The results for both the treadle and hopper groups can be explained by a single underlying mechanism, that is, the development of a set to respond (industriousness) or of a set to not respond (laziness). Presumably the treadle group learned that reinforcement is contingent on some response and they entered the autoshaping situation with a set or expectation that a similar contingency still applied. The hopper birds responded less because they anticipated a continuation of the noncontingency that existed in their first stage of training.

Our interpretation is entirely parallel to that offered by Maier *et al.* (*2*) for the learned helplessness phenomenon. If that interpretation is valid, then the significance of subjects' ability to control their environment is not restricted to commerce with aversive stimuli or to any single motivational or reward system. Thus, speculations involving presumed physiological or biochemical correlates of helplessness in terms of stress reactions (*3*) are likely to be of limited use. Further research will be necessary to validate our interpretation as well as to explore the limits of its applicability.

LARRY A. ENGBERG
GARY HANSEN
ROBERT L. WELKER
DAVID R. THOMAS
*Department of Psychology, University of Colorado, Boulder 80302*

**References and Notes**

1. M. E. P. Seligman and S. F. Maier, *J. Exp. Psychol.* **74**, 1 (1967).
2. S. F. Maier, M. E. P. Seligman, R. L. Solomon, in *Punishment and Aversive Behavior*, B. A. Campbell and R. M. Church, Eds. (Appleton-Century-Crofts, New York, 1969), pp. 299–342.
3. M. E. P. Seligman, S. F. Maier, R. L. Solomon, in *Aversive Conditioning and Learning*, F. R. Brush, Ed. (Academic Press, New York, 1971), pp. 347–400.
4. P. L. Brown and H. M. Jenkins, *J. Exp. Anal. Behav.* **11**, 1 (1968).
5. E. Gamzu and D. R. Williams, *Science* **171**, 923 (1971).
6. D. R. Williams and H. Williams, *J. Exp. Anal. Behav.* **12**, 511 (1969).
7. One subject in the control group died during the experiment and was not replaced.
8. J. B. Overmier and M. E. P. Seligman, *J. Comp. Physiol. Psychol.* **63**, 28 (1967).
9. Supported by NIH research grant HD-03486 and training grant MH-10427.

2 August 1972

## Similarity in Developmental Profile among Related Pairs of Human Infants

Wilson (*1*) presented evidence that he interprets as demonstrating that the similarity in overall level and developmental profile contour of scores on the Bayley infant test is greater for monozygotic (MZ) than for dizygotic (DZ) twins for assessments made during the first year (3, 6, 9, and 12 months) and the second year (that is, 12, 18, and 24 months) of life. The intraclass correlations expressing similarity within pairs were very high, approaching and in some cases equaling the reliability of the test.

Wilson cites the fact that the intraclass correlations for DZ twins were high for both overall level ($r = .75$ and .79, for the first and second years, respectively) and profile contour ($r = .52$ and .50), and that the size of these concordances signifies "that the differences within DZ pairs produced by gene segregation and different life experiences are comparatively small in relation to the sizable differences between pairs" (*1*, p. 917). He concludes that genetic factors are paramount in such development. However, circumstances surrounding the analysis of these data may modify this interpretation.

The degree of heritability for a trait depends on the difference between the correlations for MZ and DZ twins, not on the absolute size of the DZ correlation. For example, "broadsense heritabilities" can be calculated by taking twice the difference between $r_{MZ}$ and $r_{DZ}$ (*2*). Mating is assumed to be random in this heritability estimate, an assumption that is probably tenable since infant test scores do not relate very strongly to adult mental or personality characteristics that might influence mate selection (*3*). Wilson's data treated in this manner give heritabilities of .30 and .20 for overall level (first and second years, respectively) and .50 and .30 for profile contour. These values are lower than those reported for IQ in later childhood (*4*), and do not seem to warrant the conclusion that "infant mental development was primarily determined by the twins' genetic blueprint and that . . . other factors served mainly a supportive function" (*1*, p. 914).

The high correlations for DZ twins may derive from common nongenetic as well as genetic circumstances. In fact, since the genetic correlation for DZ twins averages .50 (they have half their genes in common), correlations between DZ twins substantially higher than .50 must reflect common environmental circumstances or assortive mating, or both these factors.

When these same analyses were performed on 142 sibling pairs from the Fels Longitudinal Study, the intraclass correlations for Gesell developmental scores were .24 and .44 for overall level (at 6 and 12 months and at 12, 18, and 24 months, respectively) and .09 and .14 for profile contour. The twin correlations reported by Wilson are two to six times these values, despite the fact that the degree of genetic overlap is the same for DZ and sibling pairs. Twins may be more similar than siblings because they share environmental circumstances (prenatal environment, stimulating familial experiences, and so forth) and because those environmental factors have their effects at the same age for twins but at different ages for siblings.

The method of analysis used by Wilson also raises some issues of interpretation. His intraclass correlations express within-pair variability relative to the appropriate total variability separately for MZ's and DZ's, before these groups are compared. Such a procedure may be justified when the variability between individuals in one group is different than that in another but this is reportedly not the case for the data presented.

A more direct approach is to contrast the within-pair variability for MZ's and DZ's by calculating the ratio of the mean square within pairs for DZ's divided by the mean square within pairs for MZ's. For Wilson's data, this procedure yields significant differences between zygosity groups for overall level during the first year ($F = 3.44$, d.f. $= 51/45$, $P < .0001$) and second year ($F = 1.74$, d.f. $=$ 46/51, $P < .03$) and for profile contour during the first year ($F = 1.94$, d.f. $= 153/135$, $P < .0001$) but not during the second year ($F = 1.18$, d.f. $= 92/102$, $P \cong .25$). Thus, when differences in similarity within pairs are considered directly, MZ's are more similar than DZ's in profile contour only during the first year of life.

The entire univariate analysis-of-variance model involving repeated measures, which was used by Wilson to estimate profile similarity (as well as the above-mentioned analyses), tends to have a positive bias toward significant effects (in this case favoring within-pair similarity) if the covariances (that is, correlations) between all pairs of scores measured on the same individuals are not equal (*5*). Such heterogeneity is almost always present in longitudinal data of this sort because adjacent-age assessments are more highly correlated than chronologically distant ones. The data in Wilson's report do not permit an estimation of the degree of positive bias with respect to significance levels or the size of intraclass correlations. A multivariate approach to the repeated variable presents a statistical alternative that does not make these covariance assumptions.

Finally, in contrast to MZ pairs which are always same-sexed, Wilson's DZ group was composed of cross-sexed as well as same-sexed pairs. Since females had higher scores than males at all ages (statistically significant at 18 months), the reported results are further qualified to an unknown extent by introducing this additional variability within DZ pairs.

From the Fels sample, we compared (i) 142 pairs of siblings, (ii) unrelated controls matched for year of birth, sex, and parental education, and (iii) unrelated pairs matched for year of birth and sex but not parental education. While siblings were more similar in the general level of their performance on Gesell tests administered at 6, 12, 18, and 24 months, there were no differences in within-pair similarity between the three groups for profile contour as assessed either by comparison of the square root of the sum of squared deviations between corresponding points or by multivariate profile analysis (*6*).

When all the data on similarity among related individuals in profile contour of mental test performance (*1, 6*) is considered, it would appear that the evidence for non-zero heritability of profile contour after the first year of life is ambiguous at best.

ROBERT B. MCCALL
*Fels Research Institute,*
*Yellow Springs, Ohio 45387*

**References and Notes**

1. R. S. Wilson, *Science* **175**, 914 (1972).
2. D. S. Falconer, *Introduction to Quantitative Genetics* (Ronald, New York, 1960), p. 185.
3. R. B. McCall, P. S. Hogarty, N. Hurlburt, *Amer. Psychol.*, in press.
4. M. P. Honzik [*Child Develop.* **28**, 215 (1957)] reports increasing correlations for general level of IQ between biological parent and adopted child as the child's age increases.
5. G. E. P. Box, *Ann. Math. Statist.* **25**, 484 (1954).
6. R. B. McCall, paper presented at the meeting of the American Psychological Association, Honolulu, September 1972.
7. ———, *Science* **170**, 644 (1970).
8. Research supported in part by NIH grants HD 04160 to the author, NIH grants FR-04437, HD-00868, and FR-00222 to the Fels Research Institute, the Grant Foundation of New York, and by the Fels Fund of Philadelphia. I thank M. Appelbaum, V. Crandall, F. Falkner, A. Jensen, and J. Loehlin for their comments on an earlier draft of this technical comment.

1 May 1972

Wilson (*1*) states that early intellectual development is "primarily determined" by genetic factors. As a major source of support for his argument, Wilson presents data indicating a high degree of within-pair concordance for overall intelligence of dizygotic twins. These data are used to illustrate the importance of between-pair variance, which is postulated to be a function of both environmental and genetic factors. Wilson then proceeds to rule out environmental variance as important by presenting results indicating insignificant correlations between socioeconomic status (SES) and early IQ. However, in ruling out environmental variance by the use of SES, Wilson ignores certain characteristics of SES which make this variable unsuitable for this purpose.

First, there is a question of the utility of SES as a valid measure of distinct environmental differences between groups. Wellman (*2*) indicated that the range of environments within any given SES level was so great as to make suspect any statements using SES as a measure of environment. Pavenstedt (*3*), studying the types of environments encountered within a given socioeconomic group, and Tulkin (*4*), investigating the role of socioeconomic status, race, and family environment in school achievement, pointed out that controlling for SES does not at all control for types of experiences encountered. Rather than measure distinct homogenous environments, SES appears to measure overlapping heterogenous environments. This overlap between supposed categories would minimize any correlation such as that between SES and intelligence. Evidence on this point can be seen in the data presented by Wilson. Because dizygotic twins are genetically similar to full sibs, the maximum expected genetic correlation between dizygotic twins should be .50 (*5*). However, the dizygotic twins' correlations for mean and overall level, presented by Wilson, are greater than this expected value (of the six mean correlations, three are $\geqq .61$ and three are $\geqq .72$; both of the two overall level correlations are $\geqq .75$). This degree of correlation indicates the operation of nongenetic factors in the dizygotic twins' mental test performance. Since the correlations between SES and IQ reported by Wilson are essentially zero-order, it seems that whatever the nongenetic factors affecting the twins' mental test performance, they are not being adequately reflected by SES indices.

The heterogenous nature of the SES variable reflects a second, more serious problem for the use of this variable as an experiential measure. As human behavioral ecologists (*6*) have pointed out, SES differences tell us nothing at all about the specific, proximal experiences being undergone by a particular child. It is these specific experiences that will affect the development of the child's intelligence—not distal labels such as SES, which may or may not reflect these experiences. This distinction between distal as opposed to proximal experiences is reflected by the available evidence on early intelligence and experience. Thus, significant relations between the distal variable SES and early intelligence level are not found. However, significant correlations are found between proximal variables, such as linguistic stimulation in the home or overstimulation of infants, and their performance on a Piaget-based scale of infant intelligence (*7*); between responsivity of the environment or stimulus variety of the home and infants' performance on the Bayley mental development scale (*8*); and between tactual stimulation of infants and their performance on the Bayley scale (*9*).

The above argument is not to be considered as denying the importance of genetic factors in the development of intelligence. The monozygotic-dizygotic comparisons reported by Wilson provide clear evidence for the operation of genetic factors in the process of intellectual development. However, before one can assign the overwhelming share of control to genetic factors, as Wilson proposes in this threshold model, one must first produce results showing that either (i) there are no differences in the specific environment of children who are discrepant in intelligence, or (ii) any differences that occur in the specific environment are unrelated to intellectual development. As of yet, evidence of this type is not available.

THEODORE D. WACHS
*Department of Psychological Sciences,*
*Purdue University,*
*Lafayette, Indiana 47907*

**References and Notes**

1. R. Wilson, *Science* **175**, 914 (1972).
2. B. Wellman, in *39th Yearbook of the National Society for the Study of Education* (Public School Publishing, Bloomington, Ill., 1940), pp. 1–21.
3. E. Pavenstedt, *Amer. J. Orthopsychiat.* **35**, 89 (1965).
4. S. Tulkin, *J. Pers. Soc. Psychol.* **9**, 31 (1968).
5. D. Falconer, *Introduction to Quantitative Genetics* (Ronald, New York, 1960), pp. 1–184.
6. B. Caldwell, *Merrill-Palmer Quart.* **16**, 260

(1970); M. Schoggen and P. Schoggen, *DARCEE Papers and Reports* (Demonstration and Research Center for Early Education, John F. Kennedy Center, George Peabody College, Nashville, Tenn., 1971), vol. 5, No. 2.
7. T. Wachs, I. Uzgiris, J. Hunt, *Merrill-Palmer Quart.* **17**, 283 (1971).
8. L. Yarrow, J. Rubenstein, F. Pedersen, paper presented at the 1971 meeting of the Society for Research in Child Development.
9. N. Solkoff, S. Yaffee, D. Weintraub, C. Blaise, *Develop. Psychol.* **1**, 765 (1969).
10. Supported by NIH research grant HD-04514. I thank F. Weizmann for his criticisms of a preliminary draft of this comment.

8 May 1972; revised 28 July 1972

McCall refers to the substantial size of the dizygotic (DZ) twin correlations for overall developmental level and profile contour, and he quotes the small portion from my report in which the relation between genetic influence and the size of the DZ twin correlations is discussed. Then he states, "The degree of heritability for a trait depends on the difference between the correlations for MZ [monozygotic] and DZ twins, not on the absolute size of the DZ correlations."

All the analyses in my report were explicitly performed to provide statistical tests of the differences between the correlations for MZ and DZ twins, as explained in the text and tables preceding the mention of the DZ correlations. The conclusions about genetic influence were based on the results of these tests, which showed significantly higher concordance for MZ twins, and the DZ correlations were interpreted in the context of these results. The analysis of the twin data was in no sense deficient of comparisons of MZ and DZ data.

McCall's introduction of the heritability ratio brings up a chronic source of confusion and misinterpretation in this research area. The ratio came originally from plant and animal breeding studies in which parentage was controlled and a planned mixture of genotypes was achieved through inbreeding or crossbreeding. When the appropriate assumptions were met, the ratio expressed the proportion of genetic variance in whatever phenotypic characteristic was being measured.

But as an expression of the genetic contribution to test-score variance on a psychological variable, the heritability ratio can be seriously misleading. There are several different formulas for the ratio which use different combinations of variance estimates and yield somewhat different results (*1*). Further, the same heritability ratio may be obtained from two sets of within-pair correlations which have quite different implications. For example, by McCall's method of doubling the difference between MZ and DZ correlations, a heritability estimate of .30 is obtained with $R_{MZ} = .90$ and $R_{DZ} = .75$; and it is also obtained for $R_{MZ} = .75$ and $R_{DZ} = .60$; for $R_{MZ} = .50$ and $R_{DZ} = .35$; and for $R_{MZ} = .20$ and $R_{DZ} = .05$. The sampling error of the difference gets larger as the correlations get smaller, so that while the first difference is significant, the next is marginal ($P = .08$), and the other differences are clearly nonsignificant. In fact, neither of the correlations in the last example is significantly different from zero. The genetic influence is hardly the same in each example, yet a heritability ratio obscures these essential differences and draws attention away from the central data, which are the within-pair correlations, their sampling error, and the test of significance for the difference between the correlations.

The limitations are serious enough that the heritability ratio should be permanently retired. Where the data warrant it, there are other more comprehensive models for estimating genetic and environmental variance components in psychological variables (*2, 3*).

A comparison of sibling pairs with DZ twins is a desirable method for evaluating the common environmental effects of "twinness" as such, if other factors such as age at testing and the tests administered are adequately controlled. The correlations reported by McCall for the Fels sibling pairs are so low, however, as to raise questions about the reliability of the scores, particularly if the tests were administered by different examiners over a long period. The Gesell does not lend itself to formal scoring as readily as the Bayley, and the reliability of the scoring method itself has been questioned (*4*).

An analysis based on intraclass correlations is preferable to an analysis of within-pair variances, because the intraclass correlation takes into account all variance between subjects, and it expresses the proportion of variance accounted for by pair membership. The statistic flows directly from the analysis-of-variance model, and it gives a measure of covariation for the twins, or the relative homogeneity of scores within pairs (*3, 5*). It also compensates for any sampling fluctuation in the distribution of scores within each zygosity group.

The within-pair variability, however, is based solely on the discrepancies within pairs, and it gives no indication of the similarity of scores within pairs. The *F*-tests proposed by McCall permit only the limited conclusion that the discrepancies are significantly larger for DZ pairs; they provide no measure (or test) of the concordance level for either group. While gene segregation should enhance the differences within DZ pairs, the effect will be limited by assortative mating; and in any event, a comparison of the two intraclass correlations will be more reliable and informative than a comparison of the two within-pair variances.

The objection to the univariate analysis-of-variance model is in error with respect to the presumed bias favoring within-pair similarity. Nonconstant covariances may affect the comparisons involving the repeated measures, in this case the various ages at which the test scores were obtained. But there is no test of ages as a main effect —the scores are standardized separately at each age, which abolishes the between-age variance. The analysis tests for the degree of homogeneity within pairs in the total score summed across ages, which is an unbiased test under any model; and it tests for the homogeneity of the score profile at different ages. The actual size of this within-pair correlation for score profile is not affected by the covariance structure between tests, nor is there an effect on the test of differences between MZ and DZ pairs. The only possible effect is on the probability that the within-pair correlation for score profile is different from zero. If the covariance structure is deviant enough, the actual $P$ value may (for example) be .025 instead of the nominal $P$ value of .01.

Where necessary, the bias can be corrected by an adjustment in the degrees of freedom, as noted originally by Box (*6*) and as confirmed by recent Monte Carlo studies (*7*). In the twin data, every within-pair correlation was significantly different from zero at $P < .001$ under the most extreme and conservative correction possible, so the results are not being inflated by a hidden bias.

Finally, McCall reports that he found no evidence of greater similarity in profile contour on the Gesell tests for sibling pairs than for unrelated pairs. The study is as yet unpublished so the results are not available for evaluation. McCall has, however, published a similar study of older siblings from the Fels sample in which he averaged adjacent-age scores to compensate for missing data (*9*). This method of moving averages obscures any age-to-age changes that should be reflected in the profile contour, and it eliminates the variance component that is supposedly being tested. Further, his statistical analysis dealt with within-pair differences in profile contour for sibs and unrelated pairs; as noted above, this does not furnish a measure of concordance for either group or a means of testing whether the concordance level is (i) greater than zero or (ii) significantly different for the two groups. Therefore, the methods and data of the unpublished study will need careful examination before the results can be interpreted.

On balance, none of the results or conclusions from the original report are changed by McCall's comments.

Wachs in his first paragraph substitutes the terms intellectual development, intelligence, and IQ for the original term infant mental development, which was specifically chosen to designate the capabilities being manifested during infancy. The point was made clearly that there are pronounced changes in developmental precocity from one age to the next, which reflect the changing capabilities being measured by the Bayley mental scale plus the idiosyncratic spurts and lags in development exhibited by each infant. The strength of the twin data is in showing that the changes in precocity occur concordantly for genetically related pairs. But there is no direct correspondence between precocity in infancy and later measures of intelligence, and Wachs' substitution of terms serves to confuse an essential distinction between the capabilities being measured at different ages.

As to Wachs' assertion that I ruled out the importance of environmental variance, my specific conclusion was that "the caretaking and stimulation needed to support infant mental development are sufficiently supplied by most home environments that fall above the level of impoverished. In all likelihood. however, there may be a cumulative latent influence absorbed from the home environment during infancy that combines with genetic predisposition and gradually becomes manifest as school age approaches; since the child's measured IQ becomes increasingly related to his parents' IQ, educational level, and socioeconomic status as he gets older" (*10*, p. 917).

Wachs objects to socioeconomic status (SES) as a measure of environmental variance because it does not control for the types of experience encountered. No assertion was made in my report that SES was a complete measure of environmental variance; it does, however, furnish a reliable method of designating differences between family environments, many aspects of which do relate immediately to the infant's experience. The problem with Wachs' argument is that while SES is a poor predictor of infant developmental status, it becomes an increasingly better predictor of intelligence in later childhood. If there is such a heterogeneous mixture of environments within each SES category, and if SES has so little relation to the types of experience which Wachs believes to be important for the development of intelligence, then where is this predictive power coming from? The fact is that as intelligence gradually stabilizes with age, the predictive power of SES improves, and one cannot dismiss SES as an unsuitable variable without acknowledging this relationship.

Wachs' reference to an expected maximum genetic correlation for DZ twins highlights another source of confusion in this area. There is a massive conceptual gap between .50 as the average proportion of shared genes for pairs of offspring drawn from randomly matched parents, and .50 as the correlation coefficient expressing the ratio of within-pair covariance to the total variance in a distribution of test scores. There are so many assumptions needed to bridge the gap that one cannot arbitrarily establish a maximum value expected for genetically related pairs, especially where assortative mating is involved. The values will ultimately be determined by large-scale, carefully controlled developmental studies including groups of differing genetic composition.

As to the distinction between specific, proximal experience as opposed to SES as a distal label or variable, the distinction is arbitrary in that any environmental variable must register as a proximal experience if it is to affect behavior. An SES category is a representation of a particular life-style, the events of which impinge directly on the child and furnish a primary dimension of his proximal experience. The proximal variables cited by Wachs appear in differential degree along the SES continuum, and while the exceptions are always a matter of interest, they do not abolish the relation between SES categories and the proximal experience dimension that Wachs favors.

It would be highly desirable to have a separate measurement scale of proximal experience to use in conjunction with SES ratings, and this would permit a more complete test of the relation between early proximal experience and school-age intelligence. The supporting results cited by Wachs are tentative and deal with changes in infant developmental status, whereas the more basic question concerns whether the effects of early experience will independently contribute to a sustained upward (or downward) shift in school-age IQ. The issue is not whether a supportive and appropriately stimulating environment should be supplied for each infant—it seems to me this answer must be affirmative under the most elementary considerations of human care—but whether the attained level of intelligence at school age is directly related to the range of variation found in family environments above the level of impoverished. This, plus carefully controlled follow-up studies of enrichment programs with impoverished infants, will identify the differential weights to be assigned to early experience in the development of intelligence.

In the interim, the original conclusion about infant mental development is reaffirmed: For the great majority of pairs, their life circumstances fall within the broad limits of sufficiency that permit the genetic blueprint to control the course of infant mental development.

RONALD S. WILSON
*Child Development Unit,*
*University of Louisville*
*School of Medicine,*
*Louisville, Kentucky 40202*

**References**

1. S. G. Vandenberg, *Psychol. Bull.* **66**, 327 (1966).
2. C. Burt, *Brit. J. Math. Statist. Psychol.* **24**, 1 (1971); R. B. Cattell, *Psychol. Rev.* **67**, 353 (19 J. L. Jinks and D. W. Fulker, *Psychol. Bull.* **73**, 311 (1970).
3. O. Kempthorne and R. H. Osborne, *Amer. J. Hum. Genet.* **13**, 320 (1961).
4. E. E. Werner, in *Sixth Mental Measurements Yearbook*, O. K. Buros, Ed. (Gryphon, Highland Park, N.J., 1965), pp. 807–809.
5. E. A. Haggard, *Intraclass Correlation and the Analysis of Variance* (Dryden, New York, 1958); D. S. Falconer, *Introduction to Quantitative Genetics* (Ronald, New York, 1960).
G. E. P. Box, *Ann. Math. Statist.* **25**, 484 (1954).
7. R. O. Collier, F. B. Baker, G. K. Mandeville, T. F. Hayes, *Psychometrika* **32**, 339 (1967).
8. D. W. Fulker, J. Wilcock, P. L. Broadhurst, *Behav. Genet.* **2**, 261 (1972).
9. R. B. McCall, *Science* **170**, 644 (1970).
R. S. Wilson, *ibid.* **175**, 914 (1972).

20 September 1972