



One Justified Criticism Plus Three Flawed Analyses Equals Two Unwarranted Conclusions: A Reply to Retherford and Sewell

Author(s): R. B. Zajonc, Gregory B. Markus, Michael L. Berbaum, John A. Bargh, Richard L. Moreland

Source: *American Sociological Review*, Vol. 56, No. 2 (Apr., 1991), pp. 159-165

Published by: [American Sociological Association](#)

Stable URL: <http://www.jstor.org/stable/2095776>

Accessed: 16/02/2011 10:44

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=asa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Sociological Association is collaborating with JSTOR to digitize, preserve and extend access to *American Sociological Review*.

<http://www.jstor.org>

ONE JUSTIFIED CRITICISM PLUS THREE FLAWED ANALYSES EQUALS TWO UNWARRANTED CONCLUSIONS: A REPLY TO RETHERFORD AND SEWELL*

R. B. ZAJONC
University of Michigan

GREGORY B. MARKUS
University of Michigan

MICHAEL L. BERBAUM
University of Alabama

JOHN A. BARGH
New York University

RICHARD L. MORELAND
University of Pittsburgh

Retherford and Sewell claim that the confluence model of intellectual development is fatally flawed because: (1) mathematical simulations of that model generate numerical values that do not coincide with data they were claimed to fit, once those data are converted to mental age values; (2) the simulations of the confluence model also fail to fit aggregate patterns in the Wisconsin Longitudinal Study; and (3) individual-level analyses of the Wisconsin data contradict predictions of the confluence model. These criticisms are shown to be either irrelevant, erroneous, or both. Retherford and Sewell further conclude that birth-order effects may be "a social phenomenon that does not exist." We disagree. Far from contradicting the confluence model, the Wisconsin data lend further support to it.

We thank Retherford and Sewell (1991, henceforward RS) for their interest in our work.¹ At least some of their criticism of the confluence model is well taken. The Wisconsin data that they present may, when analyzed more fully, shed additional light on the verisimilitude of the confluence model of intellectual development. However, we strongly disagree with their conclusions. They assert 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" (p. 156). They are wrong on both counts.

FAMILY CONFIGURATION EFFECTS ON INTELLIGENCE ARE REAL

In a series of published reports, we have presented data from different countries, different decades, using different tests of intellectual ability, at different ages of testing, and all show statistically significant relationships between family configuration variables and intelligence. For example,

* Direct all correspondence to R. B. Zajonc, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 48106-1248.

¹ RS have been gracious in sharing their data with us and have replied promptly to our inquiries, despite the fact that we are separated by five time zones.

Figure 1 displays mean intellectual performance for three data sets arrayed by birth order and family size. The data are based on three different tests of intellectual performance and were collected in three different countries by three independent teams of researchers. Figure 1a shows the Belmont-Marolla (1973) data on 386,114 Dutch recruits in World War II; Figure 1b displays Breland's (1974) sample of 794,589 eleventh-grade students from 17,608 different high schools in the United States; and Figure 1c presents mean scores on a mathematics achievement test for 82,689 Western Israeli high school students (Davis, Cahhan, and Bashi 1977). Taken together, the means in these three data sets represent over one million individual observations. The patterns are unmistakable and are highly statistically significant.² RS's full data set (shown in their Figure 2a on page 148) shows a similar pattern. Moreover, these patterns do not appear to be an artifact of unspecified "confounding variables," as RS assert. As our work and that of others (e.g., Belmont and Marolla 1973; Breland 1974) has shown, the patterns persist when various measures of socioeconomic status are introduced as control variables.

² With regard to Figure 2a, Belmont and Marolla (1973) report that the statistical significance of the birth-order effect, independent of family size, is significant at a p -value of 1×10^{-13} .

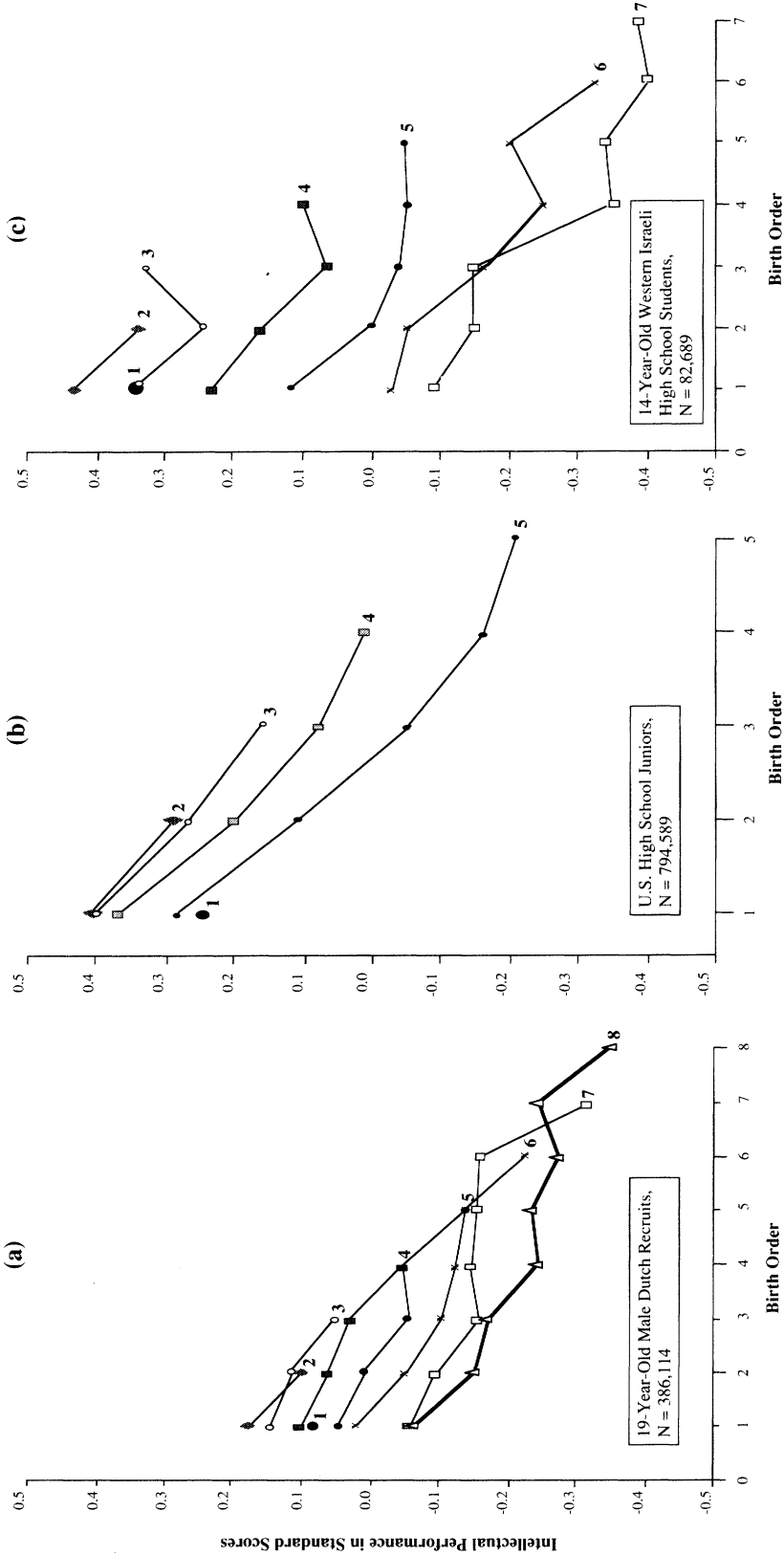


Figure 1. Average Intellectual Performance (Standard Score) by Family Size and Birth Order: Three Samples

Note: Figure 1a shows Raven Progressive Matrices test scores of 19-year-old male Dutch recruits (taken from Belmont and Marolla 1973). Figure 1b shows National Merit Scholarship Qualification Test scores for U.S. high school juniors (taken from Breland 1974). Figure 1c shows mathematics performance scores for 14-year-old Western Israeli high school students (taken from Davis, Cahan, and Bashi 1977). For purposes of comparison, the scores on the various tests were transformed to standard form.

FITTING THE CONFLUENCE MODEL TO DATA

As for the assertion that the confluence model does not hold up under scrutiny, RS base that claim on three of their findings: (1) mathematical simulations derived from the confluence theory generated numerical values that did not coincide with published data they were claimed to fit, once those data were converted to mental age values; (2) the mathematical simulations of the confluence theory also failed to fit aggregate patterns in the Wisconsin Longitudinal Study; and (3) individual-level analyses of the Wisconsin data contradicted predictions of the confluence model. We argue that these criticisms are either irrelevant, erroneous, or both.

Mathematical Simulations of the Confluence Model Were Never Intended to Generate Mental Age Scores

Zajonc and his colleagues have developed various mathematical simulations to illustrate and evaluate the theoretical ideas behind the confluence model (Zajonc and Markus 1975; Zajonc 1976). RS assert that those simulations failed to generate numerical values that resemble plausible mental age scores. Therefore, the simulations cannot reproduce test scores they were purported to fit (such as the Belmont-Marolla data) once those scores are transformed into mental ages. They then conclude that the confluence theory *itself* is fatally flawed. This would be a telling criticism if confluence modelers had claimed they were trying to model mental age. But they have made no such claim.

Repeatedly, the confluence modelers have taken pains to state that the construct of interest to them is “absolute intellectual level” (e.g., Zajonc 1976, p. 227), which, although “similar to mental age” (Markus and Zajonc 1977, p. 139) is clearly not identical to mental age. Mental age is normed so that its mean increases linearly with chronological age — indeed, mean mental age *equals* chronological age. In contrast, in the confluence model absolute intellectual level is *not* normed; it is an absolute quantity (call it “smarts” if you like, as economists refer to “utils” in their models) that is explicitly posited (based on empirical evidence) to grow *nonlinearly* as individuals mature. The hypothetical nature of the construct “absolute intellectual level” and its distinctness from mental age is evident when Zajonc (1976, p. 227) employs a numerical example in

which the intellectual levels of adult parents are set equal to “30 arbitrary units,” and the growth curve is posited to be **S**-shaped rather than the straight line that defines the developmental trajectory of mean mental age.

The mathematical simulations based on the confluence ideas were thus never intended to generate mental ages as outputs. They were intended to generate “absolute intellectual levels” measured in an arbitrary metric. Presumably, “absolute intellectual level” will correlate reasonably well with scores on a particular test of mental ability, but the metrics will differ. The appropriate gauge of a simulation’s success, then, is how well it reproduces the *pattern* of family effects — and hence, how well the simulation’s output values *correlate* with observed test scores, not whether they *equal* the observed test scores. By that appropriate standard, simulations of the confluence model have done quite well — as RS acknowledge.

To give credit where it is due, RS make a valid point in taking one aspect of Zajonc and Bargh’s (1980) simulation to task. Despite the arbitrariness of the metric of “absolute intellectual level,” if adults are assigned a score of, say, 19, then a successful simulation of the confluence model ought to generate values for offspring at maturity that are in the neighborhood of 19, and not values like 5 or 50. The Zajonc-Bargh model (with the parameter values published by them) fails in that regard, and we are grateful to RS for drawing this to our attention.

However, the more important substantive point ought not be lost in a flurry of commentary on sums of squares, nonlinear estimation algorithms, and so on. That substantive point is this: *Precise numerical values aside, the confluence model yields predictions that coincide well with a variety of complex patterns that are evident in different data sets.* For example, qualitative predictions of the Zajonc-Bargh simulation corresponded well with observed patterns among Dutch military recruits and American high school students who took the National Merit Scholarship test, two data sets for which there are (in the aggregate) clear negative correlations between birth order and test scores. At the same time, the model was compatible with data on younger populations (French and Scottish) for which there are no simple birth-order effects — a pattern not previously understood.

The values generated by the Zajonc-Bargh simulation differ from observed values by a nearly constant amount. Therefore, if strict numerical

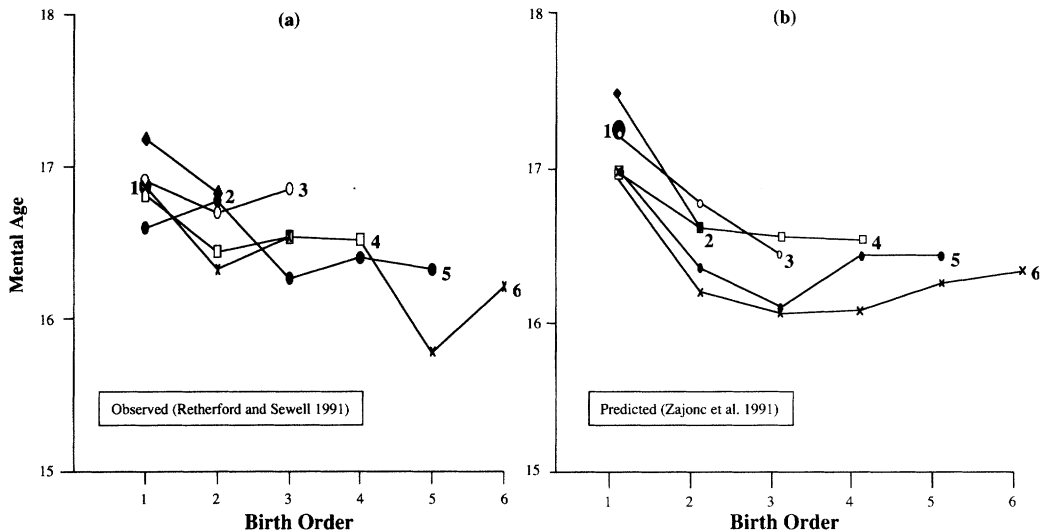


Figure 2. Observed and Predicted Mental Ages by Family Size and Birth Order: 1975 WLS Data

Note: Figure 2a shows observed data from Retherford and Sewell (1991); Figure 2b shows predicted mental age using the Markus and Zajonc (1977) version of the confluence model (parameter values estimated from Belmont-Marolla data).

correspondence with observed data is desired, the model can be rescued simply by adding a constant term to it (e.g., to the numerical evaluation of the child's intellectual environment to stand in for aspects of that environment other than the family, perhaps).

Furthermore, other published mathematical simulations of the confluence model do generate numerical values for mature offspring that compare closely with the values set for parents. In the Markus-Zajonc (1977) simulation, for example, when parents are assigned an average absolute intellectual level of 19.0, the output values generated by the simulation of the Belmont-Marolla data range between 19.4 (for the first-born of two children) and 16.0 (for the last-born of eight children), with a correlation of .95 between observed and predicted values.³

³ RS transform the mean categorized Raven test scores (published by Belmont and Marolla 1973) into mental ages. Such a transformation assumes that: (1) a one-to-one relationship exists between Raven test scores and mental age; (2) the reduction of the (unavailable) Raven scores into the (available) categorized scores has not disturbed any such one-to-one relationship; and (3) the mean and standard deviation of the categorized Raven scores for the sample of 19-year-old male military recruits is identical to that for the general population. We have no evidence to substantiate the plausibility of these assumptions, and some reason to doubt them. To that extent, we must regard the RS transformation as being more or less

Aggregate Patterns in the Wisconsin Data — Additional Support for the Confluence Model

Output of the Markus-Zajonc (1977) simulation also conforms well to the aggregated Wisconsin data as presented by RS in their Figure 2a on page 148 (which for comparison is reproduced in our Figure 2a). To demonstrate this, we employ the simulation model described in Markus and Zajonc and use the same assumptions that RS posit: no child mortality, children leave home at age 19 and parental marriages remain intact throughout the simulation (p. 146). We use the birth-interval information for the Wisconsin sample that RS provided to us. We modify slightly the "clearly somewhat arbitrary" value they assume for parents' mental age, from 18 to 18.75. Lastly, and note well, we do not "fit" the simulation to the Wisconsin data by searching for parameter values that minimize the error sum of squares (or maximize the correlation, either); instead we use the *same* set of parameter values employed by Markus and Zajonc (1977, p. 140) to fit the Dutch (and, subsequently, the Scottish) samples.

The results of the simulation are shown in Figure 2b. *It is evident that far from contradicting the confluence model, the aggregate patterns in*

arbitrary. Regardless, a slight variant on the RS transformation of the Raven scores yields values that coincide very closely (absolutely as well as relatively) to values produced by the Markus-Zajonc simulation.

the Wisconsin data support that model. The correlation between observed and simulated values is .71, and the error sum of squares is 1.65.⁴ Thus, we now have three entirely different data sets, from different populations and based on different mental tests that conform to a single simulation using a single set of parameter values. Undoubtedly, if the parameter values were altered from those originally derived to fit different data, the correspondence between the simulation output and the Wisconsin data would be even closer.⁵

Analysis of the Disaggregated Wisconsin Data Not a Test of the Confluence Model

RS perform two individual-level analyses of the Wisconsin data. First, they fit simulations based on the confluence model to a random sample of 1,015 cases (presumably using the available information on family configuration for those cases). Second, they examine the relationship between birth order and IQ for a set of 1,131 sibling pairs. They conclude that both of these analyses yield results that cast doubt on the confluence model. We disagree.

The data and general approach employed in the first of these individual-level analyses are potentially useful in evaluating the confluence model, but as it stands the RS analysis embodies some serious deficiencies that should be corrected. First, for reasons already noted, the appropriate criterion should be the correlation between confluence model predictions and mental age values, and not absolute numerical correspondence. Three other comments:

(1) As a comparison of RS's Figures 1a and 2a shows, the relationships between mental age (on the one hand) and birth order \times family size (on the

other) are noticeably different in the subset compared with the full data set. By the luck of the draw, the subset may not be representative of the larger data set. Since the total set manifests birth-order patterns while the subset does not, there must exist another subset of that data with birth-order effects that are even more pronounced than those of the total set. Why then was that particular subset selected to document absence of birth-order effects? It would be preferable to work with the latter. If parameter values are fixed (to values based on prior aggregate-level analysis), computer time should not be an obstacle to utilizing all the data.

(2) Although the students in the Wisconsin sample were all tested as high school juniors, there is some nontrivial variation in age. A quick inspection of the sample data provided to us indicated birth years ranging from 1938 to 1941. Since the dependent variable is mental age, this 2 to 3 year variation could well swamp the relatively fine-grained effects of family configuration factors. The proper strategy would be to compare simulation results and observed values separately for individuals of the same age (or birth year).

(3) We have never asserted that family configuration alone accounts substantially for intelligence. We have not found (Berbaum and Moreland 1980, 1985; Berbaum 1985) and would not expect to find a substantial amount of individual-level variation in mental age accounted for by family configuration factors alone. What we *have* said is that family configuration effects are real, but they are subtle and may require very large samples to be evident.

⁴ A slight amount of the slippage between the RS graph (reproduced in our Figure 2a) and the simulation output (in Figure 2b) may arise from a flaw in the former. A handful of individuals who are recorded in the Wisconsin data, e.g., as "first-born of five children" were in fact members of smaller sibships *at the time of testing*; the later-born sibling(s) arrived after the time of testing and could not have had any impact on the intellectual environment of the tested child. Hence, some of the values used in calculating the mean mental age of, e.g., "first-borns of five-child families," should properly have been reallocated to the category of "first-borns of four-child families."

⁵ In their footnote 1, RS criticize simulations of the confluence model (such as the Markus-Zajonc simulation) that contain an intercept term in the posited difference equation for M_{ij} , "since a nonzero intercept

would make it impossible for mental age to level off in early adulthood." Galbraith (1982, p. 189) made a similar criticism. However, a nonzero intercept in the simulation model plays exactly the same (unremarkable) role it does in an ordinary regression equation. It captures the mean effects of omitted variables on the dependent variable of interest. (Indeed, the Markus-Zajonc simulation can be rewritten as a linear regression with three parameters.) Such variables are omitted for explicit reasons. The confluence modelers have never assumed that the few factors included in the confluence model are the only ones (or even the principal ones) that influence intelligence. To the remark that a nonzero intercept "makes it impossible for mental age to level off" beyond the developmental years, our reply is that the model applies to the developmental years and not to the years beyond. On the dangers of extrapolating models into domains for which they were not intended, and in particular on the senselessness of being too literal in interpreting intercept terms, see Rao and Miller (1971, pp. 5-6).

With regard to the second of their individual-level analyses, RS's examination of sibling pairs is largely irrelevant to the confluence model. For one thing, they compare IQ's of siblings tested at different ages, but the relevant comparison would be between siblings tested *at the same level of maturation*. The confluence model makes predictions about the relative levels of absolute intelligence for siblings at comparable stages of development (i.e., at the same age for pre-adults, or at maturity). It does not make predictions about IQ's for siblings measured at different stages of development.

Second, as the simulation output in Figure 2b makes clear, *the confluence model explicitly does not predict that intellectual ability will necessarily be correlated, positively or negatively, with birth order*. To the contrary, we consider it a positive feature of the confluence model that it leads to (empirically supported) predictions about the conditions under which the intelligence-birth-order correlation may be negative, positive, or nil. We have emphasized repeatedly that birth order *per se* is not a causal factor:

Any apparent birth order effects found in a set of data are hypothesized to be artifactual in that they may be explained solely by family size and the spacing of births. With short birth intervals, increasing order of birth will be associated with lower intelligence levels With sufficiently long gaps, however, this pattern may be mitigated or even reversed, provided the new child is born at a time when the average intellectual level of the family is greater than that when the earlier siblings were born. (Markus and Zajonc 1977, p. 138)

Other work of ours suggests that the particular age at which children are tested will confound any simple hypotheses about necessarily negative correlations between birth order and intelligence (Zajonc, Markus, and Markus 1979).

It therefore eludes us why a finding of no consistent relationship between birth order and IQ of siblings should be interpreted as a contradiction of the confluence model, given that the comparisons were made between siblings at possibly different developmental stages and without regard to birth intervals.

CONCLUSION

We have shown that, if anything, the Wisconsin data presented by RS lend additional support to the confluence model. They have performed a service by pointing out the need for a constant in one mathematical simulation of the confluence

theory. What they have not done, however, is offer any alternative theory (beyond that of unspecified "confounding background variables") to account for the observed phenomena that are consistent with the confluence theory: why increasing family size has a generally negative correlation with intelligence (Zajonc 1976); why the only child departs systematically from that negative correlation (Zajonc 1976); why birth order tends to be correlated positively with intelligence at young ages but negatively at older ages, or why the birth order and intelligence correlation is mediated by the length of inter-birth intervals (Zajonc, Markus, and Markus 1979; Zajonc 1983); why twins typically score relatively lower than average on intelligence tests, and triplets lower yet (Record, McKeown, and Edwards 1970); why annual mean SAT scores covary with lagged birth rate (Zajonc 1986); and so on.

R. B. ZAJONC is Charles Horton Cooley Distinguished Professor of Psychology and Director of the Institute for Social Research at the University of Michigan. He received his Ph.D. in Social Psychology at the University of Michigan in 1955. His work has dealt with social behavior and more recently with the emotions.

GREGORY B. MARKUS is a Professor of Political Science and Research Scientist at the University of Michigan's Institute for Social Research. Much of his research focuses on voting and public opinion. He has taught in the ICPSR Summer Program in Quantitative Analysis since 1973.

MICHAEL L. BERBAUM received his B.S. (Psychology, 1970) from the University of Illinois Urbana-Champaign and a M.S. (Biostatistics, 1978) and Ph.D. (Social Psychology, 1983) from the University of Michigan. In addition, he holds an A.M. in International Relations from the University of Southern California (1973) and an M.A. in Russian and East European Studies from the University of Michigan. At present he is a Research Social Scientist at the University of Alabama's Institute for Social Science Research and an instructor in the ICPSR Summer Program in Quantitative Methods in the Social Sciences. His research interests include social interaction and decision making in groups and teams, as well as quantitative methods.

JOHN A. BARGH is Professor of Psychology, New York University. His current research focuses on automatic influences in social perception and in the appraisal of the environment, especially the unintentionality of these influences. He recently co-edited the book *Unintended Thought* (Guilford 1989) on this issue. In 1989 he received the "Early Career Distinguished Contribution Award" from the American Psychological Association.

RICHARD L. MORELAND is an Associate Professor in the Department of Psychology at the University of Pittsburgh. His research focuses on many aspects of small groups, but he is especially interested in how such groups change over time. This interest has led him to study such topics as group formation and termination, group development, and socialization processes in groups.

REFERENCES

- 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. 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.
- Breland, Hunter M. 1974. "Birth Order, Family Configuration, and Verbal Achievement." *Child Development* 45:1011-19.
- Davis, Daniel J., Sorel Cahan, and Joseph Bashi. 1977. "Birth Order and Intellectual Development: The Confluence Model in the Light of Cross-Cultural Evidence." *Science* 196:1470-72.
- Galbraith, Richard C. 1982. "Just One Look Was All It Took: Reply to Berbaum, Markus, and Zajonc." *Developmental Psychology* 18:181-91.
- Markus, Gregory B. and Robert B. Zajonc. 1977. "Family Configuration and Intellectual Development: A Simulation." *Behavioral Science* 22:137-42.
- Rao, Potluri and Roger LeRoy Miller. 1971. *Applied Econometrics*. Belmont, CA: Wadsworth.
- Record, R. G., Thomas McKeown, and J. H. Edwards. 1970. "An Investigation of the Difference in Measured Intelligence Between Twins and Single Births." *Annals of Human Genetics* 34:11-20.
- Retherford, Robert D. and William H. Sewell. 1991. "Birth Order and Intelligence: Further Tests of the Confluence Model." *American Sociological Review* 56:141-58.
- Zajonc, Robert B. 1976. "Family Configuration and Intelligence." *Science* 192:227-36.
- _____. 1983. "Validating the Confluence Model." *Psychological Bulletin* 93:457-80.
- _____. 1986. "The Decline and Rise of Scholastic Aptitude Scores." *American Psychologist* 41:862-67.
- Zajonc, Robert B. and John A. Bargh. 1980. "The Confluence Model: Parameter Estimation for Six Divergent Data Sets on Family Factors and Intelligence." *Intelligence* 4:349-61.
- 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.

MANUSCRIPTS FOR THE

ASA ROSE MONOGRAPH SERIES

The *ASA Arnold and Caroline Rose Monograph Series* invites authors to submit manuscripts for publication.

The Series welcomes a variety of sociological work: qualitative or quantitative empirical studies, and theoretical or methodological treatises. For information, or to submit a manuscript (four copies, please), contact the Editor, Professor Teresa A. Sullivan, Rose Monograph Series, 436 Burdine Hall, Department of Sociology, University of Texas at Austin, Austin, TX 78712-1088.