Psychological Review
1994, Vol. 101, No. 3, 396–413

Copyright 1994 by the American Psychological Association, Inc.
0033-295X/94/$3.00

No More Than Skin Deep: Ethnic and Racial Similarity in Developmental Process

David C. Rowe, Alexander T. Vazsonyi, and Daniel J. Flannery

Many studies adduce evidence of ethnic or racial dissimilarities in developmental outcomes (e.g., delinquency and achievement). Many researchers fail to distinguish between group average levels and developmental processes (correlations). Evidence is reviewed that developmental processes are nearly identical for U.S. Black, Hispanic, White, and Asian ethnic and racial groups. Using diverse and representative data sources, covariance matrices were computed for these ethnic groups and then compared by using a LISREL goodness-of-fit test. Not only were these matrices nearly identical but they also were no less alike than covariance matrices computed from random halves within 1 ethnic or racial group. This article documents the importance of accepting ethnic and racial similarity of developmental processes. Thus, group average level differences may result from different levels of developmental antecedents working through common developmental pathways.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

In an integrated school in an American city, two boys with similar academic problems might be taking a standardized test. Both boys fidget at their seats and strike up conversations with classmates rather than pay attention to the teacher's recitation of test directions. Both boys read a year and one half behind grade level. They may have different nonacademic gifts, one being good at sports, the other at drawing. If both of these boys were ethnically White, one would look naturally to influences in their home environments to understand their academic slowness and hyperactivity. One would expect similar influences to be responsible for both boys' conditions—one might expect to find two lower class families lacking in financial resources and two single mothers who, in their stressful lives, were unable to encourage their children's schoolwork. If genetic influences play a role in hyperactivity (Goodman & Stevenson, 1989), one might expect that their biological fathers both had childhood histories of overactivity and inattentiveness.

However, suppose one boy were ethnically Black and the other ethnically White, or one boy were ethnically Hispanic and the other White. With this knowledge, in social science, one response has been to generate different developmental explanations for the boys' behavior (Helms, 1992; Ogbu, 1991; Spencer, 1990). The ethnic minority family may instill different social values. For instance, one group may emphasize the importance of school achievement, whereas the other may emphasize peer popularity; or the minority boy may have been teased and taunted by his majority classmates, to the detriment of his self-esteem and academic performance. The aim of this article is to explore these two contrasting explanations of behavioral variation in minority and majority ethnic and racial groups: common developmental processes versus minority-unique developmental processes.

The article's focus is on English-speaking ethnic and racial minorities living in America, mainly Blacks and Hispanics compared with Whites, although one study detailed later also included Asians. We acknowledge that ethnic and racial minorities are themselves heterogeneous in genetics and social histories. Through intermarriage, genetic isolation among racial and ethnic groups has been relaxed; for instance, about 25% of the autosomal genes in American Blacks originate from intermixing with the White population (Chakraborty, Kamboh, Nwankwo, & Ferrell, 1992). Great heterogeneity also exists within the "White" majority culture. Although people of English background predominate, American Caucasians may trace family histories to France, Germany, the Netherlands, Scotland, or to other places. Once in America, these diverse racial and ethnic groups adopted English as their first language. Their unique cultural backgrounds absorbed changes from, and exerted influence on, a broader American culture.

Nevertheless, we believe that racial and ethnic classifications carry important information about peoples' genetic and cultural backgrounds and, as such, that they can be useful in scientific studies. First, these classifications carry information about evolutionary phylogenies (Stringer & Andrews, 1988). For instance, American Blacks originated on the African continent and so they are racially Negro, a group distinguishable from other races on the basis of heritable, physical characteristics. American Whites mainly originated in Western Europe and would be racially Caucasian. Asians came from China, Japan, and Southeast Asia and from a Mongoloid racial grouping distinctive from both Caucasians and Negroes. Hispanics, although racially Caucasian, mainly originated in Spain and Portugal and first resided in Mexico, the Caribbean, or South America before moving to the United States. Although the exact number of existing racial groupings depends on the particular classification system used, it is indisputable that generations of geographic separation have given rise to racial subdivisions, traceable by different biological lineages. Second, these classi-

---

David C. Rowe, Alexander T. Vazsonyi, and Daniel J. Flannery, Division of Family Studies, University of Arizona.

We thank Joseph L. Rodgers, Mark Roosa, George Knight, Stephan A. Cernkovich, Nancy Darling, Larry Steinberg, Janine Goldman-Pach, and Travis Hirschi for providing us with the covariance matrices reanalyzed for this article.

Correspondence concerning this article should be addressed to David C. Rowe, School of Family and Consumer Resources, University of Arizona, Tucson, Arizona 85721.

fications carry information about cultural backgrounds as well. For example, Anglo culture is distinctive from that which arose in the Hispanic Americas, and African culture is distinctive from European and Hispanic ones. Thus, race or ethnicity qualifies as a nominal variable, correlated with individuals' genetic and cultural histories. For instance, the dichotomous Black/White variable would correlate with extent of African genetic heritage strongly, but imperfectly. In social science, many approximately valid measures are used to capture sources of behavioral variation, and racial and ethnic classifications are no exception. Most studies use peoples' self-classifications of race or ethnicity as the basis for groupings; we would therefore accept individuals' self-classifications as a basis for their racial or ethnic group membership.

The subtle issue of the conceptual relation between developmental process and group average levels has been discussed previously in the developmental literature (McCall, Appelbaum, & Hogarty, 1973). Nonetheless, confusion often remains between these two concepts, which are at times genuinely difficult to distinguish. We define a *group average level* as a group's statistical mean on either independent variables (e.g., influences) or dependent variables (e.g., outcomes). In contrast, we refer to *developmental processes* as the biological or psychosocial mechanisms relating the independent variables (influences) to the dependent ones (outcomes). It is sometimes thought that, if Ethnic Group A has a higher mean level on an outcome than Ethnic Group B, then some different developmental process is necessarily implied but, of course, this is not so. Identical developmental processes may occur in both groups, but they may also differ in average levels on developmental antecedents (influences), so that they consequently differ in the outcome as well.

For example, blood pressure readings have been reported as higher in Black than in White men. Blood pressure is, to some degree, sensitive to dietary salt intake. Suppose the causal influence of salt on blood pressure is represented as a regression coefficient of .17. Assuming that the physiological processes linking salt to blood pressure were the same in Blacks and in Whites, then this coefficient would be .17 in both populations. Nonetheless, if Black men culturally had diets higher in salt than did White men, then this common developmental process could produce a part of the *average racial difference* in blood pressure. Suppose dietary salt levels were measured in a combined population of an equal number of Black and White men. One may discover that Black men had a mean dietary intake of .6 standard deviation units above the combined group mean ($X = 0$ for a standardized variate), whereas White men had −.6 units below it. Then the mean blood pressures of Black and White men should differ by about .2 standard deviation units (that is, $.6 \times .17 - (-.6) \times .17$). In this hypothetical blood pressure example, a common developmental process partly explains both individual differences and the average differences between racial groups.[1]

Alternatively, developmental processes could differ from one racial or ethnic group to another. In the blood pressure example, suppose, for instance, that Black men were affected by dietary salt but White men were unaffected, at least for typical exposures. In this case, the regression of blood pressure on salt intake would be .17 in Black men but .00 in White men. White men and Black men could have identical levels of dietary salt intake, but because salt intake raises blood pressure in Black men because of their greater salt sensitivity, just Black men would develop higher average blood pressure levels.

Neither example was intended to be a true analysis of salt metabolism and its relation to blood pressure; rather, they were presented here as illustrations of a subtle distinction between two contrasting alternatives: (a) group average differences due to common processes but different antecedent conditions, and (b) group average differences due to dissimilar biological or psychosocial processes.

In the next section, some theoretical arguments for the different developmental processes among ethnic and racial groups are considered. They are followed by a separate section making arguments for common developmental processes in all ethnic and racial groups.

Argument for Ethnic and Racial Group Differences

Americans belonging to different ethnic and racial groups may possess separate cultural histories that lead to different mechanisms of socialization within different groups (Harrison, Wilson, Pine, Chan, & Buriel, 1990; Steinberg, Dornbusch, & Brown, 1992). One hypothesis is that because Ethnic Group A has been socialized differently from Ethnic Group B, a difference in developmental outcome may occur as well.

Helms (1992) advanced such a cultural explanation for the lower IQ test performance of Blacks than Whites. She argued that the culture of American Blacks, despite its geographic and temporal separation from African roots, maintains significant elements of African culture through its traditions. According to Helms, at least eight cultural elements are more widespread among American Black culture than among American White culture. They include (a) spirituality, the "greater validity of the power of immaterial forces in everyday life over linear, factual thinking"; (b) movement: "personal conduct is organized through movement"; and (c) social time: "time is measured by socially meaningful events and customs" (p. 1096). Helms further connected these cultural values with difficulty on standardized tests because they either interfere with the acquisition of test-relevant material or with test-taking motivation. Thus, a preference for social time may interfere with the rigors of a timed IQ test, in which "time is a valuable commodity." Helms summarized Heath's (1989) argument that a gulf in thought and emotion separates Blacks from other ethnic groups:

> Black Americans are socialized in Black communities to develop spontaneous, creative, interactive, and expansive thinking skills. Consequently, upon reaching testable age, it is difficult for them to reconcile the contrasting socially oriented worldview of their communities with the ascetic Eurocentric view that presumably underlies test construction . . . . (p. 1097)

A second hypothesis about group differences is the direct psychological effects of discrimination. On one hand, minority groups may socialize children differently because they possess positive cultural traditions, independent of those in the majority culture; on the other hand, they may develop values that are

[1] Salt intake is not actually the explanation for racial differences in blood pressure. These differences persist after controlling statistically for dietary salt; our example was hypothetical.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

destructive in the long run but that represent a means of adapting to social discrimination.

This latter explanation features strongly in Ogbu's (1991) explanation of "involuntary" minorities' poor school achievement. He defined as involuntary those minorities brought into another society through slavery, conquest, or colonization. Because their incorporation into another society was imposed, the involuntary minority, according to Ogbu, develops a psychological identity in opposition to that of the mainstream culture. Members of an involuntary minority perceive institutionalized discrimination against them and therefore they lose faith in their ability to compete successfully in institutions such as schools. Although these minority parents may verbally value schooling as a means of social advancement, their lack of personal academic effort and success may undermine their exhortations. In these groups, young people may channel their energies away from conventional routes to success and into survival strategies of petty crime and other oppositional behavior. In their generally poor schools, they also see evidence of discrimination—that they are given less just because they are minorities. In overview, minority-unique experience with social discrimination may lead to their (relative) failure in schooling and later in conventional economic competition.

Unlike some explanations for group differences, though, Ogbu accepted the existence of individual differences, noting that "I do not claim all . . . involuntary minorities are academically unsuccessful" (1991, p. 29). In this context, he mentioned several strategies for making good academic progress while not offending peers' values contrary to achievement, such as using sports success as a "cover" for genuine academic achievement.

Ogbu's views, and others like his advocating deep cultural differences, have been widely influential in social science. In a recent issue of *Child Development* devoted to minority children, a variety of "cultural difference" explanations were advanced (for instance, Coll, 1990; Harrison et al., 1990; McLoyd, 1990; Slaughter-Defoe, Nakagawa, Takanishi, & Johnson, 1990). A sampling of these explanations includes (a) value on mother–infant enmeshment versus autonomy, (b) a culturally rooted value on power assertion among American Blacks deriving from traditional African values, (c) minority groups' emphasis on cooperative rather than on competitive views of life, and (d) traditional Asian values such as respect for authority and group cooperation. As shown in these different examples, many social scientists now argue for different developmental processes as a source of group average differences.

In summary, although the emphasis has been on differences between American Whites and Blacks, the general idea that ethnic groups may have different cultural values, given their preexisting traditions, and their conflicts with a majority culture, forms a basis for expecting group differences in causal developmental processes. In opposition to ethnic and racial difference are those arguments favoring similarity in developmental processes, as discussed next.

Argument for Ethnic and Racial Group Similarities

In social science, there is a great awareness of human universals (Brown, 1991; Buss, 1989; Russell, 1991). Two examples of cultural diversity, color names and facial expressiveness, on more recent empirical examination, show universals. Around the world, people parse the color spectrum into similar components and name them, and they can read and express the same emotions using nearly identical configurations of facial expression. The list of cultural universals extends into all behavioral domains—specific mate preferences, love of kin, preferential altruism directed toward kin, play, deceit, enduring mateships, and many more. Social scientists involved in studying psychological traits among cultures also adopt a *universalistic* perspective: that characteristics of different cultures result from organism–culture interactions and so are constrained and not purely arbitrary (Berry, Poortinga, Segall, & Dasen, 1992).

Although genetic arguments are sometimes marshaled as an explanation of ethnic average differences (Jensen, 1969; Rushton, 1988), a strong justification for universal developmental processes is also provided. About 25% of genetic loci are *polymorphic* (i.e., their genes exist in multiple forms such as the genes A, O, and B in the ABO blood group). Genetic explanations for group differences focus on those polymorphic loci at which genotype frequencies may differ among ethnic groups. In most polymorphic genetic loci, however, genetic variation is more prominent within than between ethnic and racial classifications (Lewontin, 1982). Moreover, by current estimates, about 75% of genetic loci contain no genetic variability at all across humans, genetically, we are much more alike than different.[2] At these nonvariable loci, the genes in Asians, Blacks, and Whites would be exactly identical—this shared genetic heritage makes all people belong to one human species, with many common characteristics. This great degree of genetic similarity leads to the expectation that people of different ethnic groups will be alike in many complex behavioral adaptations because the genes that form them tend to be shared by all ethnic groups. Of course, this conclusion only applies to traits with a partly genetic basis but, as argued by Berry et al. (1992), many traits may result from organism–environment interactions.

A common American culture also encourages the expectation of similar developmental processes. Although Americans of different ethnic backgrounds have some unique heritage, they also may share in common many aspects of American culture; for example, second-generation American Hispanics are probably more familiar with MTV, McDonald's, classroom schooling, presidential inaugurations, and the National Collegiate Athletic Association basketball playoff tournament than they are with Mexican games or Peruvian oral traditions. Indeed, these aspects of American culture are pervasive, reaching most Hispanics, Asians, Blacks, and the majority of Whites as well. Thus, an argument against the ethnic "difference" view may combine these two observations: (a) that biological processes in most humans would be necessarily alike, and (b) that in America they may interact with many common cultural features. Consequently, developmental processes leading to outcomes for different American ethnic groups would be highly similar.

It is clear that the arguments on the difference versus sim-

[2] The genetic difference between humans and chimpanzees amounts to changes in about 1% of total DNA. Thus, it is possible for a relatively small number of genetic changes to produce striking morphological and psychological differences between species. Nonetheless, all humans would be substantially more closely related genetically to one another than to chimpanzees or other primate "out-groups."

ilarity sides are both strong ones, deserving our serious consideration. As social scientists, we want to produce empirical data bearing on which of these contrasting alternatives best describes the development of behavioral traits.

Empirical Methods of Evaluating Developmental Processes

If we had some marvelous technology that would make developmental process visible, we could just aim this machine at children from different ethnic groups and make a direct comparison of the developmental processes leading to particular traits. Lacking such a wonderful device, we must resort to indirect strategies that examine the strength of different developmental processes by using statistical associations (e.g., correlations and covariances).

One way of testing for process differences is to compare the covariances or unstandardized path coefficients between variables that are influences and those that are developmental outcomes across ethnic groups. In general, if these statistics differ significantly from one ethnic group to another, and if the magnitude of this difference is appreciable, one may conclude (in the absence of statistical artifacts) that developmental processes also differ between groups. In contrast, if the statistics were the same for different ethnic groups, one may accept that developmental processes were the same (given a large sample so statistical power would be available to detect group differences).

Comparing Covariance Matrices

In this article, our general analytic strategy was to compare covariance matrices computed on different racial and ethnic groups. It is more efficient to test the statistical similarity of total covariance matrices rather than to investigate specific causal models. For any matrix, many different models would be possible, offering specific ones would be open to the criticism that alternative models were not considered, and we cannot be sure which model is the correct one. Comparing the statistical similarity of matrices is a model-free statistical test. If the matrices were statistically identical, however, the correct analytic procedure would be to pool them before investigating specific causal models; of course, the same causal model must apply to each racial/ethnic group. In later analyses, each matrix is about 10 × 10. They contain both widely accepted influence variables and major developmental outcomes. The influence variables may include family functioning and peer relationship variables; and the developmental outcomes may include academic achievement, conduct problems or delinquency, and depression. The research question is as follows: Are the statistical associations found in these covariance matrices the same or different in different ethnic groups?

In answering this question, several weak statistical procedures must be avoided. Merely counting the number of statistically significant correlations (i.e., greater than zero) would be a poor procedure. If one group gave a correlation of .20 ($p < .05$) and another one of .17 ($p > .05$), they might be seen as different; however, both associations may be statistically significant in large samples that provide greater statistical power. The correct statistical approach is one of testing the statistical significance of the difference between two correlation coefficients. Given the number of correlations in a 10 × 10 matrix, however, many pairwise comparisons can easily lead to Type I statistical errors (i.e., inferring nonexistent relationships). Instead of piecemeal procedures, we adopt a strategy of testing for significant differences between entire covariance matrices by using linear structural equation modeling (LISREL). Although many social science applications use correlation matrices, we follow a more stringent test of comparing covariance matrices. In the latter comparison method, the equivalence of variances is tested simultaneously with that of covariances. In addition, in group comparisons, statistical estimates of standard errors and model fit lie on firmer mathematical grounds when covariance matrices are used (Cudeck, 1989).

The analytic approach and general argument may be illustrated with a hypothetical example. Table 1 presents covariance matrices (2 × 2) for majority and minority individuals separately. The variable X_1 is a developmental influence (e.g., parental involvement); the variable Y is a developmental outcome (e.g., children's educational aspirations). The matrices suggest different developmental processes: The $X_1 - Y$ association is stronger for the minority group (where the r calculated from the covariance matrix equals .62) than for the majority group (where $r = .30$). The X_1 and Y variables also possess greater variance in the minority group (5 vs. 9.40, and 5 vs. 9.18, respectively).

What could be happening? Figure 1 shows the structural model that generated the majority and minority matrices. In the minority group, a variable that is not directly observed (X_2) exerts an influence on both X_1 and Y variation. This variable would be some factor unique to minority group members. It might be variation in exposure to racial taunts and insults, or it might be variation in individuals' reactions to discriminatory social practices (e.g., red-lined housing or job discrimination). Note that this variable must vary within the minority group; correlational analysis, of course, cannot detect an influence that is exactly constant for all group members. Whatever its exact source, it influences variation in (to use the earlier choices) both parental involvement and children's own educational aspirations. According to Wright's rules for reading a structural diagram, the correlation of X_1 and Y has the following mathematical expectation in the majority group: $r = a$; and in the minority group, $r = a + bc$, where a, b, and c are the path coefficients, as shown in Figure 1. The existence of X_2 also induces greater variance in X_1 and Y because it is causally linked to them. No-

Table 1
Hypothetical Matrices for Majority and Minority Groups

Group	X_1	Y
Majority		
X_1	5.00	
Y	1.50	5.00
Minority		
X_1	9.40	
Y	5.71	9.18

Note. $a = .3$, $b = .9$, and $c = .35$. In majority group, variance of X_1 = 5, residual on Y = 4.55. In minority group, residual on X_1 = 2.11, variance of X_2 = 9.0, and residual on Y = 5.50. Matrices were computed from a basic language program, major.bas.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

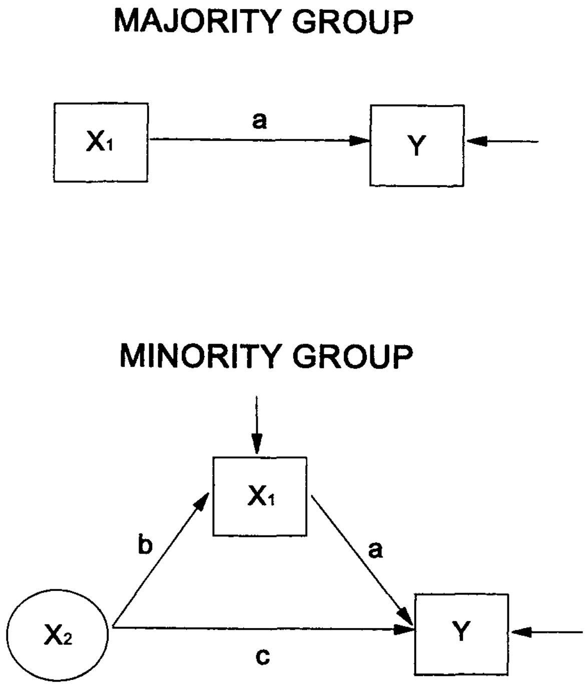

Figure 1. Path models for minority and majority groups. (Standardized path coefficients: $a = .3$, $b = .9$, and $c = .35$; Y = developmental outcome; X_1 = developmental influence; X_2 = minority-unique developmental influence.)

tice, however, that these diagrams say nothing about the groups' average levels on X_1 or Y. The analysis of developmental processes may be done independently of variables' average levels. The two perspectives may be conjoined if ethnic or racial differences in average levels were found on antecedent variables (X_1, X_2) and developmental outcomes (Y), but the focus in this article is on developmental process per se.

Table 2 gives results from a LISREL comparison of the covariance matrices in Table 1, with a hypothetical sample size of 100 majority persons and 100 minority persons. LISREL generates a covariance matrix that is the average of the two covariance matrices (when sample sizes are equal). Table 2 gives this fitted matrix, which is the average of the two covariance matrices in Table 1. The more the matrices deviate from this fitted matrix the worse is the statistical fit.

Table 2
LISREL Model Fit to Majority and Minority Matrices

Fitted matrix	X_1	Y
X_1	7.20	
Y	3.61	7.09

$\chi^2(3, N = 200) = 16.9$, $p < .001$; GFI, majority group = .88; GFI, minority group = .95; $\chi^2/N = .085$

Note. Fitted using LISREL 7, with total $N = 200$: 100 individuals in majority group, 100 in minority group. GFI = goodness-of-fit index. Model line for first group, mo nx = 2 nk = 2 lx = id td = ze; model line for second group, mo ph = in.

Assessing Goodness of Fit

In structural equation models, several indexes of goodness of fit are available. The chi-square test takes a value of zero if the fitted and observed matrices were exactly equal and a value greater than zero if they were unequal. If the chi-square is statistically significant (on the basis of the degrees of freedom for the comparison), then the hypothesis that the two matrices were equal would be rejected statistically against one that they were unequal. However, the sensitivity of the chi-square test depends directly on sample sizes, and in large samples ($Ns > 200$) it is almost impossible not to obtain a nonsignificant value of the chi-square (Green, 1992; Tanaka, 1987). For this reason, alternative indexes of fit have been developed (Green, 1992; Loehlin, 1992a). The LISREL program (Jöreskog & Sörbom, 1988) provides a goodness-of-fit index (GFI) constrained to fall between 0 and 1 that assesses the match of the expected and observed covariance matrices. GFI values greater than .90 are usually considered to provide a good fit to the observed covariance matrix (Green, 1992).[3] Another way to eliminate the effect of sample size is to calculate the chi-square divided by the sample size. Although this index lacks a general interpretation, it is useful when matrices of the same dimensionality are being compared. Because of its generality, we emphasize the GFI as the best measure of goodness of fit.

In the hypothetical example, the chi-square showed that the groups differed statistically, $\chi^2(3, N = 200) = 16.9$, $p < .001$. The chi-square per observation equalled .085. Moreover, at least in one group, the GFI was unsatisfactory (GFI = .88 in the majority group, .95 in the minority; discrepancies between larger dimensioned matrices would produce lower GFIs). The statistical differences between the minority and majority group, of course, were expected in this example because it was constructed to produce group differences.

In later statistical analyses, one additional comparison is used. The sensitive chi-square fit index as derived from different ethnic and racial groups is put against chi-squares derived from random halves of a single ethnic group. If different developmental processes exist between ethnic and racial groups, then greater statistical strain would result from the comparison of different ethnic groups than from arbitrary halves of a single ethnic or racial group.

Data Sources

In this section, we describe the principal data sources used in our analyses. We report results for six data sources. One data source, the National Longitudinal Survey of Youth (NLSY), was used twice: once when treating the surveyed persons as individuals, and a second time when matching individuals as sibling pairs from the sampled households. The results from sibling pairs are discussed separately in a later section.

[3] The accuracy of the GFI is not entirely independent of sample size, but according to Marsh, Balla, and McDonald (1988), it "performed better than any other stand-alone index (p. 396)" in a study of the influence of sample size on fit indexes. In many comparisons, we have samples sufficiently large as to avoid biases introduced by using small samples.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Data Sources Analyzed for This Study

Space limitations prohibit a detailed discussion of the reliability and validity of variables used in the following studies. However, most variables had been carefully chosen as representing accurately different developmental constructs. For example, the Home Observation for Measurement of the Environment (HOME) is one of the most widely used measures of family environment (Caldwell & Bradley, 1984). The Peabody Picture Vocabulary Test is a high-quality, nationally standardized instrument (Dunn & Dunn, 1981). An extensive literature exists in criminology showing the reliability and validity of self-report delinquency variables (Hindelang, Hirschi, & Weiss, 1981). The Youth Self-Report is a widely used self-assessment of behavior and social competence that possesses national norms on different ages and ethnic groups (Achenbach, 1991). Although no set of variables is perfect, for the most part the ones selected in the following studies satisfy current guidelines for behavioral assessment.

Tucson Substance Use Study. Table 3 gives the characteristics of each source and lists the variables with which covariance matrices were computed. The first source, the Tucson Substance Use Study, was collected by us in Tucson, Arizona. Children in grades 6–7 were recruited through a Tucson-area school district. All students attending school on a particular day were sampled. A passive consent procedure was used at most schools, so that students were only omitted from the sample if their parents had requested so in writing. Students completed the survey questionnaires during homeroom sessions under the supervision of study staff. Of the original available number of 1,437 students, 67 students (4.7%) did not participate due to parental concerns, and 134 students were absent during data collection (9.3%). The approximate participation rate was about 86%. An additional 66 surveys (4.6%) were not usable due to incomplete or missing data, defined as completion rates below 50% of all items.

The final sample had 1,022 Caucasian and Hispanic students (for an overall completion rate of 81.4%). It was about equally divided between males and females and between 6th and 7th graders. Their mean age was about 13 years at the time of data collection. About 28% of the sample were living with a single parent. Variables classified as developmental outcomes were drug use, aggression, depression, grades, impulsivity, school adjustment, and self-efficacy. Variables classified as antecedents were peer pressure, friends' drug use, parental involvement, and parental monitoring (see descriptions in Flannery, Vazsonyi, Torquati, & Fridrich, 1994).

National Longitudinal Survey of Youth (Individuals). In 1979, the original National Longitudinal Survey of the Work Experience of Youth used standard stratified probability sampling methods to locate households representative of the American population. Blacks, poor Whites, and Hispanics were oversampled to gain more detailed information on them. The original sample consisted of 14- to 21-year-olds. In 1986 and 1988, children born to women in the NLSY study were brought into the study (in 1986, ages 1–14; Baker & Mott, 1989). In this report, we focus on these NLSY children who, because of the original sample composition, were disproportionately minority and poor. In the NLSY children, the Black mothers had a slightly higher average educational level than the White or Hispanic parents. The mothers were about the same age (in their late 20s).

The NLSY is a rich data set, with many possible structures for analysis. In the individual data set, we decided to focus on middle childhood (ages 6–9), an age range not represented in our other data sources. The 10 variables used are listed in Table 3 (for detailed description, see Baker & Mott, 1989). One noteworthy measure is an abbreviated version of the HOME that was used to index the quality of the home environment (Caldwell & Bradley, 1984). This version included both interviewer observations and maternal self-reports. Children's achievement was assessed with the Peabody Individual Achievement Test (PIAT) reading, mathematics, and comprehension tests. Other measures included appear in Table 3.

National Longitudinal Survey of Youth (Siblings). As many mothers in the original NLSY sample had more than a single child, it was possible to construct biologically related sibling pairs. Details on the construction of sibling pairs are given in Rodgers, Rowe, and Li (1994). The rarer twin, cousin, and half-sibling pairs, also identified from the NLSY children, were not used in this article. In this analysis, the matrix was restricted to variables for which missing data were few in number in the 1988 NLSY data collection. To maximize sample size, siblings were permitted to be from 6 to 18 years old. The age-adjusted HOME total score was used as an index of the quality of family environment. Unlike the individual matrix, which included age and was based on a limited age range, we chose to use developmental outcomes already adjusted for age and sex differences. Problem behaviors were assessed by the age- and sex-adjusted problem behavior total score; academic achievement was assessed by the mean of the PIAT reading, math, and comprehension measures, all age-normed variables. It is important that the HOME measure was not identical for siblings because some items were child specific. Thus, a sibling matrix contained six variables: Sibling A's achievement, problem behavior, and HOME variables; and the same variables repeated for Sibling B.

Covariance Matrices Provided to This Study

California/Wisconsin Study. Students attending nine high schools in California and Wisconsin each completed two self-report questionnaires (Steinberg, Mounts, Lamborn, & Dornbusch, 1991). These schools were selected to provide a diverse sample in terms of family structure, socioeconomic status, and type of community (rural, suburban, and urban). Of the approximately 10,000 students, 9% of the students were Black, 14% were Asian, 12% were Hispanic, and 60% were White (the remainder belonged to other ethnic groups). Questionnaires were completed for about 80% of the target sample.

Bowling Green Study. The Bowling Green Study provides data on both Blacks and Whites. To avoid biases associated with school-based studies, the Bowling Green Study probabilistically sampled geographic areas in Toledo, Ohio, to locate youths between 12 and 19 years of age. Geographic stratification was based on the 1980 area census. Within strata, households were selected to identify eligible respondents, who were interviewed in their homes. A total of 942 face-to-face interviews were successfully completed. About half of the sample was female, and half was male; about the same relative division applied to race, with 45% of the sample White, 50% Black.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Data Sources for Covariance Matrices

Matrix size and variables	Sample	Source
	Tucson Substance Use Study	
Matrix 11 × 11	Tucson, Arizona	Flannery, Vazsonyi, Torquati, & Fridrich (1994)
Peer pressure	Hispanic, N = 278	
Friends' drug use	White, N = 744	
Parental involvement	Mean age = 12.7 years	
Parental monitoring	Males and females	
Self-efficacy		
Academic adjustment		
Lifetime drug use		
YSR Aggression		
YSR Depression		
YSR Impulsivity		
Grades		
	National Longitudinal Survey of Youth (Individuals)	
Matrix 10 × 10	Nationwide	Baker & Mott (1989)
Mother's education	Black, N = 549	
Age of child	Hispanic, N = 335	
HOME Cognition	White, N = 836	
HOME Emotion	Males and females	
School self-esteem	6–9 years	
Self-worth		
Math achievement		
Reading recognition		
Reading comprehension		
Problem behavior (total)		
	National Longitudinal Survey of Youth (Siblings)	
Matrix 6 × 6	Nationwide	Baker & Mott (1989)
Achievement	Black, N pair = 156	
Problem behavior	Hispanic, N pair = 128	
HOME	White, N pair = 319	
Variables repeated for each sibling	Males and females	
	6–18 years	
	Wisconsin/California Study	
Matrix 10 × 10	Wisconsin/California	Steinberg, Dornbusch, & Brown (1992)
Academic engagement 1987	Black, N = 635	
Academic engagement 1988	White, N = 3,943	
Behavioral control	Asian, N = 906	
Psychological autonomy—granting	Hispanic, N = 827	
Involvement—warmth	Males and females	
Parents' school involvement	Grades 9–12	
Parents' school encouragement 1987		
Parents' school encouragement 1988		
GPA 1987		
GPA 1988		
	Bowling Green Study	
Matrix 8 × 8	Toledo, Ohio	Cernkovich & Giordano (1992)
Parental communication	Black, N = 469	
School involvement	White, N = 409	
Attachment to teachers	Males and females	
School commitment	12–19 years	
School involvement		
Risk of arrest		
Perceived opportunity		
Delinquency		

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3 (*continued*)

Matrix size and variables	Sample	Source
	Richmond Youth Project	
Matrix 10 × 10	Richmond, Virginia	Hirschi (1969)
Participation with father	Black, N = 1,427	
Supervision by father	White, N = 1,872	
Overall GPA	Males and females	
Mother's education	Junior and senior high school	
Participation with mother		
Supervision by mother		
Peer orientation		
IQ		
Standard self-report delinquency		
Official offenses to 1967		
	Prevention Study	
Matrix 9 × 9	Southwestern city	Roosa, Tein, Groppenbacher, Michaels, & Dumka (1993)
CRPBI Acceptance	Hispanic, N = 70	
CRPBI Rejection	White1, N = 70	
CRPBI Inconsistent Discipline	White2, N = 70	
CRPBI Control	Mothers and 8- to 14-year-old children	
CRPBI Hostile Control		
Open family communication		
Problems in family communication		
Mean value on Kovacs Conduct Disorder Index		
CBCL Conduct Disorder Subscale		

Note. YSR = Youth Self-Report; HOME = Home Observation for Measurement of the Environment (Caldwell & Bradley, 1984); GPA = grade point average; CRPBI = Children's Reports of Parental Behavior Inventory; CBCL = Child Behavior Check List. Children's achievement in the National Longitudinal Survey was assessed with the Peabody Individual Achievement Test reading, mathematics, and comprehension tests.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Richmond Youth Project. This study of delinquency recruited from a population of 17,500 students entering 11 public junior and senior high schools in the Richmond, Virginia, area (Hirschi, 1969). A stratified sampling procedure yielded 4,077 students, about 45% of whom were Black. The sampling procedure yielded a group diverse in levels of family income and education.

Prevention Study. This study was part of an evaluation program that recruited children in 4th, 5th, and 6th grades into a prevention program for the children of alcoholics (Knight, Tein, Shell, & Roosa, 1992; Roosa, Tein, Groppenbacher, Michaels, & Dumka, 1993). Families were recruited in a southwestern city by using a variety of procedures (e.g., newsletters distributed to students and taken home, telephone recruitment, and door-to-door canvassing). One hundred thirty-four one-parent families and 169 two-parent families were interviewed. Family incomes ranged from less than $5,000 per year (12%) to more than $40,000 per year (14%) with the modal income range being from $5,001 to $10,000 (17%). Most families could be described as lower to lower middle class. The ethnic distribution was 60% White, 20% Hispanic, 13% Black, and 6% other. Because of their relatively larger sample sizes, we requested covariance matrices from this research group only for Whites (N = 170) and Hispanics (N = 70). On the basis of a diagnostic interview, the Prevention Study researchers found that 36% of the mothers were either problem drinkers or alcoholic, and 57% of fathers had similar diagnoses.

As shown in Table 3, extensive data were obtained on family functioning (obtained from the mothers). In addition to these parenting-style variables, two developmental outcome variables focused on childhood conduct problems. The Prevention Study sent us three covariance matrices. One was computed for the 70 Hispanics; the others were based on two groups of 70 random White families, with the stipulation that no White family appeared in both random subgroups.

Results for Covariance Matrices (Individuals)

For all covariance matrices, the general result was one of striking and consistent similarity between ethnic and racial groups. Although space does not permit a detailed description of covariance matrices, it is worthwhile to give a few details of the matrices obtained from several sources.[4] In Whites, greatest correlation in the Tucson Substance Use matrix was between academic adjustment and substance use (r = −.51). Substance use also correlated highly with aggression, friends' drug use, and parental monitoring (absolute rs = .34–.51). School grades were associated with greater academic adjustment and less susceptibility to peer pressure (rs = .38 and −.28, respectively). Self-efficacy found few correlates among the 10 remaining variables (maximum r = −.33, with depression). Overall, the matrix presents a rich set of associations for testing models of etiologic influences on substance use.

Table 4 presents the GFIs comparing these 11 × 11 covari-

[4] Except for Roosa's prevention study, the covariance matrices used here may be obtained by writing to David C. Rowe.

Table 4
Tucson Substance Use Study

Measure	Hispanic (N = 278) vs. White (N = 744)	Hispanic (N = 278) vs. White (N = 278)	Hispanic1 (N = 139) vs. Hispanic2 (N = 139)	White1 (N = 372) vs. White2 (N = 372)	Misspecified
χ^2	144.5	129.8	87.0	122.0	2,236.9
GFI (first)	.94	.96	.95	.97	.67
GFI (second)	.99	.96	.95	.97	.67
χ^2/N	.14	.23	.31	.16	4.0

Note. GFI = goodness-of-fit index. Degrees of freedom = 66. All chi-squares were statistically significant ($p < .05$).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ance matrices. Whites and Hispanics were compared twice: once with the entire samples (column 1), another time with equal sample sizes (column 2). According to the GFI, the White and Hispanic matrices were equivalent (GFIs > .90). When data sets contain unequal *N*s, the LISREL program makes the common, estimated matrix closer to the Whites' matrix that was based on the larger sample size (N = 744). When a random White sample equally large as the Hispanic sample was drawn (N = 278), then the two matrices received equal weight in the fitting process, giving closer GFIs (.96).

Despite the close, quantitative similarity of these covariance matrices, the chi-square tests yielded statistical rejections of their equality ($p < .05$). As noted earlier, in structural equation modeling it is understood that this chi-square is exquisitely sensitive to slight differences between model-expected and observed covariances (Tanaka, 1987). For this reason, we adopted other yardsticks for comparing the adequacy of model fits: the GFI, the chi-square per observation, and the statistical fit of halves of one ethnic or racial group. By randomly assigning individuals of one ethnic group to one of two equivalent subgroups, within-ethnicity/racial group covariance matrices were computed. If ethnic/racial groups differ, then the GFI, chi-square, and chi-square per observation statistics for different ethnic and racial groups should greatly exceed those on random halves of a single group.

As shown in Table 4, a comparison of random, within-ethnicity subgroups also yielded significant chi-squares (Hispanic1 vs. Hispanic2, 87.0; White1 vs. White2, 122.0, $p < .05$). These statistical rejections may reflect imperfections in social science data: The 11 variables fail to satisfy strict multivariate normality (e.g., drug use is positively skewed) and they lack exact interval scaling. Given these distributional and measurement inadequacies, any two covariance matrices may be statistically nonequivalent, as compared with two covariance matrices drawn from a true multivariate normal population. Nonetheless, on the basis of these measurement imperfections, it would be improper to postulate different causal models for the random halves of one ethnicity. Here, a better guide to matrix similarity was clearly the high GFI values, not statistically significant chi-squares (.96 for Hispanics vs. Whites in column 2; .95 for Hispanic1 vs. Hispanic2; .97 for White1 vs. White2). As shown by the GFI values, random halves of Whites or Hispanics were as statistically similar as Whites versus Hispanics. In a similar vein, the per-observation chi-square from Hispanic1 versus Hispanic2 (χ^2/N = .31) was actually greater than those from the two cross-ethnicity comparisons. In the comparisons just described, no evidence existed for differential causal processes operating within the Hispanic versus White groups.

Under any circumstances, could the statistical similarity of any two Tucson Study matrices be rejected? Given the many statistically significant variances and covariances within them, the answer would seem to be in the affirmative. To demonstrate this statistically, an analysis was conducted by comparing the Hispanic1 versus Hispanic2 groups, except that now the Hispanic2 matrix was purposely computed incorrectly, with variables entered into it in the reverse order. These two Hispanic matrices must satisfy an assumption of grossly different causal processes; hence, they should be statistically unequal. As shown in Table 4, the chi-square test in this comparison was 2,236, with GFIs equal to .67. There was no question that these two matrices fit one another poorly, so covariances and variances as found in the Tucson Study, if organized differently, can be shown to be unequal.

In the NLSY matrices, the three academic achievement level variables correlated highly (> .68). As these were raw score variables, they also correlated with age (about .65). The HOME cognition variable had statistically significant relationships with achievement and problem behavior (mean *r*s = .19 and −.20, respectively).

As shown in Table 5, the results from the NLSY were substantively similar to those from the Tucson Substance Use Study, but this time in a nationally representative study with three ethnic groups and an overrepresentation of economically disadvantaged families. As the ethnic groups were unequal in size, we drew random samples of 335 each from the White and Hispanic groups and used an equal number of Blacks. The tabled results show comparisons for these equal-sized groups: The GFI ranged from .94 to .96, all very high values. Furthermore, the cross-ethnic comparisons of Whites versus Blacks, Blacks versus Hispanics, and Hispanics versus Whites were no more different than two random halves of the Whites (GFI = .96). Other analyses, not shown here, used all individuals when comparing ethnic groups; no evidence was found for developmental differences in them either, but of course the goodness of fits were always better in the numerically larger groups. These findings again confirm the great similarity of covariance structure within and between different ethnicities.

In Table 6, Wisconsin/California data comparisons of four ethnic groups—Blacks, Whites, Hispanics, and Asians—are presented. It is evident that all pairwise comparisons produced excellent fits. The GFI equalled or exceeded .93 (mean GFI for all 12 matrices = .97). Although .93 is a very good fit, it tends

Table 5
National Longitudinal Survey of Youth

Measure	Hispanic (N = 335) vs. Black (N = 335)	Hispanic (N = 335) vs. White (N = 335)	Black (N = 335) vs. White (N = 335)	White1 (N = 418) vs. White2 (N = 418)
χ^2	140.7	209.7	171.0	193.3
GFI (first)	.96	.94	.96	.96
GFI (second)	.96	.94	.95	.96
χ^2/N	.21	.31	.26	.29

Note. GFI = goodness-of-fit index. Degrees of freedom = 55. All chi-squares were statistically significant ($p < .05$).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

to understate the degree of similarity because the matrices producing this value involved comparisons of very dissimilar group sizes: nearly 4,000 Whites versus a smaller number of either Hispanics or Blacks. Notice, too, that even over an apparent cultural distance of Blacks and Asians, the fit to a common covariance matrix remained an excellent one. The GFIs were .97 and .99 for Blacks and Asians, respectively, and the "stress" of chi-square per observation was only .12. Unfortunately, we lack comparisons for random halves of these ethnic groups; however, given the previous results and the close similarity of these matrices, little reason exists to believe that they could show any greater degree of identity of covariance pattern than found in the cross-ethnic comparisons here.

Table 7 presents the remaining data sets. The fits of Bowling Green's and Richmond Youth Project's Black and White matrices, both computed from studies of self-reported delinquency, were excellent. The Prevention Study had one of the smallest samples, and it was unique in that it was a clinical sample (about 60% of families had alcohol abuse problems). The Hispanics' matrix fit the White1 random half, χ^2 (45, N = 140) = 69.0, about as well as the two random White halves fit one another, χ^2 (45, N = 140) = 66.0. The goodness of fits were lower than what we have seen in other samples, but this difference may reflect greater sampling variation due to smaller samples and the presence of parental psychopathology, which may have influenced developmental processes (N = 70 per group). In summary, the analyses of provided data in Tables 6 and 7 reinforced what we discovered in the analyses of the Tucson and nationally based NLSY samples: a tremendous similarity of covariance structure within and between ethnic groups.

In the previous within-ethnic group comparisons, one random half of an ethnic/racial group was compared with another. Of course, many random subdivisions of any population are possible. To explore variability in the *distributions* of chi-square values, we ran a "bootstrapping" study for the NLSY Blacks (N = 549) and NLSY Whites (N = 836). One hundred random samples were drawn with replacement first from the Whites (each N = 125). These samples were used to form 50 comparisons. For each random White–White comparison, LISREL estimated a chi-square under the assumption of equality of covariance matrices. Fifty random samples (N = 125) were also drawn from the Black group and another set of 50 random samples (N = 125) from the White group. They were placed into 50 pairs of White–Black covariance matrices and, for each pair, chi-square values were estimated by LISREL.

Figure 2 presents the histograms of chi-square values for the comparisons of Whites versus Whites and Whites versus Blacks. The distributions were nearly identical, approximately normal in shape, and showed equal mean values (Whites–Whites, M = 67.4; Whites–Blacks, M = 66.8, SD = 16.6 and 14.7, respectively). Although either mean chi-square would be a statistical rejection of matrix equality, this rejection occurred when Whites' random samples were compared with one another as well as when Blacks' samples were compared with Whites'. In summary, repeated comparisons made on 200 random samples confirmed that covariance matrices across racial groups were not more dissimilar than covariance matrices *within* a racial group.

Results for Covariance Matrices (Siblings)

The NLSY data provide a unique opportunity to compare familial covariance patterns across ethnic and racial groups. These data play a somewhat different role from the individual matrices in the determination of similarity of developmental processes. In the individual matrices, a large number of measured variables were pro-

Table 6
Wisconsin/California Study

Measure	Black (N = 635) vs. White (N = 3,943)	Black (N = 635) vs. Asian (N = 906)	Black (N = 635) vs. Hispanic (N = 827)	Asian (N = 906) vs. White (N = 3,943)	Hispanic (N = 827) vs. White (N = 3,943)	Asian (N = 906) vs. Hispanic (N = 635)
χ^2	302.8	221.1	171.6	176.7	378.8	209.1
GFI (first)	.93	.96	.97	.97	.93	.97
GFI (second)	1.0	.98	.99	1.0	1.0	.98
χ^2/N	.07	.14	.12	.04	.08	.14

Note. GFI = goodness-of-fit index. Degrees of freedom = 55. All chi-squares were statistically significant ($p < .05$).

Table 7
The Bowling Green Study, Richmond Youth Project, and Prevention Study

	Bowling Green Study	Richmond Youth Project	Prevention Study		
Measure	Black (N = 409) vs. White (N = 469)	Black (N = 1,427) vs. White (N = 1,872)	Hispanic (N = 70) vs. White1 (N = 70)	Hispanic (N = 70) vs. White2 (N = 70)	White1 (N = 70) vs. White2 (N = 70)
χ^2	.50[a]	557.8	87.8	69.0	66.0
GFI (first)	1.0	.96	.91	.93	.91
GFI (second)	.90	.97	.85	.87	.92
χ^2/N	.00	.17	.63	.50	.47

Note. GFI = goodness-of-fit index. Bowling Green Study, df = 36; Richmond Youth Project, df = 55; Prevention Study, df = 45.
[a] Chi-square was nonsignificant.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

vided that may assess familial process, for instance, parental monitoring and intellectual stimulation. These measured variables, however, captured only that variation in family developmental processes directly associated with those particular variables. In behavior genetic studies, correlations computed on pairs of biological or social relatives are used to capture variation associated with different theoretical processes.

As shown in Figure 3, the correlation of Sibling A and Sibling B on a trait phenotype can be apportioned between two variance components as follows: (a) shared variation that makes family members alike in a trait phenotype (but different from persons in another family) and (b) unshared variation that operates uniquely on each individual. The latter component would make family members dissimilar in a trait phenotype. The sibling correlation represents the influence of all shared variables tied to the family unit that are a source of behavioral resemblance among siblings. If this correlation were .30, then 30% of variation in the trait phenotype is attributable to shared variation and 70% to unshared variation.

Notice that in this apportionment the correlation coefficient was not squared to yield variance explained. By using the rules of path analysis, one can see that the variance in Sibling A's trait phenotype (or B's) is just the path coefficient, z squared. Yet the correlation of Sibling A's with Sibling B's trait phenotypes also has the mathematical expectation of z^2; hence, the trait phenotype's variance explained by shared developmental processes is also the sibling correlation. In other words, the issue is not how well one can predict one sibling's trait score from the other's but, instead, how well all familial variation relates to trait variation. The remainder of phenotypic variation is attributable to unshared influences.

The advantage, then, of sibling correlations is that they detect the maximum influence of a particular type. In research inspired by a family study, shared familial variation may be assigned further to specific genes shared by siblings or to shared environmental processes, such as books available, parental surveillance of children, and so on. Before these steps, however, familial covariance matrices can be used to ask, "Is shared family variation the same in different ethnic groups?" With additional groups of relatives, the behavioral variation within ethnic groups may be further apportioned among genetic variation, shared environmental variation, and nonshared environmental variation (see Osborne, 1980).

To estimate familial effects, 6 × 6 covariance matrices were computed for NLSY Blacks, Hispanics, and Whites by using the variables listed in Table 3. All data were double entered.[5] The matrices yielded sibling correlations for HOME environment, problem behavior, and academic achievement, respectively. Besides these sibling correlations, the matrices included the ordinary intercorrelations (on individuals) among the HOME, problem behavior, and achievement as well as sibling cross-correlations, that is, the correlation of variable 1 in Sibling A with variable 2 in Sibling B (e.g., Sibling A's achievement with Sibling B's problem behavior). For the ethnic groups, pairwise comparisons were conducted (Black vs. White, Black vs. Hispanic, and Hispanic vs. White). Pairwise comparisons were also constructed from random halves of the White and Black sibling groups.

Table 8 presents the LISREL fits of an estimated covariance matrix. These fits were all excellent, with goodness-of-fit values equal to or exceeding .95. The Whites' and Blacks' random halves fit one another with near perfection (GFI = .99). The White matrix (N = 319) fitted the estimated intergroup covariance matrices equally well when compared with either the Hispanic or Black matrices (GFI = .99). In all comparisons except White versus Black, the chi-squares (per observation) were only slightly worse than the Black 1 versus Black 2 random halves. Overall these results continue to support a conclusion of racial and ethnic similarity because in all groups, the goodness-of-fit values for the compared covariance matrices were excellent.

To present the familial covariance matrix, we did another analysis fitting Hispanics, Whites, and Blacks to a single estimated matrix in a three-group LISREL solution. In this analysis, the chi-square was 60.7 (df = 42), and the respective goodness-of-fit values were .96 for Hispanics, .95 for Blacks, and .98 for Whites. Table 9 presents this LISREL-estimated, common covariance matrix in the more readable form of a correlation matrix. Substantial sibling correlations were found for all three variables (achievement, .41; problem behavior, .49; and HOME, .76). The HOME was weakly associated with better achievement and fewer problem behaviors (rs = .29 and −.24,

[5] In double-entry data, the number of observations is twice the number of sibling pairs; it equals the number of individuals. The first $N/2$ cases in a variable consist of Sibling A; the second $N/2$ cases consist of Sibling B. Entered in this way, a sibling correlation closely approximates an analysis of variance intraclass correlation. Double-entry correlations are used routinely in behavior genetic studies.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

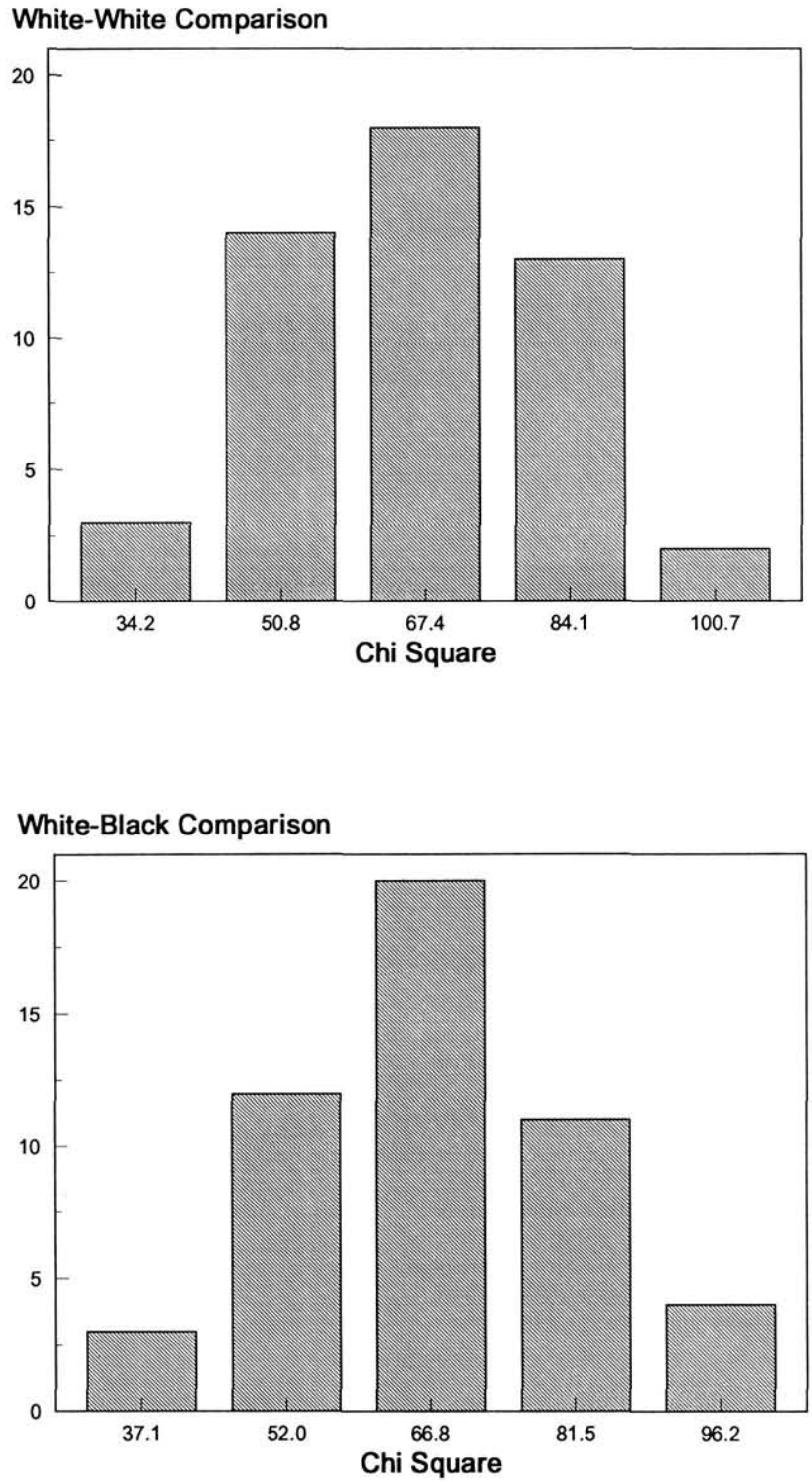

Figure 2. Monte Carlo comparisons of within-group versus between-groups covariance matrices.

respectively). The two developmental outcomes were only weakly related to one another (in individuals, −.17; within siblings, −.09). Thus, although both achievement and problem behavior were strongly familial, their causes were distinct. Because the GFIs were all very good, this pooled matrix would be the appropriate one for fitting more specific developmental models.

Discussion

Our main result was that developmental processes in different ethnic and racial groups were statistically indistinguishable. *Developmental process* refers to the association among variables in these groups and to the variables' total variances. This conclu-

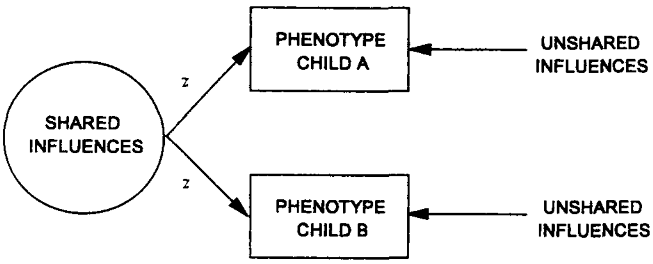

Figure 3. Path model for shared influences on siblings. z = path coefficient.

sion held for the examination of six data sources, containing a total of 3,392 Blacks, 1,766 Hispanics, and 8,582 Whites, and in one data source, 906 Asians. The patterns of covariances and variances were essentially equal when one ethnic or racial group was compared with another; moreover, this structural similarity between ethnic or racial groups was no less than that within random halves of a single ethnic or racial group. In the NLSY data source, Blacks, Whites, and Hispanics were sampled to be nationally representative but to overrepresent economically poor families; thus, these findings apply more broadly than just to middle-class populations. In a similar vein, the measures used were varied, with some matrices based entirely on adolescents' self-reports; others were based on parental reports or direct observations of the family, or both. Few studies, however, used direct observational assessments of behavioral outcomes. Outcomes included academic achievement, conduct disorders, delinquency, depression, IQ, and academic adjustment.

As discussed earlier, a test of a 10 × 10 covariance matrix not only rules out differences in measured variables but also differences in unmeasured (minority-unique) variables that causally relate to either antecedent variables or developmental outcomes. Nonetheless, our findings have failed to confirm this widespread belief: If these (unmeasured) influences existed, then covariance matrices on measured variables that resulted from them would be unequal across ethnic and racial groups. That is, a statistical model specified for one group would be misspecified for another. In a related approach, unmeasured familial influences would be expected to change within sibling-pair correlations from one ethnic or racial group to another. Again, in sibling data for large samples of different ethnic or racial groups, we found that Hispanics, Whites, and Blacks were alike in familial influences on achievement and problem behavior. Process similarity, of course, does not mean that ethnic or racial group average levels would be the same for either antecedent variables or developmental outcomes.

The Cultural Bias Argument and IQ

In overview, our position has been that the causal processes leading to individual differences in developmental outcomes may be similar across American ethnic and racial groups. Comparable issues arose in an earlier controversy that centered on the equivalence of intelligence test scores for American Blacks and Whites. The critics of IQ tests argued that they were culturally biased against Blacks (and other minorities). In its strong form, the cultural bias argument assumed a different causation of IQ within Black and White populations. In a weaker form, this argument was no more than pointing out that Blacks and Whites may experience intellectual environments that were differentially stimulating.

In its strong form, the cultural bias argument is another "difference" argument, of the type weakened by the data reviewed in this article. The argument postulated that in Blacks, IQ test scores would have the ordinary causes present in Whites as well as causes unique to Black populations (e.g., a distinct Black dialect that interfered with the acquisition of standard English). In all these hypotheses, some culturally unique influence would act to suppress obtained IQ test scores. With these mutually canceling influences, an IQ score of 90 in a Black child might underestimate his or her true intellectual ability—an ability that would have shown itself in the absence of the culturally unique influences.

Social scientists began to doubt this cultural bias explanation as evidence accumulated showing that IQ scores had the same network of correlates in Blacks and Whites (Jensen, 1980). If an IQ score of 90 actually underestimated a Black child's intellectual ability (at least over the short term), then this child would be able to show a greater ability to learn academic material than a White child with the same tested IQ. By comparing the regression lines of IQ score on later academic achievement (e.g., first-year college grades), computed separately for Blacks and Whites, researchers discovered little support for this expectation: Children with IQs of 90 got approximately the same grades (or other nonacademic outcomes), regardless of their racial groupings. In a special issue of the *American Psychologist* on IQ testing, Cole (1981) observed:

> [f]rom a large number of educational and employment studies, the

Table 8
National Longitudinal Survey of Youth (Siblings)

Measure	Hispanic (N = 128) vs. Black (N = 156)	Hispanic (N = 128) vs. White (N = 319)	Black (N = 156) vs. White (N = 319)	White1 (N = 160) vs. White2 (N = 160)	Black1 (N = 78) vs. Black2 (N = 78)
χ^2	18.5[a]	26.9[a]	41.3	12.6[a]	7.3[a]
GFI (first)	.96	.96	.95	.99	.99
GFI (second)	.98	.99	.99	.99	.99
χ^2/N	.07	.06	.09	.04	.05

Note. GFI = goodness-of-fit index. Degrees of freedom = 21. Sample size is the average number of sibling pairs per matrix.
[a] Chi-squares were statistically nonsignificant ($p > .05$).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 9
Pooled Sibling Correlation Matrix

Measure	1	2	3	4	5	6
1. Achievement 1	—					
2. Problem Behavior 1	−.17	—				
3. Home Quality 1	.29	−.24	—			
4. Achievement 2	**.41**	−.09	.27	—		
5. Problem Behavior 2	.09	**.49**	−.19	−.17	—	
6. Home Quality 2	.27	−.19	**.76**	.29	−.24	—
SD	11.8	14.4	14.7	11.8	14.4	14.7

Note. This is a double-entry sibling correlation matrix. Sibling correlations are shown in bold. Home quality was measured by Home Observation for Measurement of the Environment (HOME; Caldwell & Bradley, 1984). Home scores scaled as HOME = 1/10 × HOME.

> most common conclusion has been that many tests predict various educational and employment performances about as well for minority groups (Blacks and women being by far the most frequently studied minority groups) as for majority groups. (p. 1070)

In the same journal volume, Reschly (1981) seconded this conclusion:

> Conventional tests are nearly always found to be largely unbiased on the basis of technical criteria—for example, internal psychometric properties, factor structure, item content, atmosphere effects, and predictive validity. (p. 1098)

Or, as Berry et al. (1992) concluded,

> It appears that earlier views that sought to put the blame for unequal test scores primarily on the tests have lost much of their momentum. . . . [A] serious concern about cultural bias has become and will remain an inherent aspect of assessment, just as there is continuing concern for validity and the establishment of norms. . . . It is now generally recognized that within a society intergroup differences in test scores often are a reflection of a real state of affairs. (p. 313)

When IQ tests do show a slight bias, it lies in the opposite direction to that predicted by cultural bias theory—a single regression line slightly overpredicting Blacks' academic performance. In summary, the emerging consensus among testing experts is that cultural bias in the IQ tests themselves is no more than a minor source of group IQ difference (Barrett & Depinet, 1991; Snyderman & Rothman, 1987).

With regard to our analytic strategy, investigating the predictive validity of an IQ test would be equivalent to comparing 2 × 2 covariance matrices (e.g., variable 1 = IQ test score, variable 2 = chemistry grades) computed for the majority and minority groups separately. We know, from this research literature on "bias" in IQ testing, that these matrices must be statistically similar, as their equality would be a prerequisite to that of regression lines.

Our analysis has been more demanding, however. We have included both independent variables and developmental outcomes in larger matrices than the 2 × 2 IQ matrix, permitting greater possibility for violations of equivalence. In addition, the defenders of IQ tests merely argued that "IQ" was the same construct in majority and minority groups—that it could do equally well the task of predicting academic or job performance, regardless of individuals' ethnic identities. We make here a stronger claim about causal process for the variables examined in this article—that all the influences giving rise to individual differences in developmental outcomes are essentially equivalent in majority and minority groups. Although only one of our matrices directly included an IQ outcome, several others had proxy measures for it, such as grades or standardized achievement tests. Moreover, in all these comparisons, including the earlier sibling analysis of latent family influences, causal process appeared to be similar across ethnic and racial lines. Thus, our claim would be that IQ, as with its related variables, is not only the same construct in different ethnic or racial groups but that it would also possess identical developmental determinants in different racial or ethnic groups.

In this article, we have not tested for equivalence in means. In the IQ situation, minority populations tend to have lower mean IQ scores and lower academic performance scores than the majority population. Although the source of average, level differences between ethnic and racial groups has not been emphasized thus far in the article, it is clearly essential to integrate mean levels with our understanding of the etiology of individual differences in any complete theory of behavioral development.

Two views of ethnic and racial mean differences may be defended: (a) that they arise from different causal processes than individual differences and (b) that they arise from different antecedent levels in a common causal process. In the next two sections, both alternatives are discussed, and we give several reasons for favoring the latter.

Different Causal Processes

It is conceivable that the causal processes leading to *average* levels would be different from those creating within-group variation in behavior. This possibility is real in the mathematical sense in that *averages* and *correlations* are statistically independent. However, for this alternative to hold requires also that minority-unique Factor X contributes to average level but does not contribute to variation among individuals.

For example, consider that Down's syndrome children are mentally retarded because they inherit a chromosomal abnormality: three copies of chromosome 21. There are clearly large mean differences between normal children (IQ range 70–130) and Down's syndrome ones (IQ range 25–70). Consider, however, that other familial influences may be similar for the two groups. If mothers' IQs (range: 70–130) were plotted on the *X* axis, and children's IQs on the *Y*, the parent–child regression on IQ could be .50 for both groups of children. In this example, the Down's children's mean IQ level would be affected by a different developmental process from that influencing their variation around that mean. That is, the latter variation might have similar familial causes in both Down's and normal children.

However, this argument—influences on means separate from those on individual differences—is a strong one because it requires nearly equal exposure to and influence of the unique causal mechanism in all exposed persons. Although this may make sense for a chromosomal abnormality with devastating developmental consequences, it is more difficult to imagine that *psychosocial* processes affect all persons within a given group equally. For example, Ogbu (1987) acknowledged that not all Black Americans would experience racial discrimination in the same way:

> Of course, not everyone feels this way. Some black Americans do not identify with the oppositional identity and cultural frame of reference; some do so only marginally. (p. 165)

If minority-unique Factor X contributed to both group averages and within-group variation, then its influence should have been apparent in our earlier analyses (at least for the developmental outcomes that were chosen). The greater the number of measured variables in our matrices, the less likely some "ghost" process exists as a unique developmental mechanism in any one minority group. Indeed, the example just given—of a chromosomal abnormality—is a poor one for arguing that ethnic and racial differences exist: Down's syndrome would affect all racial and ethnic groups in the same way.

Different Levels on Common Antecedents

The other explanation of group averages is that they result from different average levels on antecedents in the pan-ethnicity, common developmental pathways. Although large average level differences in common antecedents may work through causal pathways to become smaller, there still would be socially and statistically significant differences in the developmental outcomes. From this viewpoint, a focus on "averages" or on within-group "individual variation" would be misleading; they are simply different summaries of the total variation. Turkheimer (1991) noted that, insofar as IQ variation is concerned, any developmental process that has been postulated to influence group averages should also influence within-group variation, and vice versa:

> Although the two-realms [group and individual] hypothesis is now the received view of nature and nurture . . . it is implausible to suggest that the forces shaping the IQs of groups are different from those shaping the IQs of individuals; environmental and genetic factors can affect only individuals, one at a time. . . . *There are two realms of variance, between and within groups; there is only one realm of development.* (pp. 393–394, italics added)

The "one-realm" model, with its universalistic assumptions, possesses great theoretical parsimony. It permits generalization of findings from one ethnic or racial group to another, and it eliminates hypotheses that refer to minority-unique variables. One must wonder, therefore, why so little theorizing in social science has followed this route. In the next section, we consider this issue and its possible remedies.

Genetic Differences and the IQ Issue

In the main, opposition to a "one developmental process" theory arose from the IQ controversy, namely, from the observation of lower average IQs in American Blacks than Whites. This average IQ difference, with its socially meaningful implications for securing higher education and more prestigious jobs, demanded some explanation from social scientists. One explanatory route: common process, different levels on the developmental antecedents, was acceptable so long as those antecedents were environmental (e.g., social class or home environmental quality) but not if they were genetic. The problem for social scientists has been that 50%–70% of the within-ethnic/racial group variation owes to a genetic etiology (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Plomin, DeFries, & McClearn, 1990). The common developmental pathway model merely asserts that the cause of group average differences lies in common antecedents, environmental or genetic, or both.[6] However, with scientific evidence of genetic influences on individual differences in IQ increasingly secure, the common process model automatically raised for social scientists the issue of racial or ethnic differences in genetic alleles related to IQ variation. According to Shepard (cited in Cole, 1981),

> One reason that bias in mental testing is so volatile an issue is that it involves the specter of biological determinism, i.e., whether there is a large difference in intelligence (IQ) between Black and White Americans which can be attributed largely to inherited differences. (p. 1067)

One way for social scientists to avoid the implication of genetic variation in ethnic and racial *average* differences was to postulate different causal processes for IQ in minority versus majority groups in the United States. Like a walk down a garden path, the way at first looks good, with hereditarian hypotheses about ethnic and racial differences safely silenced.

However, a pretty garden path can lead into brambles of brush and scrub oak; it may carry considerable theoretical cost. The assumption of difference meant that developmental findings from one group could not be generalized to another. It provokes a search for difference through the few statistically significant correlations among the greater number computed on minority and majority groups. Furthermore, these significant correlations are often given complex, psychosocial explanations—but with what prospect for independent replication? It also may reinforce group stereotypes, leading to a neglect of the considerable overlap between ethnic and racial groups in behavioral traits. It allows one developmental outcome—tested IQ scores—to drive thinking about many other outcomes, many of which correlate only weakly with IQ variation. In our

[6] The antecedent variables in this article are usually presumed to represent environmental influences because they are labeled as environmental variables (e.g., social class and parental monitoring). However, genetic variation has been found in measures labeled environmental when they are treated as phenotypes in behavior genetic analyses (Plomin & Bergeman, 1991). Thus, there is a further reason to hypothesize genetic influences on average ethnic and racial differences.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

view, all these liabilities have been accumulating at considerable cost to social science research.

It is not the aim of this article to settle the nature/nurture controversy for average racial IQ differences; the data do not permit a resolution that would be convincing to most social scientists (Loehlin, Lindzey, & Spuhler, 1975; Mackenzie, 1984). Nevertheless, we do want to encourage social scientists to think more in terms of common developmental pathways and less in terms of "difference."

Research Designs

One mechanism for thinking about common developmental processes is to consider offspring from interracial (ethnic) marriages. In these marriages, which are increasingly common in America, children would be genetically admixed. In addition, through contact with biological relatives on both sides, they may be culturally admixed as well. These children form a linking *bridge* between different social groups, which are not so culturally or genetically distant as widespread social stereotypes maintain. Indeed, they offer the opportunity for a unique and powerful research design for investigating the commonality of genetic and family environmental influence.

Consider, for example, the kinships pictured in Figure 4. This research design can demonstrate commonality of (familial) developmental processes. All families would be sampled through a racially or ethnically intermarried couple, who have an interracial (ethnic) child (cf. Family 2). Two additional families would be ascertained through the interracial couple and through each parent's brother or sister. Thus, Family 1 would be a minority family with at least one child. The other would be a majority (White) family with at least one child (cf. Family 3).

In these families, there are covariance matrices for parent–child, child; uncle/aunt, child; spouse of uncle/aunt, and cousins. If (familial) developmental processes were identical for all children, regardless of racial group, then covariance matrices computed from the starting point of a majority child in Family 3, an interracial child in Family 2, and a minority child in Family 1, should be identical. Genetic influences would be indicated by a child's correlation to an uncle or aunt greater than to an uncle or aunt's spouse. Family environmental influences would be indicated by equal correlations to uncle/aunt versus their spouses (e.g., the child's resemblance to them is on the basis of similar social class levels in the adult siblings' families). Specific models, allowing for spousal resemblance and for direct measures of family environment, could be constructed for the covariance matrices available in the research design of Figure 4 (Neale & Cardon, 1992).

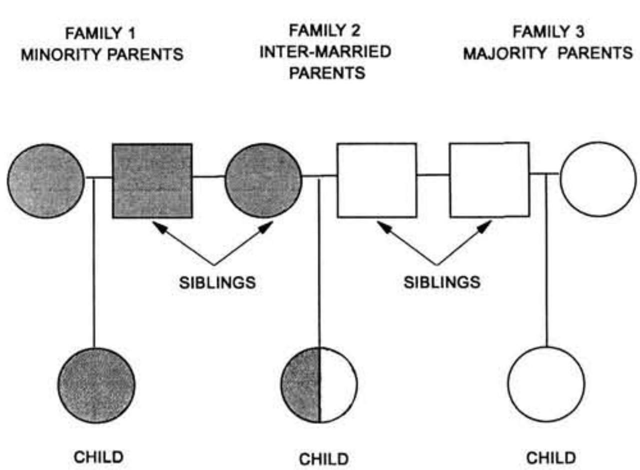

Figure 4. Family pedigree from interracial child proband.

In this research design, it is also possible to investigate the origin of racial and ethnic mean differences. If average levels as well as individual differences were genetically influenced, then the interracial child should have an average (trait) score midway between that of his or her cousins on both sides. If no genetic influence exists, then all cousins would have equal means. However, if sampling of particular families were unrepresentative, it may result in offsprings' average differences, which would then simply reflect these sampling biases. Therefore, a better test of the "mean difference" hypothesis would be to compare the intercepts of the regression of child on mid-parent scores. The intercept has the following mathematical expectation: a = child mean − b(midparent mean), where a is the intercept, b is the pooled (if developmental processes are common) regression coefficient, the mid-parent mean is the average of the mother's and father's trait scores, and child mean is the offsprings' average trait score. For a genetically influenced trait that has a higher majority than minority mean, the intercepts should order as follows: Minority Family 1 < Intermarried Family 2 < Majority Family 3.[7]

Of course, other research designs have been proposed to investigate whether racial and ethnic average differences in traits possess a partly genetic basis, including transracial adoption studies (Weinberg, Scarr, & Waldman, 1992) and nonfamilial genetic admixture studies (Mackenzie, 1984). In these research designs, the focus is on racial differences—rather than on a *commonality* of developmental processes. Nonetheless, they are methods, each with strengths and limitations, for evaluating empirically hypotheses about a genetic basis to ethnic and racial mean differences. Avoidance of this issue should not allow social science to take an enticing "garden path" of assuming development differs among racial and ethnic groups. Over the long run, in our opinion, it will be better to settle this issue with well-designed empirical studies. Perhaps research in this volatile area would move forward if researchers would heed Loehlin's (1992b) advice for behavioral genetic studies of racial differences in IQ: (a) Say clearly what your results mean and what

[7] As this research design is used for purposes of illustration, we do not detail its possible strengths and weaknesses. To mention a few issues, if a variable correlates strongly with the decision to marry interracially (ethnically), then its variance might be restricted in these families. Comparisons of interracial (ethnic) parents and others from their respective populations could reveal the degree of selection bias, and various statistical approaches may be adopted to deal with it. For other variables, self-selection may not be a problem. For the test of genetic mean levels, it would be necessary that the minority parent have little majority biological percentage in his or her own parents or grandparents. Finally, this design would require solving practical problems of recruitment of the proband families and the families of the parents' siblings. A design involving three families presents difficult practical problems because cooperation is needed from siblings of the proband family.

they do *not* mean, (b) put matters in a quantitative perspective, and (c) be tactful.

Applications

In work on ethnic and racial groups, the results of this article can suggest a few "DON'TS." The "DON'TS" all relate to statistical pitfalls in the comparison of groups:

1. When a covariance is significantly greater than zero in one group but not in another group, DON'T automatically interpret this finding as a group difference.
2. When separate multiple regression equations are computed for the majority and minority groups, DON'T attribute a group difference to different unstandardized regression weights or to different orders of extracted variables.

Although both procedures are common, they are flawed because they fail to show that a difference between majority and minority groups is statistically significant. The statistically correct procedures are either to (a) test for the significance of the difference of two unstandardized regression coefficients or (b) test Race × Variable interactions, in addition to main effects, in multiple regression equations using unstandardized variables.

Although the latter represent proper statistical tools, we believe that they, as well, may be too liberal. With neither prior hypotheses nor independent replications, we believe that researchers should accept the null hypothesis of no group differences until such time as scientifically acceptable evidence for differences is forthcoming. Furthermore, this argument extends beyond racial and ethnic group differences that have been our primary example. In studies of males and females, or of families with dual and single earners, or of children with and without day-care experiences, many of the same problems would obtain. As in the case of racial and ethnic group differences, these other group comparisons are often done without prior hypotheses about different processes, without replications across different samples, and without concern for whether correlation coefficients are actually statistically different from one another (nor do many studies compare the variance–covariance matrices, which is the more stringent and proper test). The techniques illustrated in this article offer a methodological approach that should be extended generally for the study of group differences with respect to developmental processes.

In conclusion, in the realm of IQ/achievement and social adjustment, we found that developmental processes were not specific to any racial or ethnic group. Our finding, of course, does not exclude group-unique developmental processes in other developmental outcomes not covered by our data matrices. Nonetheless, we expect that the results shown for the groups studied in this article, which were ethnically diverse (Hispanic, White, Black, and Asian) and which were also diverse in social class origin, geographic location, variables sampled, and identity of target respondents, will generalize widely: *that developmental processes are indeed invariant across U.S. racial and ethnic groups.*

Substantively, these findings imply that researchers should seek the determinants of average level differences between ethnic and racial groups in average levels of antecedent variables that act through common developmental pathways. Researchers should also be encouraged: Results they obtain for one ethnic group or in one U.S. geographic location will probably generalize to other groups and locations. Powerful generalization is the hallmark of a successful scientific enterprise; it bodes well for the future success of social science that developmental processes are alike in many subgroups of *homo sapiens.*

References

Achenbach, T. M. (1991). *Manual for the Youth Self-Report and Profile.* Burlington: University of Vermont, Department of Psychiatry.

Baker, P. C., & Mott, F. L. (1989). *NLSY child handbook 1989.* Columbus: Ohio State University, Center for Human Resource Research.

Barrett, G. V., & Depinet, R. L. (1991). A reconsideration of testing for competence rather than for intelligence. *American Psychologist, 46,* 1012–1024.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications.* New York: Cambridge University Press.

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science, 250,* 223–228.

Brown, D. E. (1991). *Human universals.* Philadelphia, PA: Temple University Press.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12,* 1–49.

Caldwell, B. M., & Bradley, R. H. (1984). *Home observation for the measurement of the environment.* Little Rock: University of Arkansas.

Cernkovich, S. A., & Giordano, P. C. (1992). School bonding, race, and delinquency. *Criminology, 30,* 261–291.

Chakraborty, R., Kamboh, M., Nwankwo, M., & Ferrell, R. (1992). Caucasian genes in American Blacks: New data. *American Journal of Human Genetics, 50,* 145–155.

Cole, N. S. (1981). Bias in testing. *American Psychologist, 36,* 1067–1077.

Coll, C. T. G. (1990). Developmental outcome of minority infants: A process-oriented look into our beginnings. *Child Development, 61,* 270–289.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317–327.

Dunn, L. M., & Dunn, F. C. (1981). *PPVT—Revised manual.* Circle Pines, MN: American Guidance Service.

Flannery, D. J., Vazsonyi, A. T., Torquati, J., & Fridrich, A. H. (1994). Ethnic and gender differences in risk for early adolescent substance use. *Journal of Youth and Adolescence, 23,* 1–19.

Goodman, R., & Stevenson, J. (1989). A twin study of hyperactivity: II. The aetiological role of genes, family relationships and perinatal adversity. *Journal of Child Psychology and Psychiatry, 30,* 691–709.

Green, J. A. (1992). Testing whether correlation matrices are different from each other. *Developmental Psychology, 28,* 215–224.

Harrison, A. O., Wilson, M. N., Pine, C. J., Chan, S. Q., & Buriel, R. (1990). Family ecologies of ethnic minority children. *Child Development, 61,* 347–362.

Heath, S. B. (1989). Oral and literate traditions among Black Americans living in poverty. *American Psychologist, 44,* 367–373.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47,* 1083–1101.

Hindelang, M. J., Hirschi, T., & Weiss, J. G. (1981). *Measuring delinquency.* Beverley Hills, CA: Sage.

Hirschi, T. (1969). *Causes of delinquency.* Berkeley: University of California Press.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1–123.

Jensen, A. R. (1980). *Bias in mental testing.* New York: Free Press.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Jöreskog, K. G., & Sörbom, D. (1988). *LISREL 7: A guide to the program and applications.* Chicago: SPSS.

Knight, G. P., Tein, Y. J., Shell, R., & Roosa, M. R. (1992). The cross-ethnic equivalence of parenting and family interaction measures among Hispanic and Anglo-American families. *Child Development, 63,* 1392–1403.

Lewontin, R. (1982). *Human diversity.* New York: Scientific American Books.

Loehlin, J. C. (1992a). *Latent variable models: An introduction to factor, path, and structural analysis* (2nd ed). Hillsdale, NJ: Erlbaum.

Loehlin, J. C. (1992b). Should we do research on race differences in intelligence? *Intelligence, 16,* 1–4.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence.* San Francisco, CA: W. H. Freeman.

Mackenzie, B. (1984). Explaining race differences in IQ: The logic, the methodology, and the evidence. *American Psychologist, 39,* 1214–1233.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410.

McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. (1973). Developmental changes in mental performance. *Monographs of the Society for Research in Child Development, 38* (Serial No. 150), 1–84.

McLoyd, V. C. (1990). The impact of economic hardship on Black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61,* 311–346.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ogbu, J. (1987). Cultural influences on plasticity in human development. In J. Gallagher & C. Ramey (Eds.), *The malleability of children* (pp. 155–168). New York: Brookes Publishing.

Ogbu, J. U. (1991). Immigrant and involuntary minorities in comparative perspective. In M. A. Gibson & J. U. Ogbu (Eds.), *Minority status and schooling: A comparative study of immigrant and involuntary minorities* (pp. 3–33). New York: Garland Publishing.

Osborne, R. T. (1980). *Twins: Black and White.* Athens, GA: Foundation for Human Understanding.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influences on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373–427.

Plomin, R., DeFries, J. C., & McClearn, G. E. (1990). *Behavior genetics: A primer* (2nd ed.). San Francisco, CA: W. H. Freeman.

Reschly, D. J. (1981). Psychological testing in educational classification and placement. *American Psychologist, 36,* 1094–1102.

Rodgers, J. L., Rowe, D. C., & Li, C. (1994). Beyond nature versus nurture: DF analysis of nonshared influences on problem behaviors. *Developmental Psychology, 30,* 374–384.

Roosa, M. W., Tein, J., Groppenbacher, N., Michaels, M., & Dumka, L. (1993). Mothers' parenting behavior and child mental health in families with a problem drinking parent. *Journal of Marriage and the Family, 55,* 107–118.

Rushton, J. P. (1988). Race differences in behavior: A review and evolutionary analysis. *Personality and Individual Differences, 9,* 1009–1024.

Russell, J. (1991). Culture and the categorization of emotions. *Psychological Bulletin, 110,* 426–450.

Slaughter-Defoe, D. T., Nakagawa, K., Takanishi, R., & Johnson, D. J. (1990). Toward cultural/ecological perspectives on schooling and achievement in African- and Asian-American children. *Child Development, 61,* 363–383.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42,* 137–144.

Spencer, M. B. (1990). Development of minority children: An introduction. *Child Development, 61,* 267–269.

Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement. *American Psychologist, 47,* 723–729.

Steinberg, L., Mounts, N. S., Lamborn, S. D., & Dornbusch, S. M. (1991). Authoritative parenting and adolescent adjustment across varied ecological niches. *Journal of Research on Adolescence, 1,* 19–36.

Stringer, C. B., & Andrews, P. (1988). Genetic and fossil evidence for the origin of modern humans. *Science, 239,* 1263–1268.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134–146.

Turkheimer, E. (1991). Individual and group differences in adoption studies of IQ. *Psychological Bulletin, 110,* 392–405.

Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: a follow-up of IQ test performance at adolescence. *Intelligence, 16,* 117–135.

Received August 5, 1993
Revision received December 20, 1993
Accepted February 10, 1994 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.