Implications of the Restricted Range of Family Environments for Estimates of Heritability and Nonshared Environment in Behavior-Genetic Adoption Studies

Mike Stoolmiller
Oregon Social Learning Center

Group and individual-difference adoption designs lead to opposite conclusions concerning the importance of shared environment (SE) for the child outcomes of IQ and antisocial behavior. This paradox could be due to the range restriction (RR) of family environments (FE) that goes with adoption studies. Measures of FE from 2 of the most recent adoption studies indicate that RR is substantial, about 67%, which corresponds to the top half of a normal FE distribution. RR of 67% cuts effect sizes and R^2 statistics by factors of 3 and 2–2.5, respectively. Because selection into an adoption study is inherently a between-family process and assuming that comparable restriction of genetic (G) influences are absent, estimates of SE, G, and nonshared influences will be substantially biased, respectively, down, up, and up by RR. Corrections for RR applied to adoption studies indicate that SE could account for as much as 50% of the variance in IQ.

The adoption study is generally considered the most powerful design for investigating genetic and environmental influences on behavior in humans (Plomin & DeFries, 1985). Plomin and De-Fries pointed out that even the most vociferous critics of human behavioral genetics acknowledge the power of the adoption design. As with all social science designs, however, the adoption study has its limitations, most of which have been well discussed in the literature. One potential limitation that deserves a second look is representativeness. The relevant question in adoption studies is How generalizable are the results? This has been a topic of much debate between behavioral geneticists and environmentalists and is the focus of this article. A brief introduction to some basic behavior-genetics terminology will be helpful to understand the debate before delving deeply into it.

Biometric model fitting produces a decomposition of the total variance of some phenotypic outcome of interest into two additive components: genetic and environmental. Environmental variance can be further partitioned into shared environment (SE) and non-shared environment (NE) variance. Other, perhaps more descriptive terms for these two components include between- and within-

family environmental variance. SE is that part of the environment that is constant across family members and makes family members alike. For example, the educational attainment of parents is constant across all siblings in a family at a given point in time and is thus an SE effect. Parenting style, to the extent that it is constant across siblings, is another example. NE is that part of the environment that varies across family members and functions to make individuals within families different from each other. For example, birth order varies across children (i.e., within but not across families), making it an NE influence. Parenting practices can also vary across siblings; to the extent that they do, they operate as NE effects. Note that most environmental determinants can vary both within and between families, meaning that they can operate as both SE and NE influences simultaneously.

Genetic variance can be further partitioned into additive and nonadditive components. Additive genetic variance is assumed to arise from many genes, each contributing a small individual effect to the overall phenotype. Nonadditive genetic variance can arise because of either dominance interactions between genes at a single locus or interactions between genes at different locations (epistasis). For either type of interaction, the quantitative phenotype obtained from the presence of the interacting genes is significantly different from what would have been expected from each gene acting individually. The population-genetic correlation between first-degree relatives on a trait in which all gene action is additive is expected to be .50. In contrast, nonadditive genetic influences can, at best, produce a population-genetic correlation of .25 if the nonadditivity is due completely to dominance interactions. As the nonadditivity shifts to epistasis, the expected genetic correlation declines toward zero (Grayson, 1989).

Assumptions about the nature of gene action on a trait have consequences for results of behavior-genetic twin designs. Grayson (1989) has demonstrated how undetected nonadditive genetic variance, either dominance or epistasis, can inflate heritability and deflate shared environmentality estimates in the classic twin de-

sign.[1] Hewitt (1989) concurred and concluded, "The bottom line is that, as has been repeatedly noted, classical twin study data alone cannot provide conclusive evidence of the absence of nonadditive genetic variance or shared family environmental effects or any number of other effects" (p. 607). Although dominance interactions can be modeled in extended twin designs, modeling epistatic interactions with human kinship data is generally considered intractable (Eaves, 1988). Largely for that reason, I focus solely on the adoption design in this article, because the correlation among unrelated children reared in the same adoptive family is a direct estimate of SE that does not depend on largely untestable assumptions about molecular gene action and is not affected by gene-environment correlation or interaction (Jensen, 1987; Plomin & Daniels, 1987).

For the classic additive-genetic model, the variance decomposition produces three proportions that sum to 1 and indicates the relative importance of SE, NE, and additive-genetic effects for any outcome of interest. This basic model has been enhanced in a number of important ways by behavioral geneticists. In fact, a consensus is emerging that it is time to move beyond the simple decomposition-of-outcome variance to more complicated and perhaps more realistic models of gene and environment transactions. In some sense, focusing on the original simple models might seem like moving backward. I do not argue that the simple models are adequate but, rather, that before moving on to more complicated models, it is important to have the best understanding possible of what the simple models reveal about the population of interest.

The debate over the generalizability of results obtained in adoption studies centers around the well-replicated findings that genetic and NE effects appear to account for most of the variance in important developmental outcomes during middle childhood, such as IQ and problem behaviors, and that SE plays a small role. Even this small role is alleged to vanish after adolescence (Plomin & Daniels, 1987; Scarr, 1992). Because of the respect that adoption studies command in terms of estimating the importance of SE, the lack of relative importance of SE has been highlighted recently by a number of behavioral geneticists. Plomin, Chipuer, and Neiderhiser (1994) concluded that, "for nearly all dimensions and disorders, shared family environment is not a major source of environmental influence" (p. 2). Scarr (1992) stated: "Being reared in one family, rather than another . . . makes few differences in children's personality and intellectual development" (p. 3). Still others have interpreted the modesty of SE influence to mean that parenting styles are relatively unimportant. Harris (1995) concluded that "children would develop into the same sort of adults if we left them in their homes, their schools, their neighborhoods, and their culture or subcultural groups, but switched all the parents around" (p. 461). Rowe (1994) recently published an entire book, *The Limits of Family Influence*, devoted to denying the importance of parenting practices. Rowe concluded:

> parents in most working-to-professional class families may have little influence on what traits their children may eventually develop as adults. Moreover, I seriously doubt that good childrearing practices can greatly reduce an undesirable trait's prevalence, whether it be IQ, criminality or any other trait of social concern. (p. 7)

These strong views, however, have not gone unchallenged.

Adoption-study critics, a group that includes both behavioral geneticists and environmentalists, have argued that adoptive fam-

ilies are not representative of the general population of families either because of the adoption situation itself (Baumrind, 1993; Jackson, 1993) or because adoptive parents are more highly educated, have higher occupational status, and are more affluent than the general population of parents (Baumrind, 1993; Boomsma, 1987; Jackson, 1993; Miles & Carey, 1997; Morton, 1987; R. J. Rose & Kaprio, 1987). Proponents of adoption studies, however, have argued that the effects of adoption itself are small once other confounding factors have been controlled (Rhea & Corley, 1994). Adoption advocates have also pointed out that although socioeconomic status (SES) is generally higher in adoptive samples, a broad range of SES is still usually represented in most adoptive samples, especially in older adoption studies, and results from these studies are consistent with more contemporary studies (Plomin & DeFries, 1985; Scarr, 1992).

The purpose of this article is to take a second and more detailed look at the representativeness of adoptive families. An adoption study involves at least three potent selection processes that have much in common. The first two, of course, involve the parents selecting themselves into the adoption process and the adoption agency selecting the parents. The third process, which may be just as important, is volunteering to be part of a study of child development. In this article, I review evidence that shows that these selection processes create a subpopulation in which the family child-rearing environment provided by adoptive parents is both above average in quality and restricted in range.

More specifically, I argue as follows: First, the range of family environments (FEs) represented in adoption studies is severely restricted compared with the larger population of American families. Second, evidence from a variety of sources indicates that FE is an important determinant of child outcomes such as IQ and antisocial behavior. Third, the selection processes for inclusion in an adoption study are inherently between-family processes, which means the between-family variance or SE variance component of the FE will be restricted. Finally, comparable restriction of NE and genetic variance in adoption studies compared with the larger population is unlikely, and, in fact, genetic variance might even be mildly inflated. Range restriction has the well-known effect of attenuating correlations. Biometric path models are based on correlations between adoptive siblings and between adoptive parents, biological birth parents, and adoptees. The range restriction of SE will attenuate correlations reflecting those causes, whereas correlations reflecting genetic and NE causes will be inflated. This pattern of differential attenuation and inflation is the reason for the apparent lack of relative importance of SE in behavior-genetic adoption studies.

It is sometimes stated that nonrepresentativeness of adoptive families can be corrected by collecting a control sample of bio-

---

[1] Put briefly, genetic nonadditivity lowers the genetic correlation for dizygotic (Dz) twins but not monozygotic (Mz) twins. The decline in genetic correlation will lead to a decline in the overall correlation for Dz twins but not for Mz twins, assuming that some of the phenotypic variance is due to genetic influence. Twice the Mz − Dz difference in correlations, the simply moment estimate of heritability from the additive twin model, is thus too big, and twice the Dz correlation minus the Mz correlation, the simple moment estimate of shared environmentality from the additive twin model, is thus too small. See Grayson (1989) for more details on this argument.

logical families matched to the adoptive families on family demographics, which presumably matches families on SE. For example, according to Scarr and Weinberg (1978): "An easier corrective for the selection of adoptive families is to have a comparison sample of biologically-related families that are similarly selected" (p. 676; see also Leahy, 1935, p. 244). In fact, most of the major adoption studies in the United States have used such a design, probably for this reason. I show, however, that the use of a matched biological control sample does not correct the biases that are due to range restriction.

Range Restriction to Family Environment in Existing Adoption Studies

Indirect Evidence of Range Restriction

Even if the selection processes in adoption were not so obvious, there are strong indications in the adoption literature that range restriction is operative. It is well known that adoption designs that focus on individual differences provide some of the strongest evidence that SE is not particularly important. It is less widely appreciated but equally well supported that adoption designs that focus on group differences provide some of the strongest evidence for the opposite conclusion, namely, that SE is very important. In the individual-difference design, the correlation between unrelated siblings in the same adoptive family is the most direct estimate of the importance of SE. In the group design, the mean level of the outcome for the adoptees is compared with a presumably genetically equivalent group reared in a very different environment, such as the birth parents, additional biological siblings of the adoptee who remain with the birth parents, or a normative value for children in the social milieu in which the adoptee would have been reared had the adoption not taken place. Locurto (1990), in a review of adoption group designs, concluded that adoption appeared to raise IQ by 12 points, or ¾ of a standard deviation, a substantial amount. Cloninger, Sigvardsson, Bohman, and von Knorring (1982) reported that the lifetime prevalence of adult arrest for adoptees in the Swedish sample was 12%, less than half the 26% prevalence of arrest for the biological fathers and quite close to the national average of 11% for Sweden at the time. For the Danish adoption sample, the adoptee lifetime prevalence of adult arrest was 16%, a little more than half the 29% prevalence of arrest for the biological fathers (Mednick, Gabrielli, & Hutchings, 1984). Fergusson, Lynskey, and Horwood (1995) found in the New Zealand Christchurch sample that adoptive adolescents as a group scored midway between children in two-parent biological and one-parent biological families on measures of antisocial behavior and substance abuse, despite the fact that the biological mothers of the adoptees did not differ significantly from the mothers in the one-parent biological families on a wide range of background variables. This group-level evidence indicates the beneficial aspects of being adopted from the lower class to the middle class and is consistent with a strong effect of SE on both IQ and deviance. The apparent contradiction in conclusions between group and individual-difference designs in adoption was considered in detail in an excellent review by Turkheimer (1991). This work, although it did not focus on range restriction per se, essentially anticipated the arguments in this article.

First, Turkheimer (1991) demonstrated that the currently popular explanation for the paradox, the so-called two-realms hypoth-

esis, is based on a misunderstanding of earlier work by McCall (1981) and is conceptually inadequate. The two-realms hypothesis holds that there are two separate and uncorrelated realms of IQ development: mean levels of IQ across populations, which are affected by the environment, and rank order of IQ within a population, which is primarily affected by genes. The two-realms hypothesis allows behavioral geneticists to claim that most of the variance in IQ within groups is genetic, whereas environmentalists can claim that FE produces large mean-level shifts across groups. The two-realms hypothesis is undoubtedly popular, because everybody wins.

Unfortunately for proponents of this position, Turkheimer (1991) argued quite persuasively that the two-realms view cannot hold for development in general, because it can be true only under the very special condition that there is no within-group variability on the relevant environmental determinants. With respect to adoption studies of IQ, the two-realms view can be true only if all adoptive FEs are functionally equivalent with respect to IQ-relevant determinants, something Turkheimer considered implausible. However, range restriction, if severe enough, could have exactly that effect. I pick up Turkheimer's argument[2] where he left off, at the point of considering it implausible that adoptive FEs are functionally equivalent, and submit that the selection forces involved in adoption studies are responsible for the range restriction that produces the near functional equivalence that makes the two-realms hypothesis appear to hold for adoption studies.

Direct Evidence of Range Restriction

The criteria of adoption agencies for placing children are not the only selection force involved in determining the composition of adoption samples. Self-selection by potential adoptive parents is probably equally as important as selection by adoption agencies. For example, Bachrach (1983) reported that in the U.S. population as a whole in 1976, based on the National Survey of Family Growth (NSFG), adoption was rare, but among women who were sterile, childless, about 30 years old, and married, 50% had adopted a child. In more recent work, Bachrach and her colleagues (Bachrach, London, & Maza, 1994; Bachrach, Stolley, & London, 1992) have shown that in the NSFG sample of women during the decade of the 1980s, both seeking and actually obtaining an adoption were significantly related to being White, to having higher educational attainment, and especially to strongly desiring more children than the woman expected to have.

The process of volunteering for a scientific study is also a strong selection process. The general characteristics of volunteer participants have been well documented (Rosenthal & Rosnow, 1975), and the problems this creates for adult twin studies has been reported on by Lykken and his colleagues (Lykken, McGue, & Tellegen, 1987; Lykken, Tellegen, & DeRubeis, 1978). In fact, one recent adoption study, the Minnesota Adoption Project (MAP), capitalized on volunteer bias as a convenient way of matching

---

[2] Actually, Turkheimer (1991) explored an alternative approach for resolving the contradiction between mean level and correlation results in adoption studies. He focused on raw regression coefficients for the measured variables that were taken as indicators of genetic and environmental effects. The advantages and disadvantages are discussed more fully in the section Statistical Consequences of Range Restriction.

biological families to the superior sociodemographic and parental IQ characteristics of the adoptive families (Scarr & Weinberg, 1978).

Thus, it is probably safe to assume that there is no serious dispute over the fact that powerful selection forces are at work in determining what kind of families will be present in an adoption study. Children simply are not placed in families characterized by obvious pathology in the parents or siblings, and it is unlikely that such families would even enter the adoption process or volunteer for a study of child development. Despite the obviousness of the selection processes involved and the controversy surrounding the extent of the generalizability of adoption studies, no study that I am aware of has documented the extent of the range restriction of the adoptive FE. Indeed, most of the existing adoption studies have very little detailed, prospective information about FE using norm-based measures, a limitation that has been discussed in detail elsewhere (Baumrind, 1993; Hoffman, 1991; Wachs, 1983).

The Colorado Adoption Project (CAP; Cherny & Cardon, 1994) and the New Zealand Christchurch samples, however, are recent exceptions. Checking for range restriction is relatively straightforward for norm-based measures in these samples. The simplest indication of range restriction is the ratio of the variance in the adoptive sample compared with the normative variance on the measure of interest. Another simple index is percentage restriction, which is merely 100 times the quantity 1 minus the variance ratio. On measures of the quality of the home environment using seven scales from the Home Observation for Measurement of the Environment (HOME) rating system (Caldwell & Bradley, 1978) at two points in time (when adoptees were 12 months and 24 months of age), CAP adoptive homes were not only above average compared with published norms but also highly restricted in variance—70% on average (the mean, 25th, 50th, and 75th quartiles of variance ratio distribution based on the 14 assessments were .30, .18, .32, and .38, respectively; Plomin & DeFries, 1985). Variance ratios are ratios of the CAP variances to the published normative variances.

The utility, however, of the normative values for the HOME provided by Caldwell and Bradley (1978) has been questioned (Plomin & DeFries, 1985). The original norming sample was recruited from Little Rock, Arkansas, and had a disproportionate number of socially disadvantaged families. Fortunately, the HOME is a widely used instrument, and published results for other samples by other investigators are available so it is not necessary to rely on only the published norms. To help put published results in perspective, it is useful to first consider the distribution of occupational status for households in the United States in 1983, which is close in time to when the CAP adoptees were growing up. I took the unemployed and people no longer seeking employment (mostly single-parent households on welfare) as the lower class; skilled blue collar, clerical, and unskilled blue collar workers as the working class; and professionals and managers as the middle class. I also based the classification of a two-parent family on the higher occupational status score in the pair. With these definitions and the information in S. J. Rose (1986), the breakdown was 10%, 53%, and 37%, respectively, for lower, working, and middle class. A sample that is not subjected to unusual selection procedures and has a similar SES distribution should be the most representative of families in the United States.

Table 1 was assembled from 443 references listed at Robert Bradley's web page.[3] Sampling from this bibliography, which is probably the most complete listing of published works that have used the HOME, was based on a quasi-random strategy. First, only works that pertained to the infant/toddler version (for comparability to CAP data) and that were published in a peer-reviewed journal or book chapter were considered eligible, which narrowed the search to 281 articles. The vast majority of references eliminated at this stage either used versions of the HOME for older children or were papers presented at conferences. Second, the article had to be available at the University of Oregon library, which further narrowed the search to 176 articles. The possibility that either of these two selection factors might introduce a substantial bias to estimates of the mean and variance of the HOME total score seems remote. Of the 176 articles, a fair number, 61, did not report means, standard deviations, or both. If a sample appeared to be predominantly middle class or was interesting for some other reason, I contacted the author to try to obtain the missing information (Bates, Olson, Pettit, & Bayles, 1982; and Wilson & Matheny, 1983, in Table 1). Of the remaining 115 articles, 34 had information of interest; the remaining 81 used the same samples as the 34 and so had no new information to contribute. In the interest of parsimony, I pooled information presented for independent subsamples in the original articles unless the subsamples were of particular interest because the standard deviations were comparable with the CAP values (Bee et al., 1982; Parks & Smeriglio, 1986; and Johnson et al., 1992, in Table 1). If an intervention was involved, only the control group values were used. For one study, as indicated in Table 1, I used the pooled within-group standard deviation derived from an F test. For the National Longitudinal Survey of Youth (NLSY) data, the mean and variance for the 45-item version was estimated from a shortened 26-item version. One potentially interesting middle-class sample (Metzl, 1980) was excluded because of the inability to reconcile the article's discrepant information concerning the standard deviation for the HOME total score in the control group.

Table 1 presents means and standard deviations for the HOME total score, sorted according to the standard deviation, and a description of the SES composition of the sample. The available SES information varied substantially across studies. In some cases, only the investigator's global description was available. Note that despite both the variability in the SES composition of the samples and the fair number of predominantly middle-class samples, CAP variances are still the most restrictive of any published study in my search. Note also that HOME means go up as one moves from lower- to middle-class samples. The correlation of SES indexes such as education and occupation with HOME scores is a well-replicated finding (Bradley, 1994). Homogeneous disadvantaged samples, as indicated in Table 1, have an average variance of about $6.4^2 = 40.96$. Despite their homogeneity on social disadvantage, however, these samples tend to be the most variable in home environments. The variance ratio for the CAP compared with the disadvantaged samples is .19; in other words, disadvantaged samples have about 5 times the FE variability of the CAP. This is an enormous degree of range restriction. The samples that come closest to matching the SES distribution for U.S. families are

---

[3] The web address is http://www.ualr.edu/~crtldept/abstract/home.html.

Table 1

Investigator, Sample Location, Means, and Standard Deviations of the HOME Total Score, and Sample Socioeconomic (SES) Description for Literature Review of Published Studies Using Bradley and Caldwell's HOME

Sample	HOME total score M	SD	Lower	Working	Middle
			SES distribution[b] (%)		
Sagi, Jaffe, Tirosh, Findler, & Harel, (1988), Israel	34.6	8.4[g]	Lower/working		
Holder-Brown, Bradley, Whiteside, Brisby, & Parette (1993), Disabled, Arkansas	32.4	7.6		None given	
Bradley et al. (1989), Little Rock, Arkansas	30.7	7.5	31	52	17
Wilson & Matheny (1983), Louisville, Kentucky twins[j]	32.3	7.5		Heterogeneous	
Reis, Barbera-Stein, & Bennett (1986), Illinois[h]	30.0	7.3	Lower (76% < 12 years of school)		
Unger & Wandersman (1988), South Carolina[h]	32.6	7.1	Lower 10.3[a]		
Weisglas-Kuperus, Baerts, Smrkovsky, & Sauer (1993), Rotterdam, the Netherlands	35.0	7.0	Lower/working (25% of sample unemployed or manual labor)		
Lozoff, Park, Radan, & Wolf (1995), Costa Rica	29.8	6.7		Working 10.5[a]	
Johnson, Houston, Texas[h,k]	28.5	6.8	48	52	0
Bee et al. (1982), low education, Seattle, Washington[h]	33.0	6.5		Working 11.6[a] (0.8)	
Stevens (1988), Atlanta, Georgia[h]	28.9	6.2	Lower 10.7[a]		
Benasich & Brooks-Gunn (1996), New York[h]	33.7	6.2	Lower/working 1.96[d] (0.9)		
Ramey, high risk, Chapel Hill, North Carolina[h,k]	27.9	6.1	39	58	3
Palti, Otrakul, Belmaker, Tamir, & Tepper (1984), control group, Israel	27.1	6.0		Working 10.5[a]	
Murray, Johnson, & Peters (1990), Omaha, Nebraska[h]	32.9	6.0	Lower/working 36.8[b] (16.4)		
Widmayer et al. (1990), Florida[h]	25.6	5.8	Lower/working		
Luster & Dubow (1992), NLSY[e,i]	31.5	5.8	Lower/working		
Barnard, Seattle, Washington[i,k]	36.5	5.7	7	29	64
Rock, Head, Bradley, Whiteside, & Brisby (1994), blind, Arkansas	29.9	5.6		None given	
Jordan (1978), St. Louis, Missouri	33.0	5.5		Working/middle	
Allen, Affleck, McGrade, & McQueeny (1981), Connecticut	34.1	5.6		Working 11.7[a] (2.2)	
Fein, Gariboldi, & Boni (1993), Italy	33.1	5.5		Working/middle 14.4[a] (4.8)	
Siegel, Hamilton, Ontario, Canada[i,k]	35.3	5.4	9	60	31
Parks & Smeriglio (1986), low & mid SES, Baltimore, Maryland	31	5.4	Lower/working 16.0[b] (2.2)		
Shonkoff, Hauser-Cram, Krauss, & Upshur (1992), New England	35.1	4.8			Middle 13.8[a] (3.1)
Luster & Rhoades (1989), teen mother, Michigan	35.5	4.6	Lower/working (teen mother's father's, education level = 9.5)		
Coll, Offman, & Oh (1987), Rhode Island	33.5	4.4			Middle 30.2[b] (8.0)
C. T. Ramey, Farran, & Campbell (1979), general population, Chapel Hill, North Carolina	38.1	4.3			Middle 15.1[b] (2.5)
Luster, Rhoades, & Haas (1989), rural New York	38.1	4.3		Working 12.7[a] (2.2)	
Lamb et al. (1988), Time 1, Sweden	36.5	4.3			Middle 6.1[c] (2.2)
Bates, Olson, Pettit, & Bayles (1982), Indiana[j]	31.4	4.2	13[f]	69[f]	16[f]
Johnson et al. (1992), Blacks, Galveston, Texas	36	4		Working 28.0[b] (11)	
Ross (1984), control, Ithaca, New York	34.9	4.0	Lower/working		
Parks & Smeriglio (1986), high SES, Baltimore, Maryland	35.2	4.0			Middle 48.8[b] (6.6)
Gottfried & Gottfried (1984), Fullerton, California[k]	36.4	3.7	0	24	76
Hollenbeck (1978), Bethesda, Maryland	35.0	3.7			Middle 15.6[a]
Bee et al. (1982), high education, Seattle, Washington	38.4	3.5			Middle 15.8[a] (1.8)
Lamb et al. (1988), Time 2, Sweden	40.0	3.4			Middle 6.1[c] (2.2)
Johnson et al. (1992), Whites, Galveston, Texas	40	3			Middle 41[b] (14)
CAP, Biological (Plomin & DeFries, 1985)	39.6	2.8			Middle 15.7[a] (2.2)
CAP, Adoptive (Plomin & DeFries, 1985)	39.1	2.8			Middle 15.7[a] (2.5)

Note. Standard deviations for SES indicators in parentheses. HOME = Home Observation for Measurement of the Environment rating system; NLSY = National Longitudinal Survey of Youth; CAP = Colorado Adoption Project.
[a] Years of education for father if available, otherwise mother. [b] Based on Hollingshead four-factor SES scores (lower < 20, 20 < working < 40, middle > 40). [c] Hollingshead occupation-status code (1–9 scale). [d] Education level (1 = *less than high school*, 2 = *high school graduate*, 3 = *some college or higher*). [e] Estimated values for 45-item scale based on 26 items used in the NLSY battery. [f] SES categories are: lower = unskilled, semiskilled, unemployed; working = skilled, white collar, college student; middle = professional, managerial. [g] Estimated from pooled within-group error from *F* test. [h] Homogeneous socially disadvantaged sample. [i] Roughly representative sample. [j] Mean and standard deviation provided by personal communication with investigator. [k] As reported in Bradley et al. (1989).

those of Barnard, Siegel, and the NLSY (see Table 1 and Bradley et al., 1989) with standard deviations around 5.5. The variance ratio for the CAP and these samples is about .26, or a factor of about 4, which is still a huge amount of restriction. To make the degree of restriction more concrete, consider Gottfried and Gottfried's (1984) samples of 24% and 76% working and middle class, respectively. This split corresponds roughly to the top, advantaged half of the U.S. distribution of occupations. To see this, note that the sample would have to include all of the middle class, which is 37% of the population, and 13% of the working class, which means that $^{37}/_{50}$, or 74%, of the sample is middle class and $^{13}/_{50}$, or 26%, is working class. These values come very close to matching Gottfried and Gottfried's SES values. The CAP is still restricted on the HOME by about 30% compared with Gottfried and Gottfried's sample. In fact, the CAP is most similar to the Swedish middle class and the highly educated middle-class subsamples of Bee et al. (1982), Parks and Smeriglio (1986), and Johnson et al. (1992), but still somewhat restricted compared with these high SES samples. On the basis of Table 1, the CAP appears to be representative of my definition of the middle class, which comprises a little more than the top third of the population. The variance restriction problem was so severe in the CAP that Plomin and DeFries abandoned the traditional scoring system and devised a new one based on 5-point rating scales and counts from videotaped interaction tasks. Although this is a sound strategy for improving the reliability of the measure by reducing error variance, it cannot create more true individual differences.

The New Zealand Christchurch sample is a population sample of 1,265 families that included 42 adoptive families as well as two-parent and single-parent biological families (Fergusson et al., 1995). Birth-mother characteristics were the same for the adoptee and single-parent biological groups on 16 measures of social background and perinatal status with the sole exception that the birth mothers of the adoptive children were 3 years younger on average. Compared with the single-parent biological families, however, the adoptive families were clearly advantaged throughout childhood in terms of childhood experiences, standards of health care, family material conditions, family stability, and mother–child interaction. In addition to higher means, adoptive families were also characterized by restricted variances, 55% on average, on all 16 above-mentioned environmental variables (the mean, 25th, 50th, and 75th quartiles of the variance ratio distribution based on 16 assessments were .45, .23, .41, and .57, respectively). To avoid overreliance on the HOME, I combined the variance ratios for the 30 measures of FE quality from the CAP (14 assessments) and New Zealand (16 assessments) studies. The mean and median of the distribution are, respectively, .38 and .34, implying about 62% or 66% restriction of variance. The distribution is positively skewed; therefore, the median of .34 is a better indicator of central tendency. Thus, I take a variance ratio of one third, or about 67% restriction, as a best estimate of the degree of range restriction of FE in adoption studies compared with the total population of American families. The variance ratio of FE in adoption studies, one third, is consistent with truncating a standard normal distribution at the median. In other words, if the FE quality distribution is approximately normal, adoption samples appear to be drawn from just the top half. This is somewhat less restricted than implied by the SES comparisons of the HOME total score discussed above with respect to Table 1, but it is probably better to

avoid relying too heavily on just the HOME in arriving at an overall figure.

Those variables that are not directly involved in the selection processes of adoption are restricted only to the degree that they are correlated with the direct determinants of the selection processes. This means that both more distal determinants of child outcomes, such as SES or parent IQ, and the child outcomes themselves, such as IQ or conduct problems, will not be as badly restricted as indicators of FE. Some adoption studies show very little or no restriction of variability on SES of adoptive parents; some behavior geneticists have argued that because of this, results from adoption studies are broadly generalizable (Plomin, Fulker, Corley, & DeFries, 1997; Scarr, 1993). Results from the CAP shown in Table 1, however, indicate that adoptive homes, despite any outward appearance of representativeness or lack of variance restriction on SES measures, are still range restricted on the FE variables that matter most in producing positive behavioral and intellectual outcomes (e.g., parenting practices). Adoption agencies screen out families with obvious parental or sibling pathology and, in the past, required both marital stability and a medical problem that prevented childbearing; these powerful screens would all serve to restrict the range of FE even if SES or income were not considered in placing children. For example, Capaldi and Patterson (1991) found that the number of marital transitions was highly correlated with both maternal antisocial behavior and parental supervision practices. In fact, Capaldi and Patterson's published data indicated that the number of marital transitions is the single best indicator of maternal antisocial behavior, better than Minnesota Multiphasic Personality Inventory (MMPI) self-report scales, drug use, police arrest, driver's license suspensions, and mother's age at the birth of her first child. Thus, maritally stable, low SES families who strongly desire to rear children and who are approved for adoption are probably still both above average and range restricted on key FE variables.

Other evidence of range restriction in adoption samples worth considering is the degree of range restriction on important child outcomes. I have shown elsewhere that the degree of restriction of the Texas Adoption Project (TAP; Horn, Loehlin, & Willerman, 1979) adoptee IQ distribution is substantial and consistent with truncation from below the SE distribution at about the 63rd percentile (Stoolmiller, 1998). In other words, based solely on the restriction evident in the IQ outcome distribution, adoptive families appear to be drawn from the top 37% of the SE distribution. Note that this replicates the comparisons based on SES for the HOME total score in Table 1.

Range restriction appears to be a major problem both with the adoptive samples in other countries and with outcomes such as antisocial behavior and substance abuse. Cloninger et al. (1982) reported that none of the adoptive parents of the 862 male adoptees in the Swedish sample had a history of arrest. This was well below the national average for Sweden (11%) and far less than the comparable rate in the biological fathers (26%). The Danish sample was similar (Mednick et al., 1984). Adoptive fathers had an arrest rate of 6%, below the national prevalence at the time for all Danish men, whereas biological fathers had an arrest rate of 29%.

Finally, no evidence has been found to suggest that the FE restriction in adoption samples cannot be readily duplicated by using either volunteer bias (Scarr & Weinberg, 1978) or family demographic matching (Plomin & DeFries, 1985). Thus, although

fertility problems might be a strong determinant of adoption-seeking behavior in married couples, they do not appear to create FEs that cannot be matched in normal biological families on a wide variety of variables considered to be the most direct determinants of positive child outcomes.

In summary, adoption studies have focused on families that tend to be White, older, more affluent, higher SES, higher IQ, maritally stable, and better functioning and that strongly desire to rear more children then they expect to bear. More importantly, these families are more uniform on these social advantages so that the variance of FE is badly restricted. The attendant range restriction (67%, or the top half of the FE distribution) renders a researcher relatively powerless to detect FE effects by using standard linear models or correlations within the adoptive sample.

Is Family Environment Significantly Related to Child Outcomes?

The fact that FE is severely restricted in adoption samples might not have much consequence if it were not an important determinant of child outcomes. Evidence from a variety of sources including behavior-genetic adoption studies, however, indicates that both FE and SE are important.

As mentioned previously, being adopted from the lower to the middle class appears to increase IQ by 12 points and to cut lifetime prevalence of arrest nearly in half. Both of these findings are indicative of the importance of FE and SE. In addition, despite the obvious range restriction, adoptive-family SES was a significant predictor of adoptee criminality, even after controlling for genetic risks, in both the Swedish and Danish samples (Cloninger et al., 1982; Mednick et al., 1984). Teasdale and Sorensen (1983) also found in the Danish data that adoptive-father income, education, and occupational prestige predicted adoptee educational attainment. The effects of family SES in the Swedish and Danish studies (despite both the inherent range restriction and the fact that SES is at best a crude and indirect indicator of FE) are not consistent with claims that FE and SE does not matter.

The strongest evidence for the importance of FE, however, comes from randomized trials in which aspects of FE have been experimentally manipulated to test a hypothesized causal connection to a child outcome. With respect to both IQ and deviance, numerous trials have been conducted and were reviewed by Brestan and Eyberg (1998), Jackson (1993), Kazdin (1987, 1994), and S. L. Ramey and Ramey (1992), among others. The evidence clearly supports a causal connection between parental intellectual stimulation and discipline and supervision practices for IQ and deviance, respectively. Note that these results do not rule out child effects or genetic effects, but they do unambiguously establish the causal status of parenting practices. A very interesting addition to the studies reviewed above is randomized trials involving Treatment Foster Care programs. Chamberlain and Reid (1998) reported that this approach is far more effective at curbing preadolescent and early adolescent delinquency than are the standard alternative treatments available. Even more important, however, is the fact that the treatment effect appeared to be mediated by differences in discipline and supervision strategies (Chamberlain, 1997). So, the findings in randomized trials with biological parents and their offspring mentioned above seem to generalize to situations in which the parents are not genetically related to the child. Note the

similarity between the foster parent results and the findings of large group differences in adoption studies. The causal status of parenting practices on IQ and deviance is not really the issue, as this has been demonstrated repeatedly in randomized trials. The important issue is how the selection processes that operate in adoption studies affect the SE and NE variability relative to the larger population of American families.

Restriction of Shared or Nonshared Variance?

The selection processes involved in adoption are by definition between-family processes. Some families are selected because of the high quality of the FE they can provide, and some families who cannot provide as good an FE will not be selected. Half the difference between sibling environment scores is an estimate of the NE for that measure (Rovine, 1994; Turkheimer & Waldron, 1998). Because many couples do not even have children when they begin the adoption process, it is hard to see how they could possibly be selected for nonshared or within-family effects. Even if it were possible, selection for any particular score on the NE (i.e., particular values for the sibling environmental difference score) makes no sense in the adoption context. It does nothing to ensure that children are not placed in bad families because there is no necessary relation between a sibling environmental difference score and the overall quality of the home environment. Thus, regardless of how any particular aspect of FE breaks down in terms of shared and nonshared effects, the shared portion of the overall FE effect will be attenuated by the between-family selection processes. The nonshared or within-family portion of the FE effect will not be altered except under unusual circumstances that are discussed shortly. These ideas are better illustrated by Figure 1.

Idealized FE data for 20 families, each with two siblings, are plotted in Figure 1. All random variation has been eliminated to visually simplify the figure. The 20 families are sorted from lowest to highest on average FE across the siblings; this sorting is depicted by the diagonal solid line. Neither the sorting nor the smooth linear increase of FE is necessary to the argument but simply lends visual clarity to the figure. The individual values of FE for Siblings 1 and 2 are, respectively, shown by the circles above and below the line for the average FE. The assignment of siblings to the Sibling 1 or 2 designation is arbitrary. The circles for sibling pairs are connected by a vertical straight line to visually emphasize the grouping of siblings within families. Note that the range of variability for siblings within families is constant across the entire range of variability of average FE. This is an important background assumption of the standard biometric model. Now, assume that an adoption sample is recruited and that only the top half of the distribution of FE is represented as a result of the selection processes at work. The issue of genetic variance restriction or inflation is not addressed in Figure 1 because adoptive parents do not contribute genes to the adoptee. The topic of genetic variance is discussed later. The selection of the top half of the FE distribution is depicted in Figure 1 by the dashed lines. The vertical dashed line separates the top 10 families from the bottom 10 families, and the horizontal dashed line gives the cutoff point on the FE scale, in this hypothetical case, zero. There is now only half the range of variability for the shared part of FE, but the nonshared part of FE remains the same. Thus, the relative magnitudes of the shared and nonshared effects of FE have shifted substantially in
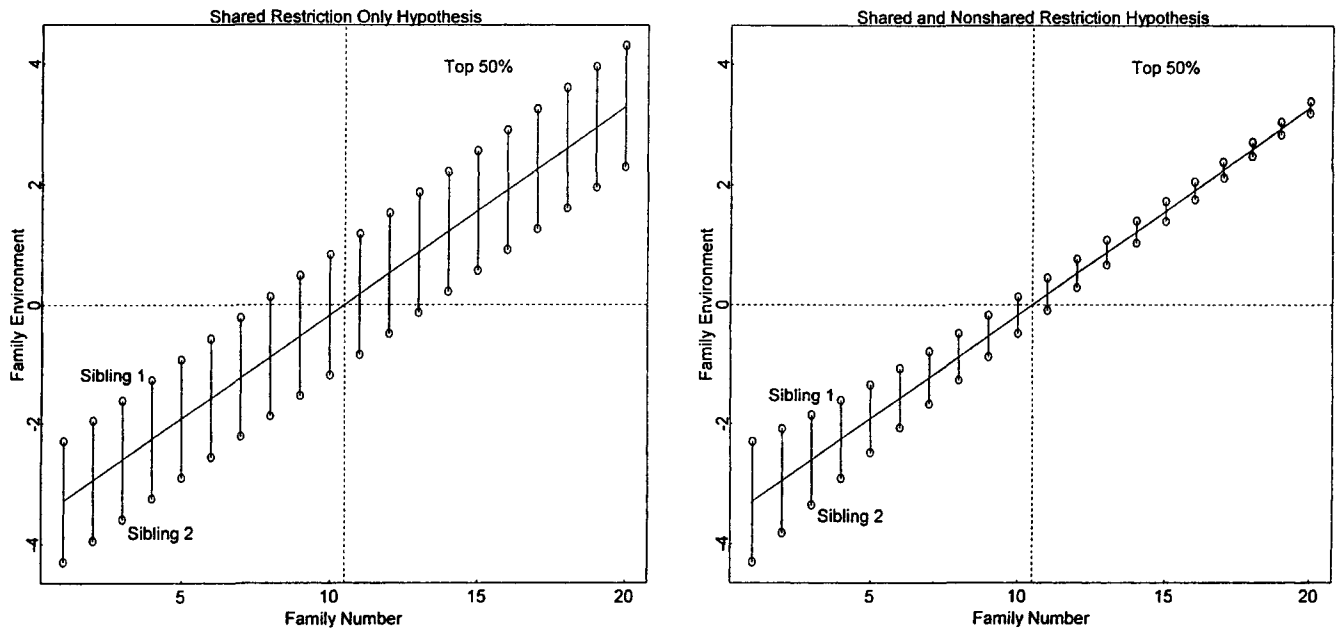
Figure 1. Restriction of shared or between-family variability in sibling environments with no effect on nonshared or within-family variability (panel on the left) and strong negative correlation between level of shared environment and variability of nonshared environment (panel on the right).

favor of the nonshared effect. In order to produce estimates that match the larger population, the range of variability of the non-shared part of FE and genetic influences would have to be similarly restricted to match the between-family selection processes in adoption. But what would bring about this restriction for the within-family effect of FE? There is nothing in the standard biometric model that would imply such restriction, and in fact, if it did happen, it would represent a violation of the standard assumption that sibling variability is constant across the range of average FE.

Figure 1 shows a situation where siblings encounter more similar FEs as the average quality of FE goes up. In terms of parenting practices, more overall skill in child rearing goes with greater consistency across siblings. The figure depicts a smooth relation, but the relation could also be abrupt and discontinuous. In this scenario, the between-family selection process results in restriction of both between- and within-family variability because of the linkage between the two. Change over time in FE could be one source of such a linkage. For example, although something like family income would be a shared influence at any given point in time, it could be important as a nonshared effect if siblings are exposed to different levels of family income at equivalent points in their respective development and if the timing of the exposure to poverty has important developmental consequences.[4] If adoptive families are more stable over time on family income, however, this source of within-family variance might be restricted. Alternatively, something like Scarr's niche-picking hypothesis would predict just the opposite, that within-family sibling environmental variability would increase as FE quality increases (Scarr, 1992) because siblings have the parental latitude to pursue their own interests and proclivities. I am not aware of any research that has demonstrated such a linkage in either direction. Plotting the average sibling

environment versus the absolute value of the difference in sibling environments is a simple graphical technique for detecting such a linkage. Careful empirical scrutiny of these opposing possibilities, however, is beyond the scope of this article but is an important topic for future research. In summary, restriction of nonshared variability cannot be safely ruled out because the correlation between the level of the SE and the variance of the NE is usually assumed to be zero rather than being empirically determined. Given that this standard assumption does hold, however, it is important to understand the implications of the between-family selection process for adoption studies.

Restriction of Genetic Variance?

Assessing the possibility of inflation or restriction of genetic variance in the adoptee population is very difficult because (a) direct assessments of all relevant genes for the phenotypes of interest are not yet available and (b) birth-parent phenotypic variance is an ambiguous indicator of genetic variance unless restriction or inflation of birth-parent environmental variance can be ruled out, which would require a direct assessment of the birth parents' relevant developmental environment (including their childhood environments). Nonetheless, it is worthwhile considering some of the issues involved and what little evidence is available.

The three issues to consider are (a) which couples are involved in an unplanned, unwanted pregnancy, (b) which couples are likely to resolve the problem by relinquishing the baby for adoption, and (c) which babies are likely to be placed. The last issue is not

---

[4] Thanks to Jeanne Brooks-Gunn for pointing out this example.

considered in detail because, for the CAP and TAP, the babies were placed at or very near birth, which makes selection based on qualities of the adopted infant unlikely.

It is clear that the vast majority of couples involved in unplanned, unwanted pregnancies are unmarried couples in their late teens (Pearson & Amacher, 1954; Plomin & DeFries, 1985). Although birth-mother information is common, almost nothing is known about birth-father characteristics in the American adoption studies. Some information on biological fathers was collected for the CAP but only on about 20% of the sample, and it is doubtful that these volunteer fathers were representative of all the biological fathers of CAP adoptees. In addition, almost nothing is known about birth-father involvement in the relinquishment process. This is unfortunate because birth fathers contribute half of the genes in question and thus constitute a wild card in any discussion of the adoptee gene pool.

It seems reasonable to assume that because the birth parents are involved in an unplanned teenage pregnancy, they are probably more antisocial and, by virtue of the negative correlation between IQ and antisocial behavior, have lower IQ scores than do birth parents in general. This is consistent both with Swedish and Danish birth-father results mentioned earlier and with research on teenage sexual activity and pregnancy (Capaldi, Crosby, & Stoolmiller, 1996). Pearson and Amacher (1954) also noted a surprisingly high level of psychopathology in their birth-mother sample, although they lacked the normative data necessary to make a careful comparison. Horn, Green, Carney, and Erickson (1975) noted elevated rates of psychopathology on the MMPI, especially psychopathic deviancy and schizophrenia, in the TAP birth-mother population and ruled out pregnancy and age as confounding factors by comparing birth mothers with married pregnant women and with 12th-grade high school girls. Birth mothers in the CAP also had elevated levels of sociopathy compared with adoptive mothers, and, incidentally, the variance for birth-mother sociopathy was almost 5 times as large as that for the adoptive mothers. Results for a small volunteer subsample of birth fathers were similar but not as marked.

With respect to IQ, Scarr and Weinberg (1978) found that the birth-mother distribution of education showed no evidence of restriction and argued against any genetic selection for the adolescent MAP adoptees. In contrast, Horn et al. (1979) raised the issue of genetic selection in the TAP, because all the adoptees came from a single adoption agency that catered to middle-class birth-mother families, and the distribution of IQ scores for the birth mothers was also restricted.

There is some indication in the literature on relinquishment that women with higher educational aspirations tend to relinquish their babies, which suggests genetic selection for higher IQ (Bachrach et al., 1992), at least with respect to birth mothers. Other results, however, indicate that women who relinquish their babies come from less cohesive families with histories of alcohol abuse (Bachrach et al., 1992). If these two opposing forces selected for opposite ends of the IQ distribution, they might tend to either cancel out or perhaps even lead to mildly inflated genetic variance. Such a tendency has in fact been observed. Pearson and Amacher (1954) found that for almost 3,000 mothers who relinquished a child in Minnesota from 1946 to 1951, the mean IQ was about 100, the same as the population normative value, but the standard deviation of the IQ distribution was 18, about 24% larger than the population

value of 16. Close inspection of the empirical distribution revealed excesses at the high and low ends and a deficiency in the middle compared with the expected normal distribution. Pearson and Amacher speculated that birth mothers in the middle range of IQ kept their babies more often than did the other mothers, because it was more socially acceptable to do so in their social milieu.[5]

Thus, the available data do not suggest genetic selection for either IQ or deviance in adoptee populations. The weight of evidence is stronger for deviance than for IQ, but, nonetheless, it suggests that, if anything, genetic variance might be mildly inflated for deviance and IQ, at least for birth mothers. These effects, however, are likely to be diluted by birth fathers because they are not usually involved in the relinquishment process. For specific adoption studies that use only one or two adoption agencies targeting middle-class birth mothers, genetic and environmental variance might be restricted, although, again, the genetic effect is probably diluted by the birth fathers. Clearly, however, better answers to these issues will have to await advances in the technology for assessing specific genetic contributions to the phenotypes in question because environmental assessments for birth parents will probably never be very feasible.

In summary, range restriction of FE quality is an obvious and inescapable aspect of adoption studies, and the causal status of FE on child IQ and deviance is well established. Selection is by definition a between-family process and will therefore attenuate the shared component of the FE effect. Whether comparable restriction of genetic and NE influences exists is unknown but seems unlikely. Although corrections to estimates of SE represent an improvement over making no corrections at all, they should probably be considered upper limits, at least until more definitive data are available.

Statistical Consequences of Range Restriction

Consider the linear regression model $y_i = B x_i + \epsilon_i$, with ϵ_i the normal homoscedastic random-error term. In the face of range restriction on the predictor, x, the regression weight, B, and error variance, σ_ϵ^2, are still estimated without bias, but the correlation, effect size (f^2), and the t or F statistic will all be biased back toward zero. The formula for effect size, f^2, and t or F statistic makes these relations explicit:

$$f^2 = \frac{B^2 \sigma_x^2}{\sigma_\epsilon^2} = \frac{t^2}{N-2} = \frac{F}{N-2} \qquad (1)$$

(Cohen & Cohen, 1983). Notice how both the effect size, f^2, and the F statistic are directly proportional to the variance of the predictor, σ_x^2. If the variance of x is restricted, then the effect size and F statistic decline proportionately. For example, in adoption

---

[5] Bachrach et al. (1992) have shown a major downward trend in relinquishment of children for adoption during the 1970s that did not stabilize until the mid-1980s. Although there may have been a number of reasons for the change, its magnitude suggests that caution should be used in generalizing from adoption studies carried out before the change to those carried out after the change. Thus, caution should be attached to the generalizability of the Pearson and Amacher (1954) findings concerning relinquishing mothers in the 1950s as compared with relinquishing mothers in the late 1980s to date.

studies with 67% range restriction of x, effect sizes and F statistics will be about $1/(1 - .67) = 3$ times too small. For correlations, the relation of the variance of x to the degree of attenuation is a little more complicated. The Pearson–Lawley formula, however, clearly highlights the role of the variance ratio in corrections for range restriction:

$$r_{full} = \cfrac{\dfrac{\sigma_{x,\,full}}{\sigma_{x,\,res}} r_{res}}{\sqrt{1 + \left(\dfrac{\sigma^2_{x,\,full}}{\sigma^2_{x,\,res}} - 1\right) r^2_{res}}}, \qquad (2)$$

where the subscripts *full* and *res* refer, respectively, to the unrestricted and restricted samples, and $\sigma^2_{x,\,full}/\sigma^2_{x,\,res}$ is the variance ratio. The attenuated R^2 for a variance ratio of .33, across the range of full correlations from .20 to .70, can be easily computed from the Pearson–Lawley formula and ranges from .36 to .50 times the full R^2. In other words, R^2 for measures of FE predicting child outcomes in adoption studies will be cut at least in half and could be cut almost to one third. In summary, range restriction of FE in adoption studies will lead to effect sizes and F statistics that are one third and R^2 statistics that are, at best, one half their magnitude in more normative samples. This has important implications both for the interpretation of existing studies and for statistical power considerations in the design of future adoption studies.[6]

Turkheimer (1991) addressed these same issues, but rather than advocating corrections for range restriction, he suggested that attention should be focused solely on the raw regression weights. The raw regression weights do not change with changes in population-genetic and environmental variances as do standardized quantities, such as heritability and shared environmentality, and thus, the raw regression weights are not affected by range restriction.

Turkheimer's (1991) strategy has the advantage of being a simple and direct solution to the problem of range restriction, but it has some disadvantages, too. The first problem is that it is a statistically inefficient approach to dealing with adoption designs in which range restriction is an inherent feature. Trying to interpret a raw regression weight without first rejecting the null hypothesis that it is zero is problematic. Thus, sample sizes have to be appreciably larger to compensate for the attenuation in effect size using Turkheimer's approach. Second, Turkheimer's strategy does not generalize well to multiple dimensions of environmental influence. Trying to decide what is and is not important by relying on raw regression weights when measurement scales vary arbitrarily across constructs is problematic. Partial R^2 statistics and standardized regression weights have traditionally been used in this context to aid interpretation, and these are affected by range restriction. Finally, and most importantly, Turkheimer argued that the best model for determining the importance of SE for IQ is the path model for adoptive sibling resemblance shown in his Figure 6 (see Turkheimer, 1991, p. 403). Unfortunately, raw regression weights are biased by range restriction for this model. In Turkheimer's Figure 6, the raw regression weight, e, for the effect of SE can be estimated by the square root of the covariance between IQ_1 and IQ_2. The Pearson–Lawley selection formula can be used to derive the covariance between IQ_1 and IQ_2 in Turkheimer's Figure 6, assuming range restriction shrinks the variance of the SE. The expression is:

$$cov_{res}(IQ_1, IQ_2) = e^2_{res} = var_{full}(IQ)\ cor_{full}(C,IQ)\ var_{res}(C), \qquad (3)$$

where $var_{full}(IQ)$ is the variance of IQ in the full population, $cor_{full}(C,IQ)$ is the correlation between SE, C, and IQ in the full population, and $var_{res}(C)$ is the variance of the SE in the restricted population. This expression demonstrates that as the variance of the SE fluctuates across populations, so does the raw regression weight for the effect of SE on IQ. The key difference between this model and the regression models, on which Turkheimer's argument was built, is that the independent variable is latent, and thus, the variance of the independent variable is unobserved and indeterminate. The variance of SE (E in Turkheimer's Figure 6 and C in my notation) has to be fixed at some arbitrary value (usually 1) in order to estimate e. Unlike an observed variable regression model for which the raw regression weight is unbiased even in the face of range restriction, the model in Turkheimer's Figure 6 provides biased estimates of the effect of SE when range restriction is operative. I agree completely with Turkheimer's emphasis on the utility of the raw regression weights, but I do not endorse Turkheimer's recommendation to ignore environmentality or heritability estimates as long as sampling issues are carefully attended to and appropriate corrections are taken when necessary.

Consequences of Range Restriction on Biometric Path Models

The model shown in Figure 2 is a classic biometric path model for sibling resemblance on some trait of interest, V, where V_1 and V_2 are measures of V for Sibling 1 and Sibling 2, respectively. The additive influence of genes on the level of the trait is captured by G_1 and G_2, NE influences are captured by E_1 and E_2, and C represents SE influences. The correlation of G_1 and G_2 is fixed at 0 in the adoption sample and at .5 in the biological sample to reflect assumptions about the genetic correlation for additive genetic effects.

What I present is an application of the Pearson–Lawley selection formula to the standard biometric sibling-resemblance model. This approach is implicit in the work of Meredith (1964) and has been developed for the Structural Equation Modeling framework by Dolan and Molenaar (1994). Curiously, however, interest in the effects of and models for selection and truncation effects in behavior genetics has been confined almost completely to method-

[6] As noted in Footnote 5, the supply of healthy babies for permanent adoption at birth dropped substantially after the mid-1980s. Increased competition among White couples for the most "desirable" babies has probably had the effect of making contemporary adoptive families even less representative of the entire population than they were prior to and during the 1970s. Because all three of the major adoption studies in the United States cited above (MAP, TAP, and CAP) included adoptive families from before the major downturn in relinquishments, they are likely to provide conservative estimates of range restriction for more contemporary studies. These issues make power considerations even more important for future adoption studies. A full consideration of power issues, however, including statistical corrections for range restriction, is beyond the scope of this article. I simply note that modern missing-data techniques (Arbuckle, 1996; Little & Rubin, 1987) can be used to obtain better estimates and to increase the power of the significance tests involved, provided a nonrestricted sample can be obtained. Work is forthcoming that describes new techniques for dealing with the problem (Stoolmiller, 1998).
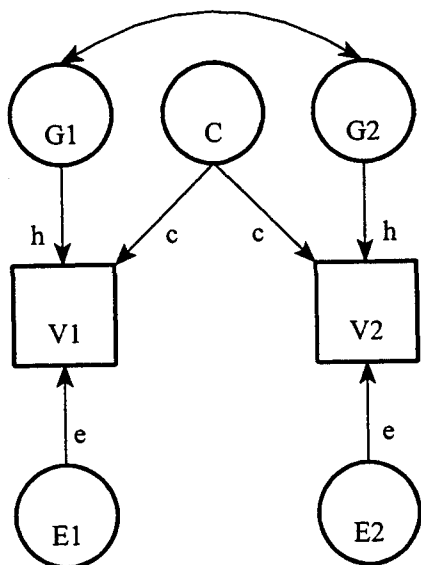
Figure 2. Classic biometric path model for adoptive and biological sibling resemblance.

ologists focused on twin designs (Lykken et al., 1978, 1987; Martin & Wilson, 1982; Neale, Eaves, Kendler, & Hewitt, 1989). This is surprising given that range restriction is an inherent part of the adoption design but need not be so for twin designs. Neale et al. verified that truncation of the distribution of the outcome of interest—for example, IQ—biased estimates of shared environmentality downward, as previously found by Martin and Wilson. Thus, prior methodological work in the twin field indicates that restriction and truncation of outcome distributions can be problematic for obtaining accurate estimates of SE, even though the truncation is not based on SE. This work suggests that if the truncation is based directly on SE, then the bias will be even more extreme, and thus, models for correcting range restriction are important for adoption studies.

I assume throughout this discussion that the models and covariance matrices represent population values. Sampling variability does not alter the substance of the argument. The model in Figure 2 also assumes that complications, such as selective placement, assortative mating, and Gene × Environment correlation and interaction, are all negligible. Although some of these assumptions

may not be reasonable, they, too, do not change the basic substance of my argument, and they introduce unnecessary complications for understanding the consequences of range restriction of SE.

Table 2 shows population variances, covariances, and correlations for the adoptive and biological siblings for the model in Figure 2 for two equivalent but alternative parameterizations. The first parameterization corresponds to the diagram in Figure 2. The latent variances are fixed at 1, and the path coefficients are estimated. In the second parameterization, the effects of the latent variables are fixed at 1, and the latent variances are estimated. Simple formulas relate the parameters in one model to the parameters in the other (Neale & Cardon, 1992). For example, $h^2 = var(G_1) = var(G_2) = var(G)$, $c^2 = var(C)$, $e^2 = var(E_1) = var(E_2) = var(E)$. The reason this complication is introduced is because the first parameterization is probably more commonly used. The second, however, makes both the assumptions about the equality of variances across adoptive and biological samples more explicit and the effects of range restriction of variances on model parameters more obvious.

First, the variances of V_1 and V_2 are assumed to be equal. This variance is referred to as the variance of V. This equivalent sibling assumption is standard, although it could be relaxed for substantive reasons such as gender or age effects. These are unnecessary complications, however, for understanding range restriction. Second, the correlation between unrelated adoptive siblings is

$$\frac{var(C)}{var(G) + var(E) + var(C)}, \quad (4)$$

which is the ratio of the variance due to SE over the total variance because $var(V) = var(G) + var(E) + var(C)$. In other words, the correlation between unrelated adoptive siblings is a direct estimate of the proportion of variance due to SE.

My review suggests that $var(C)$ in the adoption population is about one third the variance in the total population, so I express $var(C)$ in terms of the total population variance as $(1 - x_C)$, where $x_C = \frac{2}{3}$ represents the bias on the variance of C. Making this substitution, I get

$$S_A = cor_A(V_1, V_2) = \frac{(1 - x_C)var(C_T)}{var(G_T) + var(E_T) + (1 - x_C)var(C_T)}, \quad (5)$$

where the A, B, and T subscripts indicate adoptive, matched biological, and total populations, respectively, and S is the propor-

Table 2
Expected Second Moments for Adoptive and Biological Siblings

	Model parameterization	
Expected second moment	Path coefficients	Variance components
$var(V_1) = var(V_2) = var(V)$	$h^2 + e^2 + c^2$	$var(G) + var(E) + var(C)$
$cov(V_1, V_2)$, adoptive	c^2	$var(C)$
$cov(V_1, V_2)$, biological	$h^2/2 + c^2$	$var(G)/2 + var(C)$
$cor(V_1, V_2)$, adoptive	$c^2/(h^2 + e^2 + c^2)$	$var(C)/(var(G) + var(E) + var(C))$
$cor(V_1, V_2)$, biological	$(h^2/2 + c^2)/(h^2 + e^2 + c^2)$	$(var(G)/2 + var(C))/(var(G) + var(E)$ $+ var(C))$

Note. Var = variance; cov = covariance; cor = correlation; h = additive genetic path coefficient; c = shared environmental path coefficient; e = nonshared environmental path coefficient.

tion of variance due to SE. Note that additive genetic and NE variances are assumed to be the same in the adoptive, matched biological, and total populations and that the adoptive and biological populations are matched for restriction of SE. These assumptions reflect the use of a sample of biological families that are carefully matched on FE characteristics to the adoptive families. As demonstrated and discussed below, however, this does not eliminate the bias due to range restriction. Equation 5 can also be obtained by applying the Pearson–Lawley selection formula to the covariance matrix of observed and latent variables implied by the model in Figure 2. Dividing the numerator and denominator by $var(V_T)$ and noting that S_T is the proportion of variance due to SE in the total population gives

$$S_A = cor_A(V_1, V_2) = \frac{(1 - x_C)S_T}{\frac{var(G_T)}{var(V_T)} + \frac{var(E_T)}{var(V_T)} + \frac{var(C_T)}{var(V_T)} - x_C S_T}. \quad (6)$$

But note that $var(G_T)/var(V_T) + var(E_T)/var(V_T) + var(C_T)/var(V_T) = 1$, so that

$$S_A = cor_A(V_1, V_2) = \frac{(1 - x_C)S_T}{1 - x_C S_T} = \frac{S_T - x_C S_T}{1 - x_C S_T}. \quad (7)$$

Substituting the estimate of two thirds (from my literature review of the selection bias factor for SE) for x_C gives

$$S_A = cor_A(V_1, V_2) = \frac{\frac{1}{3}(S_T)}{1 - \frac{2}{3}(S_T)} = \frac{S_T}{3 - 2S_T}. \quad (8)$$

In Equation 7, it is clear that the effect of range restriction is to subtract the factor $x_C S_T$ from both the numerator, S_T, and the denominator, 1, of the unbiased estimate $S_T/1$. Because the numerator is smaller than the denominator to start with, subtracting the same thing from both results in a larger proportional decrement in the numerator than in the denominator. Thus, the biased estimate is always smaller. In fact, the only time the adoptive-sibling correlation is not biased downward is when S_T is equal to 1 or 0, that is, when all or none of the variability in V is due to SE. Equation 8 can be rearranged to solve for S_T, the proportion of variance due to SE in the total population:

$$S_T = \frac{3\ Cor(V_1, V_2)}{1 + 2Cor(V_1, V_2)} = \frac{3\ S_A}{1 + 2S_A}. \quad (9)$$

Although Equation 9 is simple, it also has a substantial negative bias (i.e., underestimates S_T) in some circumstances even when the true value of x_C is known. The estimator given in the first equation of Appendix B of Stoolmiller (1998) performs much better provided the variance of the outcome in the total population can be obtained.

Using the same algebraic maneuvers, I compute similar expressions for effects of genes and NE, although I leave out the intermediate steps. I denote the proportion of genetic variance in V_T, or $var(G_T)/var(V_T)$, as H_T and note that N_T, or $var(E_T)/var(V_T)$, is, therefore, $1 - S_T - H_T$, because all three proportions must sum to 1. Then,

$$H_A = \frac{var(G_T)}{var(V_A)} = 2\left(cor_B(V_1, V_2) - cor_A(V_1, V_2)\right) = \frac{H_T}{1 - \frac{2S_T}{3}}, \quad (10)$$

and

$$N_A = \frac{var(E_T)}{var(V_A)} = \frac{3(1 - S_T - H_T)}{3 - 2S_T} = \frac{(1 - S_T - H_T)}{1 - \frac{2S_T}{3}}. \quad (11)$$

Equations 10 and 11 can also be rearranged to solve for proportions of variance due to additive genetic and NE influences in the unrestricted population, which are, respectively,

$$H_T = H_A\left(1 - \frac{2S_T}{3}\right), \quad (12)$$

and

$$N_T = N_A\left(1 - \frac{2S_T}{3}\right). \quad (13)$$

Does a Matched Biological Control Sample Correct Biases Due to Range Restriction?

If the biological families were not carefully matched to the adoptive families on FE, then the standard model becomes inestimable because the environmental variances are different across the groups. Thus, in order to estimate the model at all, matching is important and preserves the integrity of the internal comparisons. Indeed, in some cases, the target population of interest may not be the most broadly representative population of families. In these cases, the matched biologic control family strategy is a sound one for studying interesting subpopulations, as long as such subpopulations are well defined. As is often pointed out, heritability (and environmentality) estimates apply only to the specific population at the specific time period from which the estimates were obtained. The matching strategy does not, however, compensate for the restriction of environment if one is interested in estimates that pertain to the total population. The expression to the right of the first equal sign in Equation 10 offers the clearest explanation of why this is so. The numerator, $var(G_T)$, which, as shown in Table 2, is estimated as 2 $(cov_B(V_1, V_2) - cov_A(V_1, V_2))$, is an unbiased estimate of the genetic variance for the total population, but it is divided by $var(V_A)$, the total phenotypic variance for the adoptive and matched biological populations, which is too small because it contains the restricted SE variance component. Thus, the overall expression overestimates the proportion of genetic variance in V. Note, also, that the overestimation will not be as bad for either genes or NE individually as is the underestimation of SE, because the bias will be split between both effects. If either NE or genetic effects are very small, however, the other will absorb more of the bias. For example, if genetic effects are completely absent, all the bias will be concentrated in NE and vice versa.

Specific Examples From the Adoption Literature

As an example, Horn et al. (1979, Table 13, p. 199) reported for the TAP that across four slightly different models, the average relative proportions for genetic, SE, and NE effects were, respectively, .49, .26, and .25. First, applying Equation 9, $S_T = .51$. Next, applying Equations 12 and 13, H_T and N_T are found to be, respectively, .32 and .21. I have shown elsewhere, using a different

approach based solely on the level of restriction evident in the TAP adoptee IQ distribution, that $S_T = .55$ (Stoolmiller, 1998).

As another example, consider Table 4.5 in Cherny and Cardon (1994, p. 52) based on CAP IQ data from ages 1 to 9 years. Averaging estimates across time, the genetic, SE, and NE variance proportions are, respectively, .42, .13, and .45. First, applying Equation 9, $S_T = .31$. Next, applying Equations 12 and 13, H_T and N_T are found to be, respectively, .33 and .36. Because specific information on FE range restriction is available for the CAP, it is interesting to base corrections on only this information. As shown in Table 1, the variance ratio for CAP HOME total scores is about .26 compared with representative samples implying 74% restriction. Plugging .74 in for the new value of x_C in Equation 7 and solving for S_T gives a slightly higher estimate of .36. Using the procedure in Stoolmiller (1998) based on the same CAP adoptee IQ data, a somewhat higher estimate of .55 is obtained for S_T.

For the MAP, the estimated correlation of IQ between unrelated adoptive siblings was .26, based on computing a weighted average of the correlations $-.03$, reported in Scarr and Weinberg's Table 5 (Scarr & Weinberg, 1978, p. 684) for 84 sibling pairs, and .44, reported in their Table 2 (Scarr & Weinberg, 1983, p. 263) for 140 sibling pairs. Because separate estimates for genetic and NE effects were not reported for the MAP, only the estimate of SE can be corrected. Applying Equation 9 gives an estimate of $S_T = .51$. Using the procedure in Stoolmiller (1998) based on the same MAP adoptee IQ data, an estimate of .60 is obtained for S_T.

As a final example, consider the median and weighted average correlations for unrelated adoptive siblings reared together of about .22 and .24, respectively, obtained by updating the data in Bouchard and McGue's (1981) review to include results from the adolescent MAP sample ($r = 0$, 84 pairs), the CAP ($r = .13$, 87 pairs), and the Danish adoption cohort ($r = 0$, 24 pairs; Teasdale & Owen, 1984). Applying Equation 9 gives estimates of about .46 and .49, respectively, for SE based on the median or weighted average correlation. In summary, the estimates of SE over all these examples are .51, .36, .51, and .46 for corrections based on direct assessments of FE and .55, .55, and .60 for corrections based solely on adoptive-sibling IQ distributions. In either case, the corrected estimates are substantially higher than the original estimates, and neither set supports the claim that shared FE is not important.

Adopted-Apart Twins and Siblings

Heritability estimates based on adopted-apart biological relatives, monozygotic (Mz) twins, regular siblings, or birth-mother adoptee pairs do not escape biases due to range restriction of SE. Under the same assumptions used for the model with unrelated adoptive siblings, the correlation between Mz twins, or twice the correlation between regular siblings or birth-mother adoptee pairs, is a direct estimate of the proportion of variance due to genes,

$$\frac{var(G_T)}{var(G_T) + var(E_T) + \frac{var(C_T)}{3}}. \quad (14)$$

Equation 14, however, is identical to Equation 10, because the denominator in Equation 14 is just $var(V_A)$, and thus, Equation 14 indicates that heritability estimates for related siblings adopted apart will also be seriously inflated by range restriction of SE

because the denominator is too small. This fact seems to be misunderstood in the behavior-genetics literature with respect to adopted-apart Mz twins. For example, Wachs (1991) raised this issue, although he did not explicitly label it as range restriction, with Plomin and Bergeman (1991). In their response to Wachs, Plomin and Bergeman claimed that estimates from adopted-apart Mz twins were valid so long as the rearing environments were uncorrelated. Equation 14 shows, however, that this is wrong. Correlations for birth-mother adoptee pairs may be just as seriously inflated, especially if birth mothers come from the same kind of middle-class families that the adoptee is placed with, as in the CAP and TAP. Bronfenbrenner (1986) is often cited as showing that correlations for Mz twins adopted apart into very different environments are much lower than the usual correlations reported for Mz twins adopted apart. Equation 14 shows that the elimination of range restriction of SE is probably responsible for the reduction.

So far, discussion of the effects of range restriction of SE on variance proportions in biometric models has been completely hypothetical. There are, however, a small number of adoption studies of IQ conducted by Schiff and his colleagues in France, who have attempted to obtain a more representative range of variability on at least a distal measure of FE, namely, SES. The results of Schiff's studies support the argument concerning the consequences of range restriction of SE. For example, Schiff and Lewontin (1986) found that the effect of adoption on IQs of low-SES children was very large (about 20 points as compared with biological siblings who remained with the birth mother). In comparisons with children raised in biological families with SES levels equivalent to the adoptive homes, there were no differences. Duyme (1988) extended Schiff's analysis to scholastic success and found evidence for FE effects, indexed by SES. Capron and Duyme (1989), in a full cross-fostering design based on SES, replicated the environmental effects of adoptive homes on IQ found by Schiff and also found evidence for genetic effects that were nearly equal in magnitude. The effects of SES in these studies are inconsistent with claims that SE is not important for IQ. As discussed previously, however, SES is probably not the decisive FE factor in influencing IQ, and thus, FE effects are probably still underestimated in these studies.

Summary

I have discussed in detail the range restriction of FE that is inherent in the adoption process and the impact it has on the interpretation of behavior-genetic adoption studies. I have argued that selection into adoption studies is by definition a between-family process and that thus, the SE component of variance in outcomes such as IQ and deviance is attenuated. This constitutes a serious problem because the restriction is extensive—about 67%, which implies only the top half of the FE distribution is represented in adoption studies—and because ample evidence from diverse sources indicates that FE quality has important causal effects on childhood IQ and deviance. In regressions of child outcomes on adoptive FE characteristics, range restriction of this magnitude cuts effect sizes and F statistics by a multiplicative factor of 3 and R^2 statistics by factors of 2 to 2.5. I have also argued that comparable range restriction to genetic and nonshared influences is unlikely, although direct evidence on this point is lacking. Taken together, these arguments imply that behavior-genetic adoption studies might seriously underestimate the impor-

tance of SE influences and overestimate the importance of NE and genetic influences. Corrections for range restriction applied to IQ data from recent adoption studies indicate that if adoptive families were representative of the U.S. census, SE could account for as much as 50% of the total phenotypic variance.

Across many adoption studies, two apparently contradictory sets of findings have emerged with regard to SE. The first set is that mean-level comparisons of children adopted into middle-class homes versus (a) their birth parents, (b) their biological siblings who remain with their birth mother, and (c) the population of children from the birth-mother's socioeconomic stratum consistently indicate that SE is a powerful influence on most important child outcomes. The second set is that within adoption samples, correlations of traits among adoptive siblings or among adoptive parents and adoptees are small, indicating that SE is not important. Range restriction of shared FE is a parsimonious and powerful explanation that integrates these otherwise confusing and apparently contradictory results. Range restriction of shared FE also explains why correlations for most Mz twins adopted apart are so high and suggests that environmental assessments are important for correct interpretation of adopted-apart twin studies. Range restriction of shared FE can also reconcile the discrepancy between randomized experiments that demonstrate family-based interventions have substantial and significant effects on both IQ and antisocial behavior and the apparent lack of importance of FE in individual-difference adoption studies.

It is sometimes claimed that twin and adoption designs converge on similar estimates for SE, NE, and genetic effects (Plomin et al., 1997). This might appear at first to be a problem for a range-restriction-of-adoptive-FE argument because range restriction of FE is not inherent in twin designs. As noted in the introduction, however, undetected nonadditive genetic effects can bias estimates of SE downward in twin designs. In fact, Plomin, Corley, Caspi, Fulker, and DeFries (1998) have advanced this argument to explain discrepancies between twin and adoption studies for self-reported personality dimensions. Thus, the apparent convergence of twin and adoption designs could be a consequence of different methodological problems operative in each kind of design. Alternatively, the conclusion that range restriction is not a problem for twin designs may be premature because of volunteer bias. Two twin studies are noteworthy, for example, because they involved samples that were apparently much less restricted than the usual volunteer sample. Scarr-Salapatek (1971) demonstrated with a population-based sample of twins from all schools in Philadelphia, which included a substantial number of twins from environmentally disadvantaged backgrounds, that the heritability of aptitude was much lower in the most disadvantaged third of the population than in the top two thirds. For the Louisville Twin Study (LTS), Wilson (1983) reported that a strong effort was made to recruit a representative sample from the Louisville area, and A. Metheny (personal communication, November 1998) graciously provided the mean and standard deviation for the HOME total score from the LTS (32.3 and 7.5, respectively, see Table 1), both of which are consistent with a broadly representative sample. More importantly, however, estimates of SE variance during early and middle childhood, .57, .64, .43, and .51 (Wilson, 1983, p. 305, Table 3, obtained by the moment-estimating formula $2r_{Dz} - r_{Mz}$), are roughly comparable to my estimates of SE variance based on correcting adoption studies. When twin studies do arrive at higher

estimates of SE, it is sometimes suggested that there is a special twin environmental effect (Plomin et al., 1994). The higher estimates, however, could be due to the fact that the twin studies in question have obtained a more representative sampling of FE, as appears to be the case with the LTS. A review of the existing twin literature for evidence of range restriction problems will probably be necessary before twin and adoption results can be meaningfully compared and conclusions drawn about convergence of results.

Finally, I do not question the conclusion that genes and NE have important effects on behavior in general and child outcomes in particular. Thus, I do not support the position that either is irrelevant to understanding IQ and deviance. Despite the inflation of the magnitude of their effects due to range restriction, the evidence for their importance is quite strong, and nothing in my analysis suggests that their effects are trivial. A strong current of thought, however, in contemporary behavior genetics holds that shared family influence is unimportant for nearly all traits of interest (Harris, 1995; Plomin et al., 1994; Rowe, 1994; Scarr, 1992). I believe this position has arisen largely because of both the respect that adoption studies command and the range restriction inherent in adoption studies.

Limitations and Directions for Future Research

As pointed out earlier, the corrected estimates for SE given in this article depend on a number of assumptions. One assumption, which is not readily verifiable, is that there is no restriction to genetic variance for child outcomes in adoption samples. Note, however, that all heritability calculations, whether corrected or not, rely on this same assumptions. Another assumption is that restriction of between-family variability in sibling environments does not affect the within-family variability in those same environments. As repeatedly advocated by Wachs (1983, 1987, 1996) and others, more direct assessment of the relevant sibling environments will be necessary to resolve this issue. The possibility that NE variability shrinks in better families is particularly important to investigate because a recent meta-analysis indicates that the search for specific intrafamilial sources of NE influence has been largely unsuccessful (Turkheimer & Waldron, 1998). It is crucial to determine whether this is an outcome of the restricted range of FEs in general. A number of recent works show that extrafamilial sources of economic stress can degrade FE quality, including parenting practices, which in turn affects child outcomes (Conger, Patterson, & Ge, 1995; Garrett, Ng'andu, & Ferron, 1994; Simmons, Whitbeck, Melby, & Wu, 1994). If these effects are more important at certain developmental stages than others—for example, family poverty during early childhood as opposed to adolescence (Duncan, Yeung, Brooks-Gunn, & Smith, 1998)—then change over time in FE might create substantial NE effects similar to the argument given by McCall (1983) for the case of IQ. Economically advantaged families, including adoptive families, would be less likely to be exposed to such adverse conditions, which could restrict the range of NE effects in such samples. If this kind of range restriction is substantial, then my analysis underestimates total environmentality (undercorrecting NE and overcorrecting SE) and overestimates heritability.[7]

---

[7] A detailed algebraic proof, similar to the one in this article, of this assertion is available from the author on request.

There are other possible explanations for lack of shared family influence in adoption samples. Two possibilities worth mentioning are selective placement and prenatal effects. First, selective placement based on phenotypic resemblance between birth and adoptive parents can inflate estimates of heritability and SE. The elimination of selective placement could be one reason estimates of SE are lower in some adoption studies than in others. Second, prenatal effects can also inflate heritability (for an example of significant prenatal alcohol effects on adoptee antisocial behavior, see Cadoret, Yates, Troughton, Woodworth, & Stewart, 1995) and NE effects. The prenatal environments of biologically unrelated children adopted into the same family would, in the absence of selective placement, most likely be completely uncorrelated because they are provided by unrelated birth mothers. Thus, prenatal effects would operate entirely as NE effects in adoption studies. For siblings who come from the same mother, prenatal effects are likely to have both SE and NE components (Devlin, Daniels, & Roeder, 1997). This might also contribute to lower estimates of SE effects in adoption studies.

The presence or absence of selective placement, prenatal effects, or both, however, does not rule out range restriction of the adoptive FE. All three could very well operate simultaneously and vary from measure to measure and study to study. Note, however, that selective placement was absent in the CAP (Plomin & DeFries, 1985), and the two estimates of SE for the CAP were .36 and .55 for direct and indirect corrections of range restriction, respectively, which are not much lower than corresponding estimates for the TAP and MAP. For all these reasons, I recommend treating my corrected estimates of SE as upper limits until more definitive data are available. The corrections represent an improvement over no corrections at all and indicate strongly that dismissing SE as unimportant for understanding child outcomes in the larger population of American families is premature.

Fortunately, the problems described for behavior-genetic studies with adoption samples are not insurmountable. Biases due to range restriction can be largely corrected. The details are beyond the scope of this paper, but techniques for models of sibling resemblance are available (Stoolmiller, 1998). For example, it has been hypothesized that the importance of SE for IQ declines in adolescence to zero. The strongest test of this hypothesis would be a longitudinal sibling adoption study, and, indeed, two have been recently reported (Loehlin, Horn, & Willerman, 1989; Weinberg, Scarr, & Waldman, 1992). Unfortunately for proponents of this view, in both cases, the direction of change was as hypothesized, but the amount of change in SE was small and nonsignificant. For the results of Loehlin et al. and Weinberg et al., respectively, the childhood and adolescent estimates of SE were .11 and −.09 and .31 and .19, neither of which were big enough drops to be significant with the available samples sizes. Corrections for range restriction, however, are important before drawing conclusions about change in SE from childhood to adolescence or the level of SE during adolescence. The observed correlations may be zero or even slightly negative precisely because of range restriction, and range restriction may even exaggerate whatever true change in SE exists. Thus, techniques for correcting range restriction hold the promise of producing more accurate estimates of genetic, shared, and nonshared influences from adoption samples, including estimates of hypothesized developmental change over

time. This should stimulate more interest in the adoption design as a powerful research tool for studying child development.

References

Allen, D. A., Affleck, G., McGrade, B. J., & McQueeny, M. (1981). Characteristics of the Home Observation for Measurement of the Environment Inventory in a sample of high-risk developmentally disabled infants. *Infant Behavior and Development, 6,* 53–60.

Arbuckle, J. L. (1996). Full information estimation in the presence of missing data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Erlbaum.

Bachrach, C. A. (1983). Adoption as a means of family formation: Data from the National Survey of Family Growth. *Journal of Marriage and the Family, 45,* 859–865.

Bachrach, C. A., London, K. A., & Maza, P. L. (1994). On the path to adoption: Adoption seeking in the United States, 1988. In R. P. Barth, J. D. Berrick, & N. Gilbert (Eds.), *Child welfare research review* (Vol. 1, pp. 250–272). New York: Columbia University Press.

Bachrach, C. A., Stolley, K. S., & London, K. A. (1992). Relinquishment of premarital births: Evidence from national survey data. *Family Planning Perspectives, 24,* 27–32.

Bates, J. E., Olson, S. L., Pettit, G. S., & Bayles, K. (1982). Dimensions of individuality in the mother–infant relationship at six months of age. *Child Development, 53,* 446–461.

Baumrind, D. (1993). The average expectable environment is not good enough: A response to Scarr. *Child Development, 64,* 1299–1317.

Bee, H. L., Barnard, K. E., Eyres, S. J., Gray, C. A., Hammond, M. A., Spietz, A. L., Snyder, C., & Blark, B. (1982). Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother–infant interaction. *Child Development, 53,* 1134–1156.

Benasich, A. A., & Brooks-Gunn, J. (1996). Maternal attitudes and knowledge of child-rearing: Associations with family and child outcomes. *Child Development, 67,* 1186–1205.

Boomsma, D. I. (1987). Absence or underestimation of shared environment? *Behavioral and Brain Sciences, 10,* 19–20.

Bouchard, T. J. J., & McGue, M. (1981, May). Familial studies of intelligence: A review. *Science, 212,* 1055–1059.

Bradley, R. H. (1994). The HOME inventory: Review and reflections. *Advances in Child Development and Behavior, 25,* 241–288.

Bradley, R. H., Caldwell, B. M., Ramey, C. T., Barnard, K. E., Gray, C., Hammond, M. A., Mitchell, S., Gottfried, A. W., Siegel, L., & Johnson, D. L. (1989). Home environment and cognitive development in the first 3 years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental Psychology, 25,* 217–235.

Brestan, E., & Eyberg, S. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies and 5,272 kids. *Journal of Clinical Child Psychology, 27,* 180–189.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology, 22,* 723–742.

Cadoret, R. J., Yates, W. R., Troughton, E., Woodworth, G., & Stewart, M. A. (1995). Genetic–environmental interaction in the genesis of aggressivity and conduct disorders. *Archives of General Psychiatry, 52,* 916–924.

Caldwell, B. M., & Bradley, R. M. (1978). *Home observation for measurement of the environment.* Little Rock: University of Arkansas.

Capaldi, D. M., Crosby, L., & Stoolmiller, M. (1996). Predicting the timing of first sexual intercourse for adolescent males. *Child Development, 67,* 344–359.

Capaldi, D. M., & Patterson, G. R. (1991). Relation of parental transitions

to boys' adjustment problems: 1. A linear hypothesis. II. Mothers at risk for transitions and unskilled parenting. *Developmental Psychology, 27,* 489–504.

Capron, C., & Duyme, M. (1989, August). Assessment of effects of socio-economic status on IQ in a full cross-fostering study. *Nature, 340,* 552–554.

Chamberlain, P. (1997, April). *The effectiveness of group versus family treatment settings for adolescent juvenile offenders.* Paper presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Chamberlain, P., & Reid, J. B. (1998). Comparison of two community alternatives to incarceration for chronic juvenile offenders. *Journal of Consulting and Clinical Psychology, 66,* 624–633.

Cherny, S. S., & Cardon, L. R. (1994). General cognitive ability. In J. C. DeFries, R. Plomin, & D. W. Fulker (Eds.), *The Colorado adoption project* (pp. 46–56). Cambridge, MA: Blackwell.

Cloninger, C. R., Sigvardsson, S., Bohman, M., & von Knorring, A.-L. (1982). Predisposition to petty criminality in Swedish Adoptees: II. Cross-fostering analysis of gene–environment interaction. *Archives of General Psychiatry, 39,* 1242–1247.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis in the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Coll, C. T. G., Offman, J., & Oh, W. (1987). The social ecology and early parenting of Caucasian adolescent mothers. *Child Development, 58,* 955–963.

Conger, R. D., Patterson, G. R., & Ge, X. (1995). It takes two to replicate: A mediational model for the impact of parents' stress on adolescent adjustment. *Child Development, 66,* 80–97.

Devlin, B., Daniels, M., & Roeder, K. (1997, July). The heritability of IQ. *Nature, 388,* 468–471.

Dolan, C. V., & Molenaar, P. C. M. (1994). Testing specific hypotheses concerning latent group differences in multi-group covariance structure analysis with structured means. *Multivariate Behavioral Research, 29,* 203–222.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review, 63,* 406–423.

Duyme, M. (1988). School success and social class: An adoption study. *Developmental Psychology, 24,* 203–209.

Eaves, L. J. (1988). Dominance alone is not enough. *Behavior Genetics, 18,* 27–33.

Fein, G. G., Gariboldi, A., & Boni, R. (1993). Antecedents of maternal separation anxiety. *Merrill-Palmer Quarterly, 39*(4), 481–495.

Fergusson, D. M., Lynskey, M., & Horwood, L. J. (1995). The adolescent outcomes of adoption: A 16-year longitudinal study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 597–615.

Garrett, P., Ng'andu, N., & Ferron, J. (1994). Poverty experiences of young children and the quality of their home environments. *Child Development, 65,* 331–345.

Gottfried, A. W., & Gottfried, A. E. (1984). Home environment and mental development in middle class children in the first three years. In A. W. Gottfried (Ed.), *Home environment and early cognitive development: Longitudinal research* (pp. 57–115). New York: Academic Press.

Grayson, D. A. (1989). Twins reared together: Minimizing shared environmental effects. *Behavior Genetics, 19,* 593–604.

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review, 102,* 458–489.

Hewitt, J. K. (1989). Of biases and more in the study of twins reared together: A reply to Grayson. *Behavior Genetics, 19,* 605–608.

Hoffman, L. W. (1991). The influence of the family environment on personality: Accounting for sibling differences. *Psychological Bulletin, 110,* 187–203.

Holder-Brown, L., Bradley, R. H., Whiteside, L., Brisby, J. A., & Parette, H. P., Jr. (1993). Using the HOME inventory with families of children

with orthopedic disabilities. *Journal of Developmental and Physical Disabilities, 5*(3), 181–201.

Hollenbeck, A. R. (1978). Early infant home environments: Validation of the home observation for measurement of the environment inventory. *Developmental Psychology, 14,* 416–418.

Horn, J. M., Green, M., Carney, R., & Erickson, M. T. (1975). Bias against genetic hypotheses in adoption studies. *Archives of General Psychiatry, 32,* 1365–1367.

Horn, J. M., Loehlin, J. C., & Willerman, L. (1979). Intellectual resemblance among adoptive and biological relatives: The Texas adoption project. *Behavior Genetics, 9,* 177–201.

Jackson, J. F. (1993). Human behavioral genetics, Scarr's theory, and her views on interventions: A critical review and commentary on their implications for African American children. *Child Development, 64,* 1318–1332.

Jensen, A. R. (1987). Unconfounding genetic and nonshared environmental effects. *Behavioral and Brain Sciences, 10,* 26–27.

Johnson, D. L., Swank, P., Howie, V. M., Baldwin, C. D., Owen, M., & Luttman, D. (1992). Does HOME add to the prediction of child intelligence over and above SES? *The Journal of Genetic Psychology, 154,* 33–40.

Jordan, T. E. (1978). Influences on vocabulary attainment: A five-year prospective study. *Child Development, 49,* 1096–1106.

Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin, 102,* 187–203.

Kazdin, A. E. (1994). Interventions for aggressive and antisocial children. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 341–382). Washington, DC: American Psychological Association.

Lamb, M. E., Hwang, C.-P., Bookstein, F. L., Broberg, A., Hult, G., & Frodi, M. (1988). Determinants of social competence in Swedish preschoolers. *Developmental Psychology, 24,* 58–70.

Leahy, A. (1935). Nature, nurture and intelligence. *Genetic Psychology, 17,* 241–306.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

Locurto, C. (1990). The malleability of IQ as judged from adoption studies. *Intelligence, 14,* 275–292.

Loehlin, J. C., Horn, J. M., & Willerman, L. (1989). Modeling IQ change: Evidence from the Texas Adoption Project. *Child Development, 60,* 993–1004.

Lozoff, B., Park, A. M., Radan, A. E., & Wolf, A. W. (1995). Using the HOME Inventory with infants in Costa Rica. *International Journal of Behavioral Development, 18,* 277–295.

Luster, T., & Dubow, E. (1992). Home environment and maternal intelligence as predictors of verbal intelligence: A comparison of preschool and school-age children. *Merrill-Palmer Quarterly, 38,* 151–175.

Luster, T., & Rhoades, K. (1989). The relation between child-rearing beliefs and the home environment in a sample of adolescent mothers. *Family Relations, 38,* 317–322.

Luster, T., Rhoades, K., & Haas, B. (1989). The relation between parental values and parenting behavior: A test of the Kohn hypothesis. *Journal of Marriage and the Family, 51,* 139–147.

Lykken, D. T., McGue, M., & Tellegen, A. (1987). Recruitment bias in twin research: The rule of two-thirds reconsidered. *Behavior Genetics, 17,* 343–362.

Lykken, D. T., Tellegen, A., & DeRubeis, R. (1978). Volunteer bias in twin research: The rule of two-thirds. *Social Biology, 25,* 1–9.

Martin, N. G., & Wilson, S. R. (1982). Bias in the estimation of heritability from truncated samples of twins. *Behavior Genetics, 12,* 467–472.

McCall, R. B. (1981). Nature–nurture and the two realms of development: A proposed integration with respect to mental development. *Child Development, 52,* 1–12.

McCall, R. B. (1983). Environmental effects on intelligence: The forgotten

realm of discontinuous nonshared within-family factors. *Child Development, 54,* 408–415.

Mednick, S., Gabrielli, W. F. J., & Hutchings, B. (1984, May). Genetic influences in criminal convictions: Evidence from an adoption cohort. *Science, 224,* 891–893.

Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29,* 177–185.

Metzl, M. N. (1980). Teaching parents a strategy for enhancing infant development. *Child Development, 51,* 583–586.

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture of human aggression. *Journal of Personality and Social Psychology, 72,* 207–217.

Morton, N. E. (1987). Shared environment and cultural inheritance. *Behavioral and Brain Sciences, 10,* 33–34.

Murray, A. D., Johnson, J., & Peters, J. (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language, 17,* 511–525.

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

Neale, M. C., Eaves, L. J., Kendler, K. S., & Hewitt, J. K. (1989). Bias in correlations from selected samples of relatives: The effects of soft selection. *Behavior Genetics, 19,* 163–169.

Palti, H., Otrakul, A., Belmaker, E., Tamir, D., & Tepper, D. (1984). Children's home environments: Comparison of a group exposed to a stimulation intervention program with controls. *Early Child Development and Care, 13,* 193–212.

Parks, P. L., & Smeriglio, V. L. (1986). Relationships among parenting knowledge, quality of stimulation in the home and infant development. *Family Relations, 35,* 411–416.

Pearson, J. S., & Amacher, P. L. (1954). Intelligence test results and observations of personality disorder among 3594 unwed mothers in Minnesota. *Journal of Clinical Psychology, 12,* 16–21.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on "environmental" measures. *Behavior and Brain Sciences, 14,* 373–386.

Plomin, R., Chipuer, H. M., & Neiderhiser, J. M. (1994). Behavioral genetic evidence for the importance of nonshared environment. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 1–31). Hillsdale, NJ: Erlbaum.

Plomin, R., Corley, R., Caspi, A., Fulker, D. W., & DeFries, J. (1998). Adoption results for self-reported personality: Evidence for nonadditive genetic effects. *Journal of Personality and Social Psychology, 75,* 211–218.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10,* 1–16.

Plomin, R., & DeFries, J. C. (1985). *Origins of individual differences in infancy.* New York: Academic Press.

Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent–offspring adoption study. *Psychological Science, 8,* 442–447.

Ramey, C. T., Farran, D. C., & Campbell, F. A. (1979). Predicting IQ from mother–infant interactions. *Child Development, 50,* 804–814.

Ramey, S. L., & Ramey, C. T. (1992). Early educational intervention with disadvantaged children—To what effect? *Applied and Preventative Psychology, 1,* 131–140.

Reis, J., Barbera-Stein, L., & Bennett, S. (1986). Ecological determinants of parenting. *Family Relations, 35,* 547–554.

Rhea, S.-A., & Corley, R. P. (1994). Applied issues. In J. C. DeFries, R. Plomin, & D. W. Fulker (Eds.), *Nature and nurture during middle childhood* (pp. 295–309). Cambridge, MA: Blackwell.

Rock, S. L., Head, D. N., Bradley, R. H., Whiteside, L., & Brisby, J. (1994). Use of the HOME inventory with families of young visually

impaired children. *Journal of Visual Impairment and Blindness* (March–April), 140–151.

Rose, R. J., & Kaprio, J. (1987). Shared experience and similarity of personality: Positive data from Finnish and American twins. *Behavior and Brain Sciences, 10,* 35–36.

Rose, S. J. (1986). *The American Profile Poster.* New York: Pantheon Books.

Rosenthal, R., & Rosnow, R. L. (1975). *The volunteer subject.* Cambridge, MA: Harvard University Press.

Ross, G. S. (1984). Home intervention for premature infants of low-income families. *American Journal of Orthopsychiatry, 54,* 263–270.

Rovine, M. (1994). Estimating nonshared environment using sibling discrepancy scores. In E. M. Hetherington, D. Reiss, & Plomin R. (Eds.), *Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 33–61). Hillsdale, NJ: Erlbaum.

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior.* New York: Guilford Press.

Sagi, A., Jaffe, M., Tirosh, E., Findler, L., & Harel, J. (1988). Maternal risk status and outcome measures: A three-stage study in Israel. *Child Psychology and Human Development, 19,* 145–157.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63,* 1–19.

Scarr, S. (1993). Biological and cultural diversity: The legacy of Darwin for development. *Child Development, 64,* 1333–1353.

Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review, 43,* 674–692.

Scarr, S., & Weinberg, R. A. (1983). The Minnesota adoption studies: Genetic differences and malleability. *Child Development, 54,* 260–267.

Scarr-Salapatek, S. (1971, December). Race, social class, and IQ. *Science, 174,* 1285–1295.

Schiff, M., & Lewontin, R. (1986). *Education and class.* Oxford, England: Oxford University Press.

Shonkoff, J. P., Hauser-Cram, P., Krauss, M. W., & Upshur, C. C. (1992). Development of infants with disabilities and their families: Implications for theory and service delivery. *Monographs of the Society for Research in Child Development, 57*(6), 1–163.

Simmons, R. L., Whitbeck, L. B., Melby, J. N., & Wu, C. (1994). Economic pressure and harsh parenting. In R. D. Conger & G. H. Elder, Jr. (Eds.), *Families in troubled times: Adapting to change in rural America* (pp. 3–19). New York: Aldine De Gruyter.

Stevens, J. H., Jr. (1988). Social support, locus of control, and parenting in three low-income groups of mothers: Black teenagers, Black adults, and White adults. *Child Development, 59,* 635–642.

Stoolmiller, M. (1998). Correcting estimates of shared environmental variance for range restriction in adoption studies using a truncated multivariate normal model. *Behavior Genetics, 28,* 429–441.

Teasdale, T. W., & Owen, D. R. (1984, June). Heredity and familial environment in intelligence and educational level: A sibling study. *Nature, 309,* 620–622.

Teasdale, T. W., & Sorensen, T. I. A. (1983). Educational attainment and social class in adoptees: Genetic and environmental contributions. *Journal of Biosocial Science, 15,* 509–518.

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

Turkheimer, E., & Waldron, M. C. (1998). *Nonshared environment: A theoretical, methodological, and quantitative review.* Manuscript submitted for publication.

Unger, D. G., & Wandersman, L. P. (1988). The relation of family and partner support to the adjustment of adolescent mothers. *Child Development, 59,* 1056–1060.

Wachs, T. D. (1983). The use and abuse of environment in behavior-genetic research. *Child Development, 54,* 396–407.

Wachs, T. D. (1987). The relevance of the concept of nonshared environ-

ment to the study of environmental influences: A paradigmatic shift or just some gears slipping? *Behavioral and Brain Sciences, 10,* 41–42.

Wachs, T. D. (1991). Nature and nurture: A shaky alliance. *Behavioral and Brain Sciences, 14,* 411–412.

Wachs, T. D. (1996). Known and potential processes underlying developmental trajectories in childhood and adolescence. *Developmental Psychology, 32,* 796–801.

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.

Weisglas-Kuperus, N., Baerts, W., Smrkovsky, M., & Sauer, P. J. J. (1993). Effects of biological and social factors on the cognitive development of very low birth weight children. *Pediatrics, 92,* 658–665.

Widmayer, S. M., Peterson, L. M., Larner, M., Carnahan, S., Calderon, A., Wingerd, J., & Marshall, R. (1990). Predictors of Haitian-American infant development at twelve months. *Child Development, 61,* 410–415.

Wilson, R. S. (1983). The Louisville twin study: Developmental synchronies in behavior. *Child Development, 54,* 298–316.

Wilson, R. S., & Matheny, A. P., Jr. (1983). Mental development: Family environment and genetic influences. *Intelligence, 7,* 195–215.

ORDER FORM

Start my 1999 subscription to *Psychological Bulletin!*

ISSN: 0033-2909

Send me a Free Sample Issue ❑

❑ Check Enclosed (make payable to APA)

___ $76.00, APA Member/Affiliate _____
___ $153.00, Individual Non-Member _____
___ $315.00, Institution _____
TOTAL AMOUNT ENCLOSED $ _____

Charge my: ❑ VISA ❑ MasterCard ❑ American Express
Cardholder Name _____
Card No. _____ Exp. date _____

Signature (Required for Charge)

Credit Card _____

Subscription orders must be prepaid. (Subscriptions are on a calendar basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.

Billing Address _____
City _____ State _____ Zip _____
Daytime Phone _____

SEND THIS ORDER FORM TO:
American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

AMERICAN
PSYCHOLOGICAL
ASSOCIATION

SHIP TO:
Name _____
Address _____
City _____ State _____ Zip _____
APA Customer # _____

Or call (800) 374-2721, fax (202) 336-5568.
TDD/TTY (202)336-6123. Email: subscriptions@apa.org

GAD99

PLEASE DO NOT REMOVE – A PHOTOCOPY MAY BE USED