximum likelihood could be employed if the problem of formulating a likelihood function for the confluence model could be solved. Although there is no single "correct" estimation procedure that should be used in this context, minimizing the sum of squared residuals is a widely used method that is appropriate.

year at a time, also. (This is necessary because we need to know the predicted mental age of each sibling who is still in the household at the start of each year in order to calculate average family intellectual environment.)

(6) Continue the simulation until the index child (i.e., the primary respondent for whom we have an intelligence score) reaches the age at testing. At this point, we have both a predicted mental age, $\hat{M}$, and an observed mental age, M, for the index child. For the other siblings, we have only a predicted mental age. Retain predicted and observed mental ages only for the index child. (In the case of within-families data, we can retain predicted and observed mental ages for all siblings for whom we have both.)

(7) Compute the sum of squared residuals, $\Sigma(M\text{-}\hat{M})^2$, where the summation ranges over all respondents in the sample. (Note again that siblings are omitted from this summation because we have no observed mental ages for them, only predicted mental ages.)

(8) Now use some kind of iterative algorithm to find a set of values of k, ω_1, and ω_2 that reduces the sum of squared residuals.

(9) Repeat steps 1 through 8 until the sum of squared residuals cannot be made any smaller, using some convergence criterion. The fitted values of k, ω_1, and ω_2 are the values in the last iteration.

The same nine steps are used with Method 2, except that the fitting criterion consists of maximizing $r_{\hat{M}M}$ instead of minimizing $\Sigma(M\text{-}\hat{M})^2$. A fit can be obtained by Method 2, even though Method 2 is flawed and the fit invalid.

Note that many of the assumptions that typically must be invoked in the second step may not be realistic. Moreover, because the simulation proceeds one year at a time, birth intervals are essentially rounded to whole numbers of years, which again is not very realistic. This raises the question of why the simulations are not done month- by-month instead of year-by-year. Zajonc and his colleagues do not speak to this question. The reason may be computer time, because even the year-by-year simulations can take considerable time to converge to a solution.

Fitting the confluence model to aggregate-level data. The procedure for fitting to aggregate-level data, which is not explained clearly in any published paper on the confluence model, is to treat each family-size/birth-order category as if it were an individual respondent. For example, suppose we consider the category of respondents for whom family size is 5 and birth order is 2. Our aggregate-level data provide an average mental age for this category, which we treat as the mental age of the second of five children. We can either calculate or assume some average childspacing for this category.[4] We then proceed as before to fit the model, executing the same nine steps. The number of "respondents" or "index children" equals the number of cell entries for mean mental age in the aggregate-level table cross-classified by family size and birth order.

DATA

The Wisconsin Longitudinal Study (WLS), begun in 1957, is based on a random sample of more than 10,000 Wisconsin high school graduates. These graduates were followed up in 1964 and 1975, with response rates approaching 90 percent. The data include a complete roster of all the living siblings of the original sample members, their sexes, dates of birth, and educational achievements. Information is also available on their parents' educational, occupational, and economic status at the time the respondents were seniors in high school. For more information on these surveys, see Sewell and Hauser (1980). Data on the measured intelligence of the original sample members were obtained from the Wisconsin Testing Service, which for many years administered the Henmon-Nelson Test of Mental Ability (1954 revision) to all Wisconsin high school eleventh graders (for more information on this test, see Retherford and Sewell 1988). Also available are test scores and family structure data for a randomly selected sibling of over one-fifth of the respondents. (For further details about the sibling sample, see Hauser, Sewell, and Clarridge (1982).)

Mental ability in the Dutch data is available only in the form of Raven test scores, reported in six categories, from 1 (high) to 6 (low). The original Raven scores were based on many items. The test scores were apparently grouped into six categories by the Dutch military, but the procedure used is not available. The mean score is 2.82 with a standard deviation of 1.43 (calculated from data in Belmont and Marolla 1973; also reported in Zajonc and Bargh 1980a).

[4] We conducted a sensitivity analysis that shows that results from the confluence model are strongly affected by small changes in the childspacing assumptions. A summary of this sensitivity analysis can be obtained from the authors upon request.

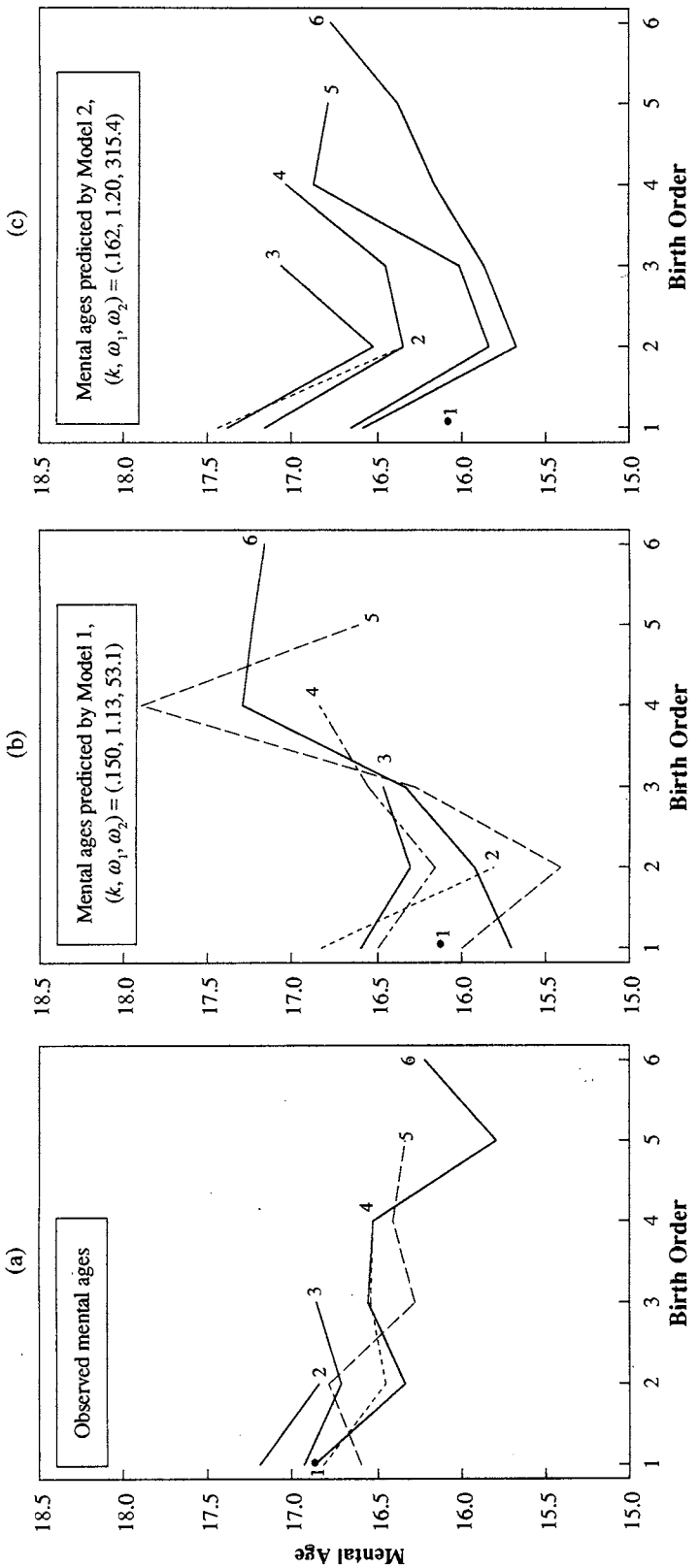


Figure 2. Fits of the Confluence Model to Aggregate-Level Data: 1975 WLS Sample (N = 6,669)

The Raven scores, denoted by R, are transformed into mental ages by converting them first into IQ scores, then into mental ages. Under the assumption that IQ has a mean of 100 and a standard deviation of 15, the conversion of R to IQ is

$$\mathrm{IQ} = 15\,\frac{2.82 - R}{1.43} + 100 \qquad (15)$$

where R denotes a cell mean for a particular family-size/birth-order category. The quantity R - 2.82 is inverted in equation 15 because the Raven test associates high scores with low intelligence. The IQ scores are then converted to mental ages using equation 2.

RESULTS

We fit the confluence model to (1) aggregate-level data from the 1975 WLS sample, (2) individual-level data from the 1975 WLS sample, and (3) individual-level data from the WLS sample of sibling pairs, using Method 1, nonlinear least squares.[5]

Aggregate-Level Data

The 1975 WLS sample is a sample of individuals. To make comparisons with analyses of the Dutch data, we grouped the data on individuals into a table of mean mental age by family size and birth order (Figure 2a). Because of small numbers of cases at the larger family sizes, the data are truncated at family size 6.

Figure 2a, pertaining to observed mental ages, suggests a slightly negative effect of birth order on measured intelligence, consistent with Zajonc, Markus, and Markus's (1979) claim that negative effects are to be expected among children tested at ages above 14 years. The vast majority of WLS respondents were tested at about 16 years of age (their junior year of high school). However, the slightly negative effect of birth order in Figure 2a could be an artifact of not controlling for confounding background variables that are correlated with family size and birth order.

Figures 2b and 2c show the result of a nonlinear least squares fitting of the confluence model to the data portrayed in Figure 2a. The fitting procedure assumed that all respondents were tested at age 16, that children leave the household on their 19th birthday, and that all parents uniformly have a mental age of 18. The mean birth intervals for each family size that were used in the fitting were calculated directly from the WLS data. The fits of Model 1 and Model 2 to these data are poor, as indicated by visual comparison of Figures 2b and 2c with Figure 2a. In addition, the fit for Model 1 differs considerably from the fit for Model 2, indicating that the confluence model is extremely sensitive to model specification.

For each fitted model, we calculated the sum of squared residuals, $\sum(M-\hat{M})^2$, as a measure of goodness of fit. The sum of squared residuals is 12 for Model 1 and 5 for Model 2. Visual comparisons lead us to view these values as large, indicating a poor fit.

To gain some perspective on the size of the sum of squared residuals, we also examined a third model that we call the "naive model." This model is specified as $M = \bar{M}$, where $\bar{M}$ is the grand (unweighted) mean of the mean mental ages plotted in Figure 2a. Obviously this model explains nothing, since it predicts simply that mental age equals the grand mean. It serves, however, as a useful benchmark against which the confluence model can be evaluated. The sum of squared residuals for the naive model is 2, which is smaller than the corresponding values of 12 for Model 1 and 5 for Model 2 of the confluence model. In other words, based on a least-squares fitting criterion, which is appropriate here, a naive model that explains nothing fits better than a complex model that is supposed to explain everything.[6] Thus, the benchmark results for the naive model show how poorly the confluence model fits the data. (One can calculate $r_{\hat{M}M}$ as -.39 for Model 1 and .37 for Model 2, but, as explained earlier, these correlations are inappropriate as measures of goodness of fit.)

[5] Method 2, which is the method of maximizing $r_{\hat{M}M}$, is not used, primarily because Method 2 is invalid, but also because a computer algorithm for Method 2 is not available. Zajonc and Bargh (1980a) did not use a computer algorithm; they simply tried out different combinations of values of k, ω_1, and ω_2 until they found a combination that seemed to maximize $r_{\hat{M}M}$ (Bargh, personal communication). Not surprisingly, their solutions are only approximate, as indicated by Table 1 in their paper, which shows a fitted ω_2 value of 150 for four of the six data sets to which they fitted the confluence model. Had a systematic algorithm been used, identical values of ω_2 for four out of six independent data sets almost certainly would not have been obtained.

[6] To those unfamiliar with nonlinear models, it may seem counterintuitive that the naive model could fit better than the three-parameter confluence model. The Appendix provides an example of a very simple nonlinear model (unrelated to the confluence model) that illustrates that such a result is indeed possible.

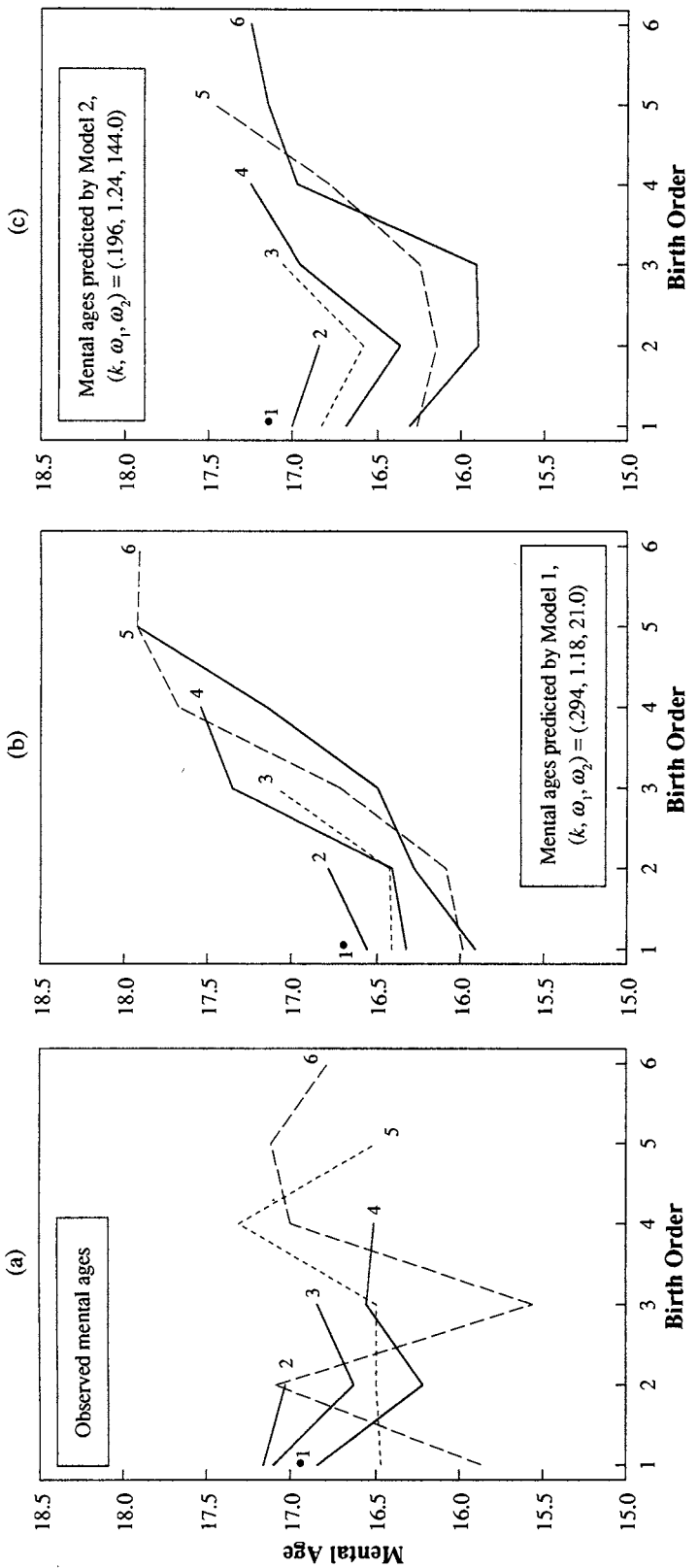


Figure 3. Fits of the Confluence Model to Individual-Level Data: 1975 WLS Subsample (N = 1,015)

Individual-Level Data

Although the confluence model has been applied to aggregate-level data, it seeks to explain individual behavior and should therefore be applied to data on individuals. Aggregate-level data, such as Zajonc and his associates have used, are inadequate to test theoretical propositions about individual behavior because (1) they greatly exaggerate the actual correlations produced by individual-level data (Robinson 1950; Velandia et al. 1978) and (2) they do not ordinarily permit the testing of alternative explanations (Rodgers 1988).

Tests based on between-family data. The WLS samples allow us to apply the confluence model to both between-family data and within-family data at the individual level. By between-family data, we mean data that include no more than one individual from any given family. Thus, comparisons of intelligence scores between any two birth orders necessarily involve persons in different families. The principal disadvantage of between-family data is that families tend to differ on a host of background factors that may confound the analysis.

To reduce excessive computer time in converging to a solution for the confluence model parameters, we selected a random subsample of 1,015 cases from the larger WLS sample. Figure 3a, which is identical in format to Figure 2a, shows observed mental age by family size and birth order for the subsample. Because of small numbers of cases at higher family sizes, Figure 3a shows greater random fluctuation than Figure 2a. Except for sampling error, the two figures should coincide.

Figures 3b and 3c show predicted mental ages derived from the confluence model applied to the subsample at the individual level for Model 1 and Model 2. The plotted points in these figures represent averages of predicted mental ages for individuals in each birth-order/family-size category. Comparison of Figures 3b and 3c with Figure 3a indicates that the fit of the confluence model to the data is very poor. Comparison of Figure 3b with Figure 3c shows that the fit for Model 1 differs considerably from that for Model 2, again indicating that results from the confluence model are highly sensitive to model specification.

Analysis of the sum of squared residuals also indicates a poor fit. The sum of squared residuals is 6,429 for Model 1 and 5,880 for Model 2. By way of comparison, the sum of squared residuals for the naive model, $M = \overline{M}$ (this time $\overline{M}$ is calculated from the mean mental ages plotted in Figure 3a), is 5,722, again indicating that the naive model fits better than either Model 1 or Model 2 of the confluence model. In this case, however, the improvement is small. (One can calculate $r_{\hat{M}M}$ as .01 for Model 1 and .08 for Model 2, but, as explained earlier, these correlations are invalid as measures of goodness of fit.)

Tests based on within-family data. A better test of the confluence model utilizes data on more than one sibling within each family. For the within-family analysis, we selected a random subsample of 507 sibling pairs (1,014 individuals) from the WLS sample of sibling pairs and applied the confluence model in the same way as before.

Figure 4, which is identical in format to Figures 2 and 3, shows observed and predicted mental ages by family size and birth order for the subsample. Figure 4a shows an irregular pattern that differs from the patterns in Figures 2a and 3a, probably due mainly to sampling variability. Comparison of Figures 4b and 4c with Figure 4a indicates that the fit of the confluence model to the data is very poor. Moreover, the fit for Model 1 differs considerably from the fit for Model 2, again indicating that results from the confluence model are highly sensitive to model specification. Model 2 fits somewhat better than Model 1, but neither fits as well as the naive model. The sum of squared residuals is 6,481 for Model 1, 5,895 for Model 2, and 5,674 for the naive model. (One can calculate $r_{\hat{M}M}$ as .01 for Model 1 and .07 for Model 2, but, as explained earlier, these correlations are invalid as measures of goodness of fit.)

REPLICATION OF THE FIT OF THE CONFLUENCE MODEL TO THE DUTCH DATA

None of our tests of the confluence model with Wisconsin Longitudinal Study data confirms the model. Consequently, we question the validity of Zajonc and Bargh's (1980a) close fits of the confluence model to a variety of other data sets. We therefore replicated Zajonc and Bargh's fit to the Dutch data. Zajonc and Bargh's results are reproduced in Figure 5. The fit, based on families with eight or fewer children, appears to be very close.

From Galbraith's earlier work (Galbraith 1982a, 1982b, 1982c, 1983), we knew that there were problems with Zajonc and Bargh's results. Galbraith took the fitted values of the model parameters (k, ω_1, and ω_2) reported by Zajonc and Bargh (1980a) for six data sets (including the

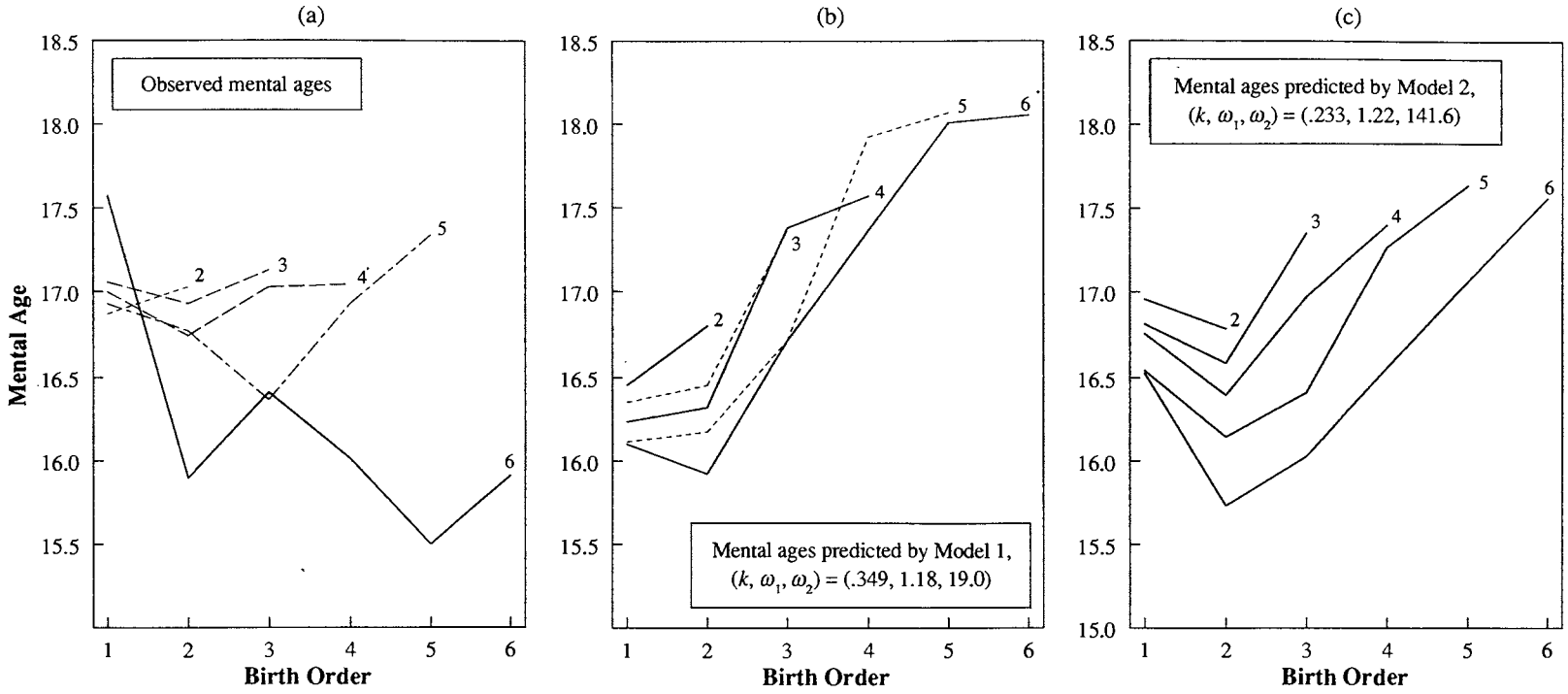


Figure 4. Fits of the Confluence Model to Individual-Level Data: 1975 WLS Sample of Sibling Pairs (N = 507 pairs)

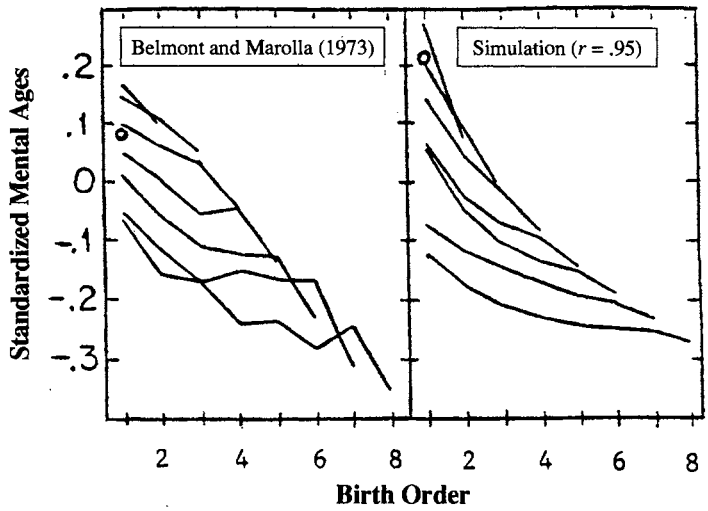


Figure 5. Observed and Predicted Standardized Mental Ages for Aggregate-Level Dutch Data

Note: Observed mental ages are in the left panel, and predicted mental ages are in the right panel. r denotes the correlation between predicted and observed mental ages.

Source: Zajonc and Bargh (1980a, Fig. 1).

Dutch data), entered these values into the Model 2 equations, generated predicted mental ages at the chronological age of testing (19), and converted these mental ages into IQ scores using the approach described earlier in this paper. We replicated Galbraith's procedure using the Dutch data and obtained the predicted mental ages shown in the second panel of Table 1. The predicted mental ages range from slightly below four years to slightly above six years. In contrast, the observed mental ages in the first panel of the table (graphed in Figure 1) range from slightly below 18 years to about 19.5 years. IQ scores, calculated from the predicted mental ages using equation 1, range from 19 to 33 and are absurdly low. These results, which are considerably more complete than those presented by Galbraith, agree with his results where comparisons can be made.[7]

Until now, the disagreement between Table 1 and Figure 5 remains unexplained. The disagreement is explained by an error in the calculation of the predicted mental ages in the right half of Figure 5. Figure 5 is based on standardized mental ages, calculated as $(M\text{-}\bar{M})/s_M$ in the case of observed mental ages and $(\hat{M}\text{-}\bar{M})/s_M$ in the case of predicted mental ages, where s_M denotes the standard deviation of M. In this calculation, $\bar{M}$ should be taken as the mean of the *observed* mental ages, and s_M should be taken as the standard deviation of the *observed* mental ages, regardless of whether M or $\hat{M}$ is being standardized. This must be done if the standardized observed mental ages and the standardized predicted mental ages are to be compared in a meaningful way. Had this been done, the very large differences between observed and predicted mental ages in Table 1 would have been preserved. But Zajonc and Bargh standardized the predicted mental ages by using the mean and standard deviation of the *predicted* mental ages instead of the mean and standard deviation of the observed mental ages (Bargh, personal communication). Because the standardization of predicted mental ages was done incorrectly, the right half of Figure 5 is incorrect. Our earlier conclusion based on Table 1, that the fit to the Dutch data based on Method 2 is very poor, stands.

Despite the very large differences between observed and predicted mental ages in Table 1, the correlation between observed and predicted mental ages in this table is indeed .95, as reported by Zajonc and Bargh. As already explained, however, the correlation is an inappropriate measure of goodness of fit, and the method of maximizing $r_{\hat{M}M}$ is not a valid fitting method.

[7] For only children, first-born children in two-child families, and second-born children in two-child families, Galbraith (1982c) obtained predicted mental ages of 5.66, 6.15, and 5.13, respectively, whereas we obtained 5.68, 6.23, and 5.15. The source of these small discrepancies is unclear.

Table 1. Observed and Predicted Mental Age by Family Size and Birth Order: Dutch Males

	Birth order								
Family size	1	2	3	4	5	6	7	8	9
Observed mental ages									
1	19.27								
2	19.49	19.29							
3	19.43	19.32	19.14						
4	19.30	19.18	19.09	18.88					
5	19.16	19.02	18.85	18.88	18.61				
6	19.05	18.84	18.69	18.66	18.65	18.34			
7	18.86	18.68	18.52	18.57	18.53	18.53	18.11		
8	18.83	18.56	18.52	18.32	18.33	18.19	18.31	17.99	
9	18.51	18.46	18.25	18.18	18.06	18.00	17.89	18.02	17.77
Predicted mental ages									
1	5.68								
2	6.23	5.15							
3	5.85	5.37	4.79						
4	5.56	5.13	4.86	4.52					
5	5.18	4.79	4.58	4.48	4.24				
6	5.11	4.67	4.41	4.28	4.22	4.08			
7	4.55	4.35	4.21	4.09	3.99	3.96	3.88		
8	4.29	4.06	3.90	3.80	3.75	3.72	3.71	3.61	

Notes: The observed mental ages were calculated from Belmont and Marolla's (1973) data and are graphed in Figure 1. Values of (k, ω_1, ω_2) for calculating predicted mental ages were taken from Table 1 of Zajonc and Bargh (1980). Childspacing estimates, up to the eighth child, were taken from Table 1 of Markus and Zajonc (1977). When entered into the Model 2 equations, these parameter estimates and childspacing estimates yield the predicted mental ages shown above. Because of truncation at family size 8 in Markus and Zajonc's childspacing estimates, the predicted mental ages shown here are also truncated at family size 8. (Zajonc and Bargh presented predicted mental ages in standardized form, shown in Figure 5, not in the unstandardized form shown here.)

In sum, Zajonc and Bargh's fit to the Dutch data is actually very poor. The excellent fit they obtained is an artifact of an incorrect standardization procedure. But even if they had found a good fit, the results would still have to be discarded, because the method of fitting the confluence model by maximizing $r_{\hat{M}M}$ is fatally flawed.

Nonlinear Least-Squares Fits of the Confluence Model to the Dutch Data

Because Method 2, the method of maximizing $r_{\hat{M}M}$, yields invalid results, we revert to Method 1 using nonlinear least squares to fit the confluence model to the Dutch data. Zajonc and Bargh (1980a) do not present results for the method of nonlinear least squares, although they state in a footnote that the method of nonlinear least squares "does not generate more accurate parameter estimates than the one we employed."

Figure 6 shows our nonlinear least-squares fits to the Dutch data for Model 1 and Model 2. Neither fit is good when compared with the graph of the observed data in Figure 1, but at least the predicted and observed mental ages are in the same range. In this sense, the parameter estimates obtained by nonlinear least squares are considerably more accurate than those published by Zajonc and Bargh, despite the authors' claim to the contrary.

Visual comparison of Figures 1 and 6 suggest that both Model 1 and Model 2 fit the data poorly. This conclusion is reinforced by examination of the sum of squared residuals. The sum of squared residuals is high for both models — 66 for Model 1 fitted by nonlinear least squares and 40 for Model 2. The sum of least-squares residuals is 7,313 for Model 2 fitted by the method of maximizing the correlation. By way of comparison, the sum of squared residuals for the naive

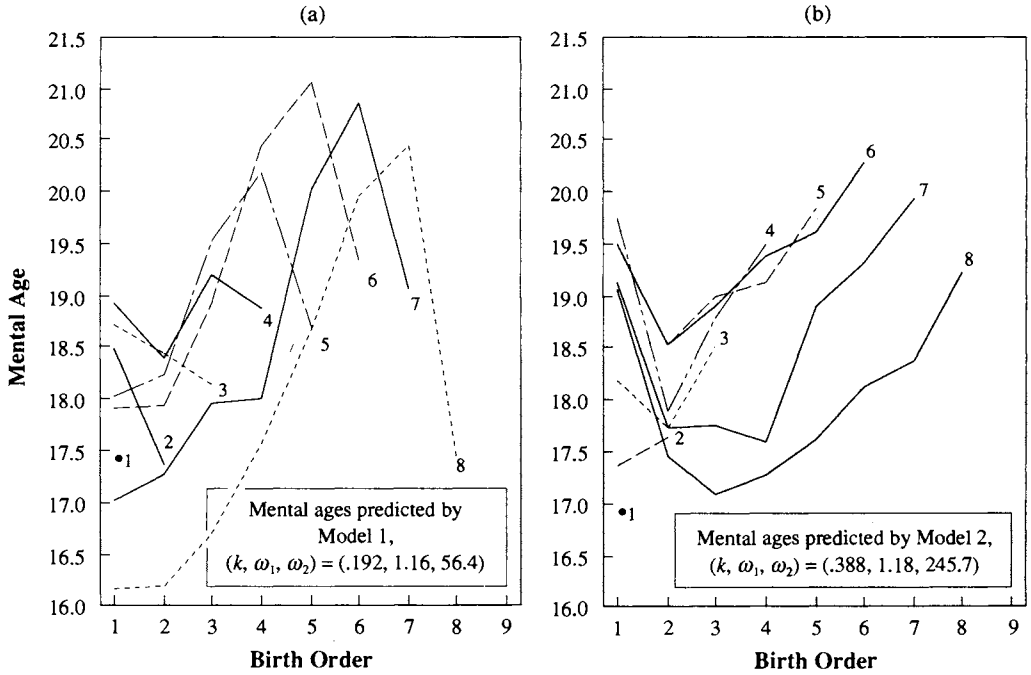


Figure 6. Nonlinear Least Squares Fits of the Confluence Model to Aggregate-Level Data: Dutch Males

model is 5. Thus, the naive model fits the Dutch data much better than does the confluence model. The fit of the confluence model to the Dutch data is very poor indeed. (One can calculate $r_{\hat{M}M}$ as -.16 for Model 1 and -.16 for Model 2, where both models are fitted by nonlinear least squares, but these correlations are invalid as measures of goodness of fit.)

IS THERE ANY INFLUENCE OF BIRTH ORDER IN THE WISCONSIN DATA?

Our reanalysis of the Dutch data leads us to conclude that confluence theory, as represented by its mathematical form, does not explain the patterns in the Dutch data. There may be a relationship between birth order and measured intelligence in the Dutch data, but, if so, the confluence model clearly does not account for the relationship. Nor does the confluence model explain any relationship between birth order and intelligence that may exist in the Wisconsin data.

Because our findings for the confluence model are so negative, we conducted a simple test to determine whether there is any relationship between birth order and intelligence in the Wisconsin data. We suspected not, because so many other studies have failed to detect such a relation when individual-level data are analyzed.

Our analysis is based on a subset of the WLS sibling sample, consisting of 1,131 pairs for whom (1) test scores were available for both the primary respondent and the random sibling, (2) all ever-born siblings survived to 1975, and (3) the primary respondent (though not necessarily the random sibling) was living with both parents in 1957, approximately one year after the testing of the primary respondent. Restrictions (2) and (3) were imposed in order to standardize and clarify the behavioral meaning of birth order insofar as possible.

For this subsample, we created a variable, ΔIQ, defined as $IQ_L - IQ_H$, where L denotes the lower birth-order member of the pair and H denotes the higher birth order member of the pair. We then computed mean values of ΔIQ, denoted as $\overline{\Delta IQ}$, for different pair types, as defined by the birth orders of the two members of the pair, and tested whether these means differed significantly from zero. Because the two members of a sibling pair have common parents and grew up in the same household, genetic and environmental factors are approximately controlled. Therefore, it was not necessary to introduce control variables into the analysis. Our analytical strategy is similar to that used by Galbraith (1982a), who also used a sample of sibling pairs.

Table 2. Mean Within-Pair IQ Differences: WLS Sample of Sibling Pairs

Pair type indexed by birth order	Mean value of ΔIQ	*N*	*p*-value
1, 2	-.34	533	.60
1, 3	-1.07	160	.44
2, 3	.24	185	.83
1, 2 or higher	.06	802	.91
Earlier than last, last	-.45	746	.43
First, middle	.72	454	.35
First, last	-1.06	279	.26
Middle, last	-.15	398	.86

Notes: ΔIQ denotes the within-pair IQ difference, calculated as IQ_L - IQ_H, where *L* denotes the lower birth order member of the pair and *H* denotes the higher birth order member of the pair. *N* denotes the number of pairs. The *p*-value is the observed level of significance.

Results are shown in Table 2 for those pair types with adequate numbers of cases. In no case does $\overline{\Delta IQ}$ come close to differing significantly from zero. There appears to be a slight tendency for first- and last-born siblings to have higher IQs than middle-born siblings, and for last-born siblings to have slightly higher IQs than first-born siblings, but again, none of these differences is anywhere near statistically significant. In sum, birth order has no discernible effect on measured intelligence in this sample of sibling pairs.

CONCLUSIONS

Our tests of the confluence model with data from the Wisconsin Longitudinal Study (WLS) provide no confirmation of confluence theory. Whether we use aggregate-level data, as Zajonc and his colleagues did, or individual-level data, we find no support for the theory.

Although Zajonc and Bargh (1980a) found birth-order effects in six aggregate-level data sets and obtained what appear to be excellent fits of the confluence model to these data, their method of fitting the model by maximizing the correlation between observed and predicted mental ages is flawed. Their method superimposes a meaningless linear model on the original confluence model, thereby unintentionally transforming a three-parameter model into a five-parameter model. This hybrid fitting procedure minimizes a sum of squared residuals, but the residuals that appear in the sum are the wrong ones. Therefore, Zajonc and Bargh's fitting procedure is statistically unsound, and all results obtained from it are in error.

Zajonc and Bargh obtained seemingly close fits only because they compounded their error by incorrectly standardizing predicted mental ages. When predicted mental ages are correctly standardized, observed and predicted mental ages are far apart, despite the high correlation between them.

An alternative, statistically sound, fitting procedure is nonlinear least squares. When we applied this method to the Wisconsin data and the Dutch data, we found that the confluence model fitted the data very poorly. A purposely naive model that predicts mental age simply as the grand mean of mental age actually does better than the highly complex confluence model in that the naive model yields a smaller sum of squared residuals between observed and predicted mental ages. We also found that the results of nonlinear least squares fitting of the confluence model are highly sensitive to changes in model specification, further eroding confidence in the confluence model.

In the context of the confluence model, the sum of squared residuals is an appropriate measure of goodness of fit, whereas the correlation (or its square) between observed and predicted mental ages is not. One should not fit the model by nonlinear least squares and then use the correlation coefficient as a measure of goodness of fit.

Although we cannot explain the presence of birth-order effects in the Dutch data and other aggregated data, it is clear that the confluence model does not explain them. Moreover, the fact that the direction of these birth-order effects varies from one data set to another suggests that the effects may not be general or repeatable. We suspect that they are an artifact of inadequate controls for confounding background variables.

Because of the failure of the confluence model to explain birth-order effects in either the WLS samples or the Dutch data, we undertook an alternative analysis of the WLS sample of sibling pairs to determine whether there are any statistically significant differences by birth order in the measured intelligence of siblings. There are none. A key strength of this sibling-pair analysis is that it adequately controls for confounding variables.

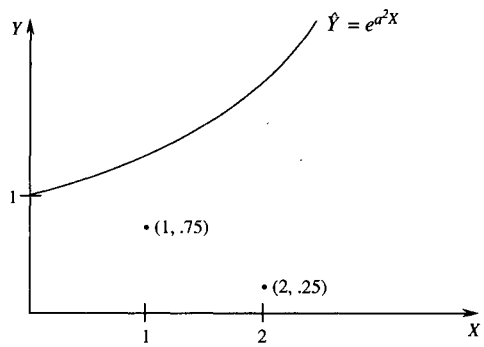
On the basis of our findings for both the WLS data and the Dutch data, we conclude that confluence theory, despite its ingenuity and intuitive appeal to many social scientists, does not hold up under careful scrutiny. It may even be a theory that attempts to explain a social phenomenon that does not exist.

ROBERT D. RETHERFORD is Research Associate and Assistant Director at the Population Institute, East-West Center, and affiliate graduate faculty in the Department of Sociology, University of Hawaii, Manoa. His current work focuses on evolutionary aspects of the demographic transition, and on analyses of current population trends in Asia and the Pacific.

WILLIAM H. SEWELL is Vilas Research Professor Emeritus, University of Wisconsin, Madison. He continues his research activities there and at the Population Institute, East-West Center, in Honolulu, where he is a Visiting Research Fellow during the winter months. His most recent publication is "Some Reflections on the Golden Age of Social Psychology," Social Psychology Quarterly *52 (1989), pp. 169-85.*

Appendix. A Simple Example Illustrating That a Naive Model, $Y = \bar{Y}$, Can Fit Better Than a Nonlinear Model

Consider the nonlinear model $\hat{Y} = e^{a^2X}$, to be fitted to two data points, (1, .75) and (2, .25), shown below. Obviously, $\Sigma(Y-\hat{Y})^2$ is minimized by setting $a = 0$, i.e., the least-squares solution is $\hat{Y} = e^{0X} = 1$, which is a horizontal line. (A rigorous proof, though simple, is omitted.)



For the naive model,

$$\Sigma(Y-\hat{Y})^2 = \Sigma(Y-\bar{Y})^2$$
$$= (.75 - .50)^2 + (.25 - .50)^2$$
$$= .125$$

whereas for the nonlinear model,

$$\Sigma(Y-\hat{Y})^2 = (.75 - 1)^2 + (.25 - 1)^2$$
$$= .625 .$$

Since .125 < .625, the naive model fits better than the nonlinear model.

Note that applying one of the usual formulae (valid for linear models) for the square of the correlation coefficient to the nonlinear model yields

$$1 - \frac{\Sigma(Y-\hat{Y})^2}{\Sigma(Y-\bar{Y})^2} = 1 - \frac{.625}{.125}$$
$$= -4 .$$

Since a negative result is impossible according to the usual interpretation of the formula, this example illustrates that one cannot generally apply the formula to nonlinear models.

REFERENCES

Adams, Bert N. 1972. "Birth Order: A Critical Review." *Sociometry* 35:411-39.

Belmont, Lillian and Francis A. Marolla. 1973. "Birth Order, Family Size, and Intelligence." *Science* 182:1096-101.

Berbaum, Michael L. 1985. "Explanation and Prediction: Criteria for Assessing the Confluence Model." *Child Development* 56:781-84.

Berbaum, Michael L., Gregory B. Markus, and Robert B. Zajonc. 1982. "A Closer Look at Galbraith's 'Closer Look.'" *Developmental Psychology* 19:174-80.

Berbaum, Michael L. and Richard L. Moreland. 1980. "Intellectual Development Within the Family: A New Application of the Confluence Model." *Developmental Psychology* 16:506-15.

_____. 1985. "Intellectual Development within Transracial Adoptive Families: Retesting the Confluence Model." *Child Development* 56:207-16.

Berbaum, Michael L., Richard L. Moreland, and Robert B. Zajonc. 1986. "Contentions over the Confluence Model: A Reply to Price, Walsh, and Vilburg." *Psychological Bulletin* 100:270-74.

Blake, Judith. 1981. "Family Size and the Quality of Children." *Demography* 18:421-42.

_____. 1989. *Family Size and Achievement.* Berkeley and Los Angeles: University of California Press.

Dixon, Wilfred J., Morton B. Brown, Laszlo Engelman, Mary Ann Hill, and Robert I. Jennrich. 1988. *BMDP Statistical Software Manual.* Berkeley: University of California Press.

Ernst, Cecile and Jules Angst. 1983. *Birth Order: Its Influence on Personality.* Berlin: Springer-Verlag.

Galbraith, Richard C. 1982a. "Sibling Spacing and Intellectual Development: A Closer Look at the Confluence Models." *Developmental Psychology* 18:151-73.

_____. 1982b. "Just One Look Was All It Took: Reply to Berbaum, Markus, and Zajonc." *Developmental Psychology* 18:181-91.

_____. 1982c. "The Confluence Model and Six Divergent Data Sets: Comments on Zajonc and Bargh." *Intelligence* 6:305-10.

_____. 1983. "Individual Differences in Intelligence: A Reappraisal of the Confluence Model." *Intelligence* 7:185-94.

Galton, Francis. 1874. *English Men of Science: Their Nature and Nurture.* London: Macmillan.

Grotevant, Harold D., Sandra Scarr, and Richard A. Weinberg. 1977. "Intellectual Development in Family Constellations with Adopted and Natural Children: A Test of the Zajonc and Markus Model." *Child Development* 48:1699-703.

Hauser, Robert M. and William H. Sewell. 1985. "Birth Order and Educational Attainment in Full Sibships." *American Educational Research Journal* 22:1-23.

Hauser, Robert M., William H. Sewell, and Brian R. Clarridge. 1982. "The Influence of Family Structure on Socioeconomic Achievement: A Progress Report." Working Paper 82-59. Center for Demography and Ecology, University of Wisconsin, Madison.

Lindert, Peter H. 1977. "Sibling Position and Achievement." *Journal of Human Resources* 12:198-219.

Markus, Gregory B. and Robert B. Zajonc. 1977. "Family Configuration and Intellectual Development: A Simulation." *Behavioral Science* 22:137-42.

McCall, Robert B. 1985. "The Confluence Model and Theory." *Child Development* 56:217-18.

Melican, Gerald J. and Leonard S. Feldt. 1980. "An Empirical Study of the Zajonc-Markus Hypothesis for Achievement Test Score Declines." *American Educational Research Journal* 17:5-19.

Mercy, James. A. and Lala C. Steelman. 1982. "Familial Influence on the Intellectual Attainment of Children." *American Sociological Review* 47:532-42.

Olneck, Michael R. and David B. Bills. 1979. "Family Configuration and Achievement: Effects of Birth Order and Family Size in a Sample of Brothers." *Social Psychology Quarterly* 42:135-48.

Outhit, M. C. 1933. "A Study of the Resemblance of Parents and Children in General Intelligence." *Archives of Psychology* 149:1-59.

Page, Ellis B. and Gary M. Grandon. 1979. "Family Configuration and Mental Ability: Two Theories Contrasted with U.S. Data." *American Educational Research Journal* 16:257-72.

Price, Gary G., Daniel J. Walsh, and William R. Vilberg. 1984. "The Confluence Model's Good Predictions of Mental Age Beg the Question." *Psychological Bulletin* 96:195-200.

Retherford, Robert D. and William H. Sewell. 1988. "Intelligence and Family Size Reconsidered." *Social Biology* 35:1-40.

Robinson, William S. 1950. "Ecological Correlation and the Behavior of Individuals." *American Sociological Review* 15:351-57.

Rodgers, Joseph L. 1981. "Effects of Family Configuration on Mental Development." University Microfilms International No. 81-25,613.

_____. 1984. "Confluence Effects: Not Here, Not Now!" *Developmental Psychology* 20:321-31.

_____. 1988. "Birth Order, SAT, and Confluence: Spurious Correlations and No Causality." *American Psychologist* 43:476-77.

SAS Institute, Inc. 1987. SAS/STAT Guide for Personal Computing, Version 6 Edition. Cary, NC: SAS Institute, Inc.

Schooler, Carmi. 1972. "Birth Order Effects: Not Here, Not Now!" *Psychological Bulletin* 78:161-75.

Sewell, William H. and Robert M. Hauser. 1980. "The Wisconsin Longitudinal Study of Social and Psychological Factors in Aspirations and Achievements." Pages 59-99 in *Research in Sociology of Education and Socialization* (vol. 1), edited by A.C. Kerckhoff. Greenwich, CT: JAI Press.

Steelman, Lala C. 1985. "A Tale of Two Variables: A Review of the Intellectual Consequences of Sibship Size and Birth Order." *Review of Educational Research* 55:353-86.

_____. 1986. "The Tale Retold: A Response to Zajonc." *Review of Educational Research* 56:373-77.

Steelman, Lala C. and John T. Doby. 1983. "Family Size and Birth Order as Factors on the IQ Performance of Black and White Children." *Sociology of Education* 56:101-09.

Steelman, Lala C. and James A. Mercy. 1980. "Unconfounding the Confluence Model: A Test of Sibship Size and Birth-Order Effects on Intelligence." *American Sociological Review* 45:571-82.

_____. 1981. "Family Structure and IQ: Artifact of Condition at Birth?" *American Sociological Review* 46:486-88.

Velandia, Wilson, Gary M. Grandon, and Ellis B. Page. 1978. "Family Size, Birth Order, and Intelligence in a Large South American Sample." *American Educational Research Journal* 15:399-416.

Zajonc, Robert B. 1975. "Birth Order and Intelligence: Dumber by the Dozen." *Psychology Today* 8:37-43.

_____. 1976. "Family Configuration and Intelligence." *Science* 192:227-36.

_____. 1983. "Validating the Confluence Model." *Psychological Bulletin* 93:457-80.

_____. 1986a. "Family Factors and Intellectual Test Performance: A Reply to Steelman." *Review of Educational Research* 56:365-71.

_____. 1986b. "The Decline and Rise of Scholastic Aptitude Scores: A Prediction Derived from the Confluence Model." *American Psychologist* 41:862-67.

Zajonc, Robert B. and John Bargh. 1980a. "The Confluence Model: Parameter Estimation for Six Divergent Data Sets on Family Factors and Intelligence." *Intelligence* 4:349-61.

_____. 1980b. "Birth Order, Family Size, and Decline of SAT Scores." *American Psychologist* 35:662-68.

Zajonc, Robert B. and Gregory B. Markus. 1975. "Birth Order and Intellectual Development." Psychological Review 82:74-88.

Zajonc, Robert B., Hazel Markus, and Gregory B. Markus. 1979. "The Birth Order Puzzle." *Journal of Personality and Social Psychology* 37:1325-41.