TWIN STUDIES IN BEHAVIORAL RESEARCH: A SKEPTICAL VIEW

Leon J. Kamin

Department of Psychology
University of Cape Town
Rondebosch 7701, South Africa
e-mail: kamin@cybertours.com

Arthur S. Goldberger

Department of Economics
University of Wisconsin
Madison, Wisconsin 53706 USA
e-mail: asgoldbe@facstaff.wisc.edu

May 2001

ABSTRACT

   We review in detail two major ongoing research projects that employ samples of twins reared apart (and in one case, twins reared together). The studies attempt, via model-fitting, to estimate proportions of genetic and environmental variance for many human traits. We discuss problems concerning the representativeness of samples, the accuracy and reliability of the data, the extent of contact of nominally separated twins, the measurement of selective placement effects, and the particular model-fitting procedures. The two studies agree in their conclusions, but we do not find the conclusions to be convincing. We suggest that no scientific purpose is served by the flood of heritability estimates generated by these studies.

CONTENTS

TWIN STUDIES IN BEHAVIORAL RESEARCH: A SKEPTICAL VIEW

INTRODUCTION

A cascade of reports by behavior geneticists has poured out in recent years, purporting to calculate the heritability -- the genetic share of phenotypic variance -- of dozens of human traits. Readers have been informed, e.g., that the heritability of controllable life events is 53% among women and 14% among men, while the heritability of uncontrollable life events is 22% among women and zero among men (46). The heritability of inhibition of aggression is 12% (35), of openness to experience 40%, conscientiousness 29%, agreeableness 12% (1), of morningness-eveningness 54% (21). The heritability of right-wing authoritarianism is 50% (30), that of IQ 70% (4) or 81% (39), that of femininity 28%, and of responsibility 63% (6). Milk and soda intake are in part heritable, but not the intake of fruit juice or diet soda (11).

These numbers all derive from research designs based upon pairs of adult twins. The classical twin method compares the observed phenotypic correlations of raised-together monozygotic and same-sexed dizygotic twins (MZT's and DZT's). When, as is almost invariably the case, the MZT's resemble each other more than do the DZT's, the difference is attributed to the excess genetic similarity of the MZT's. Assuming that mating is at random, gene action is entirely additive, and environmental similarity is independent of zygosity, the MZT and DZT phenotypic correlations are modeled as

$$r_1 = h^2 + c^2 \qquad\qquad 1.$$

$$r_2 = (1/2)\ h^2 + c^2, \qquad\qquad 2.$$

where h^2 is the genetic component and c^2 is the shared environmental component. Then doubling (r_1-r_2), the difference between those two correlations, is said to estimate heritability. Such an inference rests squarely upon the "equal environments assumption" -- the assumption that MZ's have shared no more relevant environmental experiences than have DZ's. This assumption is at best questionable (22, 24); MZ's have long been known to

experience environments that are in very many respects more similar than those of DZ's, and the etiologically relevant environmental variables for many traits cannot be readily specified.

The uncertainties surrounding the equal environments assumption are presumptively avoided by the study of separated twin pairs, who have been reared apart since infancy. Assuming that mating is at random, that placement of pair members into rearing homes has resulted in independent environments, and again that all gene action is additive, the phenotypic correlations for separated monozygotic twins (MZA's) and separated dizygotic twins (DZA's) are modeled as

$$r_3 = h^2 \qquad\qquad 3.$$
$$r_4 = (1/2)\ h^2 \qquad\qquad 4.$$

Then r_3, the MZA correlation observed for a trait, directly estimates the heritability of the trait. Doubling r_4, the observed correlation for DZA's, provides another estimate, and doubling (r_3-r_4), the difference between the MZA and DZA correlations, provides yet another. The shared environmental component is estimable by (r_1-r_3) or by (r_2-r_4). When data on MZT, DZT, MZA, and DZA pairs are all available, more elaborate modeling is possible. Nonadditive gene action would reduce the DZ factor from 1/2, positive assortative mating would increase it.

Twins are necessarily of the same age, and MZ's and same-sexed DZ's are of the same sex. So any age and sex effects on a trait will inflate their correlations. It is standard practice to remove such effects by statistical adjustment of the trait scores. Or, if the age-sex distributions of the MZ's and the DZ's coincide, then the double-the-difference estimators might automatically offset such inflation.

We focus in our review on two major ongoing projects that utilize data from samples of separated twins, sometimes combined with data from twins who have been reared together. Each of the projects has produced dozens of

research reports, spanning a wide range of disciplines. The studies are representative of the best that modern human behavior genetics has to offer.

The Minnesota Study of Twins Reared Apart (MISTRA), apart from producing a copious technical literature, has had popular impact via world-wide newspaper, magazine, and television accounts. Its design has the virtue of simplicity. Identical twins separated early in life are assumed to have had uncorrelated environments; thus their behavioral and psychological similarities in adulthood are assumed to be attributable to their genetic identity. With the further assumptions that the twins represent the range of genetic and environmental variation in the general population, the correlation of MZA's for any trait is taken as an estimate of the heritability of the trait in the population. In addition to MZA's, MISTRA has utilized a control group of DZA's.

The Swedish Adoption/Twin Study of Aging (SATSA), despite its advantages both in design and procedure, has not received the media attention of MISTRA. SATSA drew its twin subjects from a national Twin Registry. First, respondents to a mailed questionnaire who indicated that they had been separated before the age of eleven were recruited. Then a control sample of reared-together twins was drawn, matched to the separated twins for sex, age, and presumed zygosity (34). This four-group design, as noted above, allows for several estimates of heritability, in addition to the simple estimate based on the MZA r itself. Indeed, SATSA's modeling of the four groups jointly attempts to estimate shared and nonshared components of environmental variance, to capture selective placement, and also to distinguish between additive and nonadditive gene action.

Both studies are longitudinal; subjects are assessed on repeated occasions. Both studies analyze a wide range of traits; some are cognitive, others concern personality and temperament, and still others tap recollections of childhood environment. The heritability of a trait may vary with age, and

alternative indicators of an underlying concept -- "cognitive ability," or "sociability," e.g. -- need not produce the same heritability estimate. As measurement occasions and indices proliferate, so too do the research reports. Several traits may be modeled jointly to detect overlap in their genetic determinants. With so many occasions, indicators, and combinations available to choose from, it is likely that some of the studies' published conclusions, supported by conventional significance testing, are in fact reflections of chance fluctuations.

The two studies arrive at similar conclusions. Each finds substantial heritability for virtually all traits examined, with heritability higher for cognitive than for personality measures. For separated twins, measures of frequency of contact are unrelated to within-pair similarity, and other measures of rearing environment fail to detect substantial similarities. When environmental effects on a trait are detected, they are almost entirely assignable to unique experiences, those not shared by family members.

We now consider the studies in turn, raising similar questions about each. How representative, of what population, are the samples included in the study? How accurate and reliable are the data? How extensive has the separation of the nominally separated twins in fact been? How adequately has the possibility of selective placement been taken account of? How consistent with standard statistical practice have the procedures used been?

THE MISTRA STUDY

MISTRA has concluded that the heritability of IQ is approximately 70% (4). The heritabilities of various personality traits have been estimated in the range of about 30% to 50%, with some said to display substantial nonadditive genetic variance (6). The effect of environment on personality traits was asserted to be almost entirely that of nonshared experiences (7), a conclusion derived from a comparison of MISTRA's own separated-twin data with data on reared-together twins from other studies.

MISTRA's Sample

To avoid the criticism that previous MZA samples had been biased by a failure to include dissimilar MZA pairs, MISTRA "vigorously" recruited "all reared-apart twins, regardless of known or presumed zygosity and similarity" (4). Thus, "we are not subject to that bias" (5). But this effort could not and did not avoid a bias toward location and inclusion of pairs who knew of each other's existence and who had already been in some contact. Many of the earliest pairs "were self-recruited, attracted by reports of reunited twins appearing in the press" (14). Others were referred by friends, relatives, and adoption workers. Media accounts of the study invariably stressed an almost eerie similarity of the twins, despite a presumed total lack of contact. The publicity was "essential to our success in recruitment" (27). The media blitz surrounding MISTRA may well have pushed volunteers toward exaggerating both the degree of their separation and the similarity of their subsequent life experiences.

In any event, despite MISTRA's vigorous effort to recruit pairs regardless of zygosity and similarity, the early recruits were overwhelmingly identical twins. The numbers of MZ and same-sexed DZ pairs are approximately equal in the population, but between its 1979 inception and 1981, MISTRA obtained 30 MZA and only 9 same-sexed DZA pairs (27). When separately reared twins are recruited from population Twin Registries, same-sexed DZA's outnumber MZA's (23, 37). That imbalance has been regarded as a consequence of greater parental reluctance to separate identical twins (37). Clearly, MISTRA initially attracted disproportionate numbers of highly similar twin pairs. MISTRA's MZ correlations for personality traits are considerably higher than those obtained in Scandinavian studies. The difference has been attributed by SATSA authors (38) to recruitment procedures; the Scandinavian samples were drawn from population Twin Registries.

The disproportionate representation of MZA's in MISTRA has been reduced in recent years, but not eliminated. A decision to include opposite-sexed pairs expanded the DZA sample. By 1996 MISTRA was reporting on 65 MZA pairs, 38 same-sexed DZA pairs, and 16 opposite-sexed DZA pairs (12). Most pairs come from the United States, but 40% are from the United Kingdom, and others from Australia, Canada, China, New Zealand, Sweden, and Germany. So it is unclear to what population(s) heritability estimates calculated from MISTRA's haphazard samples can be generalized.

Separation and Contact

MISTRA's information on age of separation, degree of contact, and similarity of experience is gathered in the course of a "Life History Interview," each twin being questioned separately. The coded data used in analyses combine the information thus obtained into a single variable, total contact time for the pair. McCourt et al (30) say that "Any discrepancies are resolved in the direction of later separation, earlier reunion, and more contact," but provide no information about the frequency and magnitude of discrepancies. Some MISTRA papers indicate that pairs with periodic contact during childhood were assigned an age-at-separation score of zero, but TJ Bouchard (personal communication, Nov. 8, 1995) states that, on checking the data, this was not the case. He wrote that "We must have coded this way at one time and then decided that it would be misleading."

MISTRA's measure of total contact time credited 20 weeks of contact to twins who "met for a week at Christmas and for a week in the summer each year over a 10-year period" (6). The same 20 weeks of contact would have been credited to a pair who had been separated at the age of 20 weeks and never had further contact. Total contact time, thus measured, ranged between 3 weeks and 24 years. Can so crude a measure capture the extent of a pair's influence on each other, or validly index the effects of contact on similarity of experience?

Beyond the inevitable inaccuracies of retrospective data, MISTRA is vulnerable to another source of unreliability. Watson (48) writes of an MZA pair: "At the end of their week in Minneapolis one final coincidence emerged. Both had told Bouchard the same lie. 'We both said we wanted to be opera singers and neither of us can sing a note,' Barbara confessed. And they both broke into peals of laughter, yet again." This anecdote, intended to emphasize the similar behavior of MZA's, makes it clear that twins could and did lie about themselves to the investigators. Given the nature of the publicity surrounding the project, whatever shading or prevarication occurred would have tended to exaggerate the extent of a pair's separation and of their current similarities. In another context, to explain an unexpected finding, Eckert et al (13) referred to "a tendency for twins reared together to compare notes when asked about developmental milestones." Perhaps twins reared apart, who come to Minneapolis for a week-long assessment, also compare notes?

Apart from actual contact, selective placement may result in members of a twin pair being reared in highly similar, albeit separate, environments. For example, some of MISTRA's MZA pairs have been reared in related branches of the same biological family, others in homes of very similar social status. If selective placement has occurred, a shared environmental term should be added to equation 3. To assess this possibility, MISTRA calculated correlations for variables describing the homes in which pair members were reared. These "placement correlations" for standard demographic variables (parental socioeconomic status (SES), education, etc.), while positive, tended to be small, and the placement variables were only weakly associated with cognitive and personality measures of the twins. MISTRA thus infers that any contribution of selective placement to twin similarity must be trivial. But this neglects the fact that the twins come from several countries, with differing educational systems and SES distributions. Their birth cohorts also differ. So there is reason to doubt that MISTRA's demographic variables can

serve individually as valid indices of the effective similarity of rearing environments. In situations where multiple indicators of an underlying factor are available, social scientists frequently construct a combination of the indicators. For example, canonical correlation analysis of the placement variables might pick up resemblance that is not apparent from simple correlations. While MISTRA researchers have used multiple-indicator modeling for another purpose (20), we have not seen them using it to assess environmental similarity.

MISTRA also attempts to measure more subtle and detailed aspects of the home environment. The twins answered a "Physical Facilities in Childhood Environment Questionnaire" that required a yes or no reply to whether each of 41 items was present in the childhood home (31). Factor analysis of the replies produced four factors, labeled Material Possessions, Scientific/Technical, Cultural, and Mechanical. Twin correlations on each factor score were calculated; two were significant. When IQ scores were related to the factor scores, again two correlations were significant (4). Curiously, one of the two (Cultural) was negatively correlated with IQ scores. MISTRA infers from all this that selective placement was effectively inoperative. But closer inspection of the composition of the four factors would make one hesitant to view them as psychologically meaningful dimensions of the home environment. Presence of a pet dog or cat was classified in the Mechanical factor, but other animal pets fell into the Cultural factor. Having five or more magazine subscriptions, or a foreign cookbook, fell into Material Possessions, but a library of more than 200 books was Scientific/Technical, and a world atlas was Cultural. A sewing machine, and a flower or vegetable garden, were each Cultural; farm equipment was Scientific/Technical. A photographic darkroom was Scientific/Technical, but photographic equipment was Cultural. The relevance of these "factors" to a meaningful concept of selective placement is dubious, and although the factor scores are indeed

combinations of individual variables, no attempt was made to find the variable combination that best predicts the phenotypes.

Given the limited reliability and validity of its measures of contact and separation, MISTRA's failure to find significant correlations of those measures with the pairwise similarity of MZA's is not surprising. However, the measures did discriminate between the MZA and DZA samples. The mean age at separation, the time apart until first reunion, and total contact time all differed significantly between the two groups. MZA's had been separated at an earlier age, had spent less time apart before their first reunion, and had twice as much total contact time (12). Most of this extra contact time of MZA's occurred after their first reunion, and the time between first reunion and being tested by MISTRA was longer for the MZA's. There was thus more opportunity for mutual interaction between MZA than between DZA pairs. This confounds the attempt to estimate heritabilities by comparing MZA and DZA r's. (The earlier age at separation of MZA's seems anomalous, but a similar result was noted in the SATSA study; see below).

Bouchard et al (5) asserted that "There is no reason to believe that DZ twins separated early in life are likely to be placed in homes less similar than those of MZ twins." There may be some evidence to the contrary. Hur & Bouchard (20) submitted results of two questionnaires dealing with childhood family environment to a factor analysis. For one factor, labeled "Support," the average MZA and DZA r's were .41 and -.01, respectively. These results might be interpreted as indicating that selective placement was substantially greater for MZA's than for DZA's, leading to more similarity of outcomes for the MZA's. Hur & Bouchard, however, take another tack: if the self-reports are accurate, then they reflect distinct families responding similarly to identical DNA -- and if they are inaccurate, then they reflect genetically-based biases in perception. In either case, "there are important biological foundations for individual differences in perceptions of family environments

... retrospective perception can be influenced by genetic differences". The same analysis produced a second factor, labeled "Organizational and Cultural," for which the MZA and DZA r's averaged .05 and .06. Why should heredity influence perceptions of "Support" but not of "Organizational and Cultural"?

Cognitive Ability

For cognitive measures, MISTRA researchers have been slow to publish DZA r's despite their emphasis on the importance of DZA's as a control group. For the Wechsler IQ test -- one of the standard intelligence tests -- Bouchard et al (4) reported an MZA r of .69 for 48 pairs, suggesting a heritability of approximately 70%. They indicated that space limitations and the smaller size of the then-available DZA sample (30 pairs) led them to focus on the MZA sample alone. The first published DZA r for Wechsler IQ was for a subsample, comprised of those twins who had taken both the Wechsler test and an ego development test (33). These r's were .75 for 35 MZA pairs, and .47 for 26 DZA pairs. For a measure of "Verbal Reasoning", the MZA and DZA r's were .46 and .53. Newman et al (33) calculated from these data that the heritability of IQ was 76%, and that of Verbal Reasoning was 65%. It may seem surprising to see heritability of 65% reported for a measure on which the fraternal twin correlation, r_4, is larger than the identical twin correlation, r_3. But the Newman et al estimator is indeed sensible, being a weighted average of r_3 and $2r_4$. An alternative estimator, suggested by the logic of using DZA's as a control group, is $2(r_3-r_4)$; that would have given heritabilities of 56% for IQ and -14% (sic) for Verbal Reasoning.

McCourt et al (30) reported correlations for "general cognitive ability" (GCA) for a subsample of 38 DZA pairs and 39 MZA pairs who had also taken a questionnaire measuring "right-wing authoritarianism" (RWA). The GCA measure was derived from the first principal component (FPC) of MISTRA's battery of 29 brief tests of "special mental abilities." The MZA and DZA r's were .74 and .53. McCourt et al used the GCA measure merely as a control for possible

effects of cognitive ability on RWA; heritability of GCA itself was not estimated. Doubling the difference between MZA and DZA r's would suggest a heritability of 42%. Bouchard et al (4), asserting IQ heritability to be approximately 70%, had reported an MZA r of .78 for the FPC, based on 43 pairs, but gave no DZA figure. McCourt et al wrote that they "use the term general cognitive ability (GCA) in place of intelligence because formal intelligence tests scores were not used in this study." They did indicate that their GCA score correlated .74 with full-scale scores on the Wechsler, and "was thus considered to provide a valid measure of general cognitive ability", but they neither provided MZA or DZA r's for the Wechsler test itself, nor indicated why they used the GCA rather than the Wechsler.

Bouchard (2) reported that "more than 80" pairs of MZA's and "almost 60" pairs of DZA's had been studied by MISTRA. Responding to a query from one of us, Bouchard wrote (personal communication, Oct. 29, 1997) "I can't pass on the IQ results for our MZA's or DZA's because I have not published them yet. Indeed I have not even calculated them." This dilatoriness does not increase our confidence in MISTRA's claims; it stands in contrast to their treatment of personality data. DiLalla et al (12) had in 1996 published and analyzed in detail personality r's based on 65 MZA and 54 DZA pairs. As of our writing, the only DZA Wechsler IQ r released by MISTRA is based on the subsample of 26 pairs who also took the ego development test.

The results for another MISTRA cognitive test, a computer-administered "British IQ" (27), have been even more scantily reported. The MZA r for 42 pairs was .78 (4); no DZA r was reported. Lykken (27) indicated that for twins reared together the MZT and DZT r's were .78 and .14. He attributed the extremely low DZT r to "emergenesis," a conveniently hypothesized extreme form of nonadditive genetic variance. Perhaps additional evidence for emergenesis will appear when MISTRA reports a DZA r for its "British IQ".

A third cognitive score employed by MISTRA is the FPC of a battery of 28 subtests of "special mental abilities." (This measure is very similar to the GCA measure utilized by McCourt et al (30), but the latter was derived from a larger sample that included spouses and relatives of the twins). The r for 43 MZA pairs was .78; no DZA r was reported (4). However, MZA and DZA r's have been reported for four factors derived from 15 of the subtests in the battery: Verbal Reasoning, Spatial Ability, Perceptual Speed and Accuracy, and Visual Memory (31). The mean r's for the 5 subtests in the Verbal Reasoning factor were .44 for MZA's and .49 for DZA's. The authors wrote cautiously: "the MZA twins were no more similar than were the DZA twins on subtests from the Verbal Reasoning cluster ... the relatively large DZA correlations have been accounted for by statistical variability."

Bouchard et al (8), examining all 28 subtests, replaced the four computer-derived factors by four quite similarly named "domains." For the 6 subtests in their Verbal Ability domain the mean r's were .51 for MZA's and .37 for DZA's; the anomaly of r_4 greater than r_3 had disappeared. However, in the course of substituting intuitively constituted "domains" for the mechanical results of factor analysis, the authors reassign subtests from factors to domains, with some subtests assigned to a catch-all category labeled "Other." Four of the five subtests that the 1989 factor analysis had assigned to Verbal Reasoning were declared to be "Other" in 1990; for each of them, r_4 had exceeded r_3. The only 1989 Verbal Reasoning subtest retained in the 1990's Verbal Ability domain was the only one for which r_4 had been less than r_3.

Even if one accepts the reassignments, comparison of the MZA and DZA r's indicated at best a relatively modest heritability for the "special mental abilities." Bouchard et al (8), noting "the relatively large DZA correlations," wrote that the data "highlight the important role of shared environmental factors as determinants of special mental abilities." For

general intelligence, MISTRA had reported "an absence of any significant effect due to SES or other environmental measures..." (4).

Model-fitting

MISTRA's model and estimation are actually framed in terms of variances and covariances rather than correlations. Our presentation runs in terms of correlation coefficients for convenience. With that understanding, a sample size-weighted average of r_3 and $2r_4$ does provide an optimal estimate of heritability. In MISTRA's hands, this method detects substantial heritability under an astonishing variety of conditions. Table 1 presents the observed MZA and DZA r's for several traits, together with MISTRA's estimates of heritability (h^2).

INSERT TABLE 1 HERE

The heritability estimates hover around 50%, regardless of whether the DZA r is larger than, equal to, or smaller than the MZA r. The DZA r may be literally zero, or it may be fully equal to the genetic correlation of DZ twins; no matter, the heritability estimates are almost invariant. In standard statistical practice, persistent substantial discrepancies between r_3 and $2r_4$ might be viewed as evidence against the model specification. But for MISTRA, such evidence is generally dismissed by ad hoc appeals to nonadditive genetic variance, to the effects of assortative mating, and/or to an assertedly small DZA sample size.

THE SATSA STUDY

SATSA's approach is distinctive in that it systematically exploits data from the four twin groups (MZT, DZT, MZA, DZA) obtained in a single study. For cognitive ability, SATSA's estimate of 81% heritability (39) is a trifle higher than MISTRA's, while for personality traits its estimates tend to be a little lower (7). SATSA finds much of the genetic variance for both cognitive and personality traits to be nonadditive, and, like MISTRA, concludes that effects of selective placement and of post-separation contact are negligible.

When effects of environment do appear, they are generally attributable to nonshared environment. That MISTRA's and SATSA's conclusions generally conform, despite the difference in their populations and approaches, might be construed as mutual reinforcement.

SATSA's Sample

SATSA began by recruiting from the Swedish Twin Registry almost 1,000 pairs of same-sexed twins who had indicated, in response to a mailed questionnaire, that they had been separated before the age of eleven. Zygosity was at first determined by the response to a questionnaire item about physical similarity, and later checked by blood testing. Then a control sample of reared-together twins drawn from the Registry was matched to the reared-apart twins on sex, age, and presumed zygosity (34).

Separation and Contact

The core SATSA sample consists of the 758 pairs who returned a detailed 1984 questionnaire that included a number of personality scales. Despite the initial matching, twins reared apart (TRA's) were less likely than those reared together (TRT's) to return the mailed questionnaire. The core sample contains 351 TRA's and 407 TRT's (34, 37). No information has been provided about demographic or other differences between responding and nonresponding pairs. Perhaps the pairs who did respond, and are thus included in the SATSA study, were in closer contact, and thus able to make a joint decision about whether to respond. TRT's are presumably in closer contact than TRA's (12), which could account for their greater response rate. Under this reasoning, those TRA's with the most contact, those who may correlate highest in phenotypes, predominate in the sample. If so, contrasts between TRT and TRA r's would understate the effect of shared environment.

Cederlof et al (9) had also used the Swedish Twin Registry, obtaining a different sample of TRT's aged 35 to 75, all residing in southwestern Sweden. They found that, even within this limited geographical area, MZ twins are more

likely than DZ's to be living close to one another in adulthood. So some part
of the greater phenotypic similarity of MZT's over DZT's may be attributable
to more similar environments; and if MZA's are in more contact than DZA's, as
was the case in MISTRA, the same confounding of environmental and genetic
differences occurs among TRA's. When the data are derived from mailed
questionnaires, the closeness of residence and frequency of contact of MZ's
suggests another source of bias: collaboration on answers. Cederlof & Lorich
(10), discussing physical symptoms reported by twins in the Registry,
suggested that MZ's "may tend to report symptoms more concordantly than is
really the truth."

For information about the twins' backgrounds, including the age,
duration, and extent of their separation, SATSA relied on the twins' self-
reports, largely gathered through mailed questionnaires. Many of the twins in
the core sample were quite elderly; the mean age was 58.6 with 20% over age
70. There is evidence that the inconsistency of self-reports increases with
age. For the entire Twin Registry a tabulation is given, by birth cohorts, of
the reported age at separation (34). Disagreement of pair members by more than
two years increased smoothly from 11% among those born after 1950 to 26% among
those born before 1900. For the core sample itself, SATSA gives no information
about discrepancies in reported age of separation, or about how such
discrepancies were handled. Presumably some pairs also disagreed when
reporting age at first contact after separation, frequency of contact, etc.,
but again no relevant data have been provided.

SATSA "established a minimum requirement of separation *for purposes of
rearing* prior to 10 years of age" [emphasis in original](29). For the core
sample, the average age at separation was 2.8 years, with 48% separated during
the first year of life and 82% by their fifth birthday (44). It might seem
that the earlier the age at separation, the less correlated the twins'
environments will have been. But SATSA analyses indicate at best only weak

relations between age at separation and behavioral similarity, and the relations that do occur are sometimes in a counterintuitive direction. This, however, should be interpreted in the light of an observation made by Pedersen et al (36): "some of the earliest separated twins were neighbours and reared by related individuals, whereas some of the later separated twins seldom had contact after separation." In fact, 44% of the separated twin pairs in the Registry had been reared by biologically related families (34). The most common pattern was for the biological mother to rear one twin, while her sibling or parents reared the other. Twins reared in related families doubtless experience more similar environments than do those reared in unrelated families. The positive association between early separation and being reared in related families thus dilutes the observed association between early separation and behavioral similarity. The reported age at separation tells us very little -- it is in fact misleading – about the degree of actual separation of the twins. For such information we turn to other measures reported by SATSA.

The relevant data have been reported only for the combined sample of TRA's, not for MZA's and DZA's separately. The mean age at separation was 2.8 years. The number of years between age at nominal separation and first subsequent contact ranged from less than 1 year to more than 70 years, and averaged 10.9 (36). So on average the twins were back in contact at age 13.7, with many having resumed contact at a much earlier age: "the majority of twins are separated at an early age and relatively fewer were separated for the greater portion of their lifespan." The average age at testing was 58.6, implying that an average of 44.9 years had elapsed between re-establishment of contact and testing. These precise numerical values are derived from the fallible recollections of elderly twins, but it is obvious that separation was far from complete, with ample opportunity for contact and for shared environmental influences. Of course, one cannot expect observational studies

to meet the strict criteria of controlled experiments, but it is not unreasonable perfectionism to ask for thorough exploration of possible biases due to nonrandom assignment to environmental influences.

To assess whether the extent of separation was related to the twins' behavioral similarity, SATSA employed three measures: age at separation, years separated, and an index of "degree of separation" (36). This index, based on the twins' responses to a questionnaire, summed the scores for eight items, each coded 0 or 1, with 1 indicating more separation. The index had some peculiar properties. A pair who attended the same school, spent every weekend together, and vacationed together for more than 10 weeks every year would receive a summed score of 0 for those three items, appropriately indicating a very low degree of separation. But another pair, who had met only once in their lifetime, exchanged two letters a year, and spoke once a year on the phone would receive the same summed score of 0 for those three items.

We have found no analysis, in any SATSA paper, of the relation between any single item of the scale and the twins' behavioral similarity. The index might mask significant effects of single items; having been reared by related families, scored 0 or 1, is just one of the eight scale items. There is no report about possible differences between MZA's and DZA's for the scale as a whole, or for any scale item. Such information is relevant to the question of whether selective placement was greater for MZA's. We know that parents were more reluctant to separate MZA's (37); they may well have placed MZA's into more similar environments (e.g., the home of a neighboring relative) than DZA's. SATSA reports low correlations of its three separation measures with behavioral similarity, but apparently has made no attempt to construct a combined index that maximally relates to behavioral similarity.

Even if adequately separated, twins may have been reared in objectively similar homes; i.e., they may have been selectively placed. For SATSA's data on SES, we draw on a personal communication from GE McClearn (Oct. 8, 1997).

For TRA's, the r's for highest occupational level of a rearing parent were .27 for MZA's and .27 for DZA's, suggesting quite limited resemblance of childhood environments, and no difference between MZA's and DZA's. But for TRT's, the r's were .74 for MZT's and .51 for DZT's. Now these TRT pairs had been reared together, so they are describing a single set of parents, their own. Yet their level of agreement is hardly high, and the MZT r is higher than the DZT r. Elsewhere SATSA interprets higher MZ than DZ correlations of self-reports of childhood environment as evidence of genetic influence on retrospective perception (42). If that is the case, can such self-reports also serve as measures of objective environmental similarity?

Model-fitting

SATSA's main model for heritability estimation extends the one sketched earlier by splitting the genetic component h^2 into two parts. For MZT, DZT, MZA, DZA in turn, the phenotypic correlations are modeled as

$$r_1 = a^2 + d^2 + c^2 \qquad\qquad 5.$$

$$r_2 = (1/2)\ a^2 + (1/4)\ d^2 + c^2 \qquad\qquad 6.$$

$$r_3 = a^2 + d^2 \qquad\qquad 7.$$

$$r_4 = (1/2)\ a^2 + (1/4)\ d^2 \qquad\qquad 8.$$

Here a^2 is the additive genetic component, d^2 is the nonadditive ("dominance") genetic component, and c^2 is again the shared environmental component. While the simpler model expected the reared-apart ratio, r_4/r_3, to be 1/2, the extended model expects the ratio to be in the interval [1/4, 1/2], falling near the lower or upper end of the interval according as the nonadditive or additive component predominates. The reared-together ratio, r_2/r_1, might fall in the same interval, but would exceed 1/2 to the extent c^2 is nontrivial. Comparing TRT and TRA r's again estimates c^2 as (r_1-r_3) or (r_2-r_4). (SATSA's model and estimation are actually framed in terms of variances and covariances; our presentation again runs in terms of correlation coefficients for convenience). Once nonadditive genetic variance is present, the double-

the-difference rules no longer suffice to estimate heritability, although r_3 itself still does. Various estimates of a^2 and d^2 are obtainable from subtler contrasts. SATSA's formal estimation procedure in effect averages the available estimates in a statistically efficient manner. In view of their avowed reliance on formal statistical methodology, it is astonishing to observe that SATSA researchers hardly ever provide standard errors or confidence intervals for the prime targets of their investigations, the variance components.

Having adopted a tightly structured framework, the SATSA group might be credited by some for providing a coherent set of analyses that do not rely on the ad-hockery that is available to those social and behavioral scientists who are uninformed by genetic theory. But in practice, across a wide spectrum of psychological traits, SATSA typically reduces the model of equations 5.-8. by dropping one or more of the components. That happens when fitting the full model to the observed correlations would generate negative estimates of a^2, d^2, and/or c^2. In particular, it is very rare to find both a^2 and d^2 estimated for the same trait. Almost inevitably, SATSA reports either that all the genetic variance is additive, or that it is all nonadditive. For situations where observed rDZ/rMZ ratios run much higher than 1/2, SATSA modelers rely on a fallback position, adopting a variant of equations 5.-8. in which all the terms in d^2 are replaced by a common term, s^2 say, referred to as the "selective placement" (or, confusingly, the "correlated environments") component. But the possibility that environmental similarity differs between MZT's and DZT's is never entertained within the modeling.

For many personality traits, the rDZ/rMZ ratios run very low. For example, for emotionality-fear (44) this occurs for both TRT's and TRA's: the observed r's were MZT .49, DZT .08, MZA .37, DZA .04. The explanation offered is that "nonadditive genetic variance is important." Model-fitting concluded that the heritability of emotionality-fear was 39%, all nonadditive. In the

same study, the observed r's for sociability were MZT .35, DZT .19, MZA .20, DZA .19, and its heritability was estimated at 24%, all additive. Now there is no psychological reason for gene action to be entirely nonadditive for emotionality-fear, and entirely additive for sociability. An alternative interpretation of the pattern of correlations would be that MZ's experienced more similar environments than DZ's for one trait, but not the other. Admittedly, there is no psychological reason why environmental effects should operate in this manner, but the same reservation applies to the rationalization in terms of additivity of gene action.

When confronted by refractory data, SATSA investigators occasionally recognize the relevance of extra environmental resemblance for identical twins. The recognition is hedged by circumlocutions. In an analysis of men's occupations (26) the MZT r of .82 was much larger than those for the remaining groups (DZT .36, MZA .44, DZA .44). The authors wrote, "This pattern could be interpreted as a special MZT-effect." Renaming an environmental effect as a special effect does not change its essential nature. Even though $r_3 = r_4$, modeling concluded that the heritability was 60%.

The "special MZT-effect" sometimes goes under another name. Pedersen et al (37) described "an assimilation effect for MZT that inflates their similarity and is misread as nonadditive genetic variance." They suggested that "Data from twins reared apart can help in disentangling violations of the unequal (sic) environments assumption for twins reared together from nonadditive genetic variance." The logic is that if the ratios r_4/r_3 and r_2/r_1 are both less than one-half, nonadditivity is left as the only surviving explanation. In fact, however, SATSA's model-fitting often turns up a nonadditive component when an rDZ/rMZ ratio is more than one-half. For example, Pedersen et al (39), for the Information scale of the Wechsler test, reported these r's: MZT .78, DZT .20, MZA .55, DZA .33. There is no suggestion of nonadditivity among TRA's, but modeling produced a heritability of 60%, all

nonadditive. The very high MZT r was not in this instance attributed to an "assimilation" or "special" effect.

For a "Life Events" scale (41), the reported r's were MZT .28, DZT .15, MZA .49, DZA .05. Here there is no evidence for nonadditive gene action among TRT's, but modeling indicted heritability of 40%, all nonadditive. The surprising fact that the MZA r_3 was larger than the MZT r_1 was described as "an unexpected environmental finding," suggesting a "contrast effect in that, when reared together, identical twins become polarized in their perception of life events. The DZA correlation is lower than the DZT correlation, which suggests that such a contrast effect occurs only for identical twins." No reason was offered for why such an effect should occur only for identical twins; nor was any reference made to an earlier SATSA assertion (35) that an rDZ/rMZ ratio less than one-half might be attributed to "contrast effects that operate within pairs of fraternal twins but not identical twins." To complicate matters further, Tellegen et al, in a MISTRA paper (47) have speculated about a possible "assimilative co-twin influence" that may reduce genetically determined within-pair differences of DZT's but not of MZT's.

Thus, in SATSA's view, contrast (or polarization) effects can occur exclusively within MZ's, exclusively within DZ's, or presumably within both. The contrast effect, which decreases TRT similarity, is opposite in direction to the assimilation or special effect, which increases MZT similarity (or possibly, in MISTRA's view, DZT similarity). Contrast, if operating only within MZ's, will lower heritability estimates, but if operating only within DZ's it will increase heritability estimates and suggest nonadditivity. There are no a priori psychological grounds to expect which effect, if any, might operate within which zygosity for a given trait, and no post hoc explanations either. Which effects are appealed to depends upon the numerical constellation of correlations that happen to be observed. With assimilation, MZT contrast, DZT contrast, and nonadditive gene action all readily available for service,

it will not be easy to divert SATSA's researchers from their model-fitting

exercises with any empirical information.

There is an almost frivolous character to some of SATSA's analyses,

perhaps most clearly visible in a paper (45) that reports heritability

estimates for the traits of optimism (23%) and pessimism (27%). The two

supposedly independent traits were assessed by a mailed 8-item questionnaire.

Four items assessed optimism -- e.g., "I always look on the bright side of

things." The other four items assessed pessimism -- e.g., "I hardly ever

expect things to go my way." Is it reasonable to suppose that answers to such

a questionnaire will provide data of genetic and biological significance?

Gender and Age

SATSA reports significant mean differences between genders and across

age groups for many of its measures. The gender and age effects are removed

from raw scores by statistical adjustment. Even after such adjustment,

heritability may differ between genders and across age cohorts. SATSA papers

sometimes, but not always, explore this possibility. There are no hard and

fast rules as to when that happens, which allows SATSA considerable

flexibility in assessing the import of its findings.

Two papers on "Life Events" illustrate the point. The first paper (41)

makes no reference to gender or age effects, and reports that model-fitting

yielded heritability of 40% for the total scale, with heritabilities ranging

from 18% to 43% for five sub-scales. The second paper (46) fitted separate

models to the genders, with results that differed substantially; heritability

was now significant only among females. The dramatic gender differences are

illustrated in Table 2, which presents heritabilities for sub-scales, for the

best-fitting model when genders are pooled (41), and for the gender-specific

models (46). The gender-specific models indicated radically different modes of

gene action. For the genders combined, the genetic component of 43% for

controllable events was entirely nonadditive. For females, the heritability of

53% was 37% additive and only 16% nonadditive; for males there was no additive variance, just 14% nonadditive variance.

INSERT TABLE 2 HERE

Another illustration is provided in a pair of papers concerned with self-reported measures of health. The "SUMILL" scores were based on answers to questions about 51 chronic health problems, and "SRHEALTH" scores were based on answers to four broader health questions. The first paper (18) divided twin pairs into four age groups. The best-fitting model for SUMILL indicated that all genetic variance was nonadditive, with heritabilities of 43% for ages less than 50, 40% for ages 50-59, 54% for ages 60-69, and zero for ages 70 and above. For SRHEALTH the best-fitting model indicated that all genetic variance was additive, with heritabilities of zero for ages under 50, zero for ages 50-59, 29% for ages 60-69, and 26% for ages 70 and above. There was no mention of possible gender differences.

The second paper (25) fitted gender-specific models. The best-fitting models now indicated that for both health measures and both genders, all genetic variance was additive. For SUMILL heritability was 46% for females and 27% for males; for SRHEALTH it was 26% for females and zero for males. No reference was made to the earlier assertion that heritabilities for the two health measures varied significantly (in opposite directions) with age. We are left with estimates that the heritability of self-reported health measures ranges from 54% to zero, and that gene action may be entirely additive or entirely nonadditive, depending upon the age bracket and/or the gender with which we are concerned. We have not been told by SATSA whether gender differences occur in all age groups, or whether similar age effects occur within each gender.

Cognitive Ability

We proceed to SATSA's analyses of "general cognitive ability." The major article on this topic (39) was based upon 67 MZT, 89 DZT, 46 MZA, and 100 DZA

pairs, recruited from the core sample and tested in person. The average age
was 65.6 years. From a battery of 13 brief subtests of various cognitive
abilities, the FPC was taken as a measure of general cognitive ability. The
obtained r's were MZT .80, DZT .22, MZA .78, DZA .32. The correlations for
TRT's are no higher than those for TRA's, suggesting no effect whatever of
shared environment. The DZ r's are considerably less than half the MZ r's,
suggesting nonadditive genetic variance. The best-fitting model in fact
attributed all the genetic variance (81%) to nonadditive gene action. The
remaining 19% was assigned to nonshared environment. Rather surprisingly,
Pedersen et al cautioned that their modeling results "should not be
overinterpreted to mean that all of the genetic variance is nonadditive, as
the twin design has only modest power to discriminate the relative importance"
of the two genetic components. Still more surprisingly, eight years later,
Finkel & Pedersen (15) asserted that a basic assumption of their model was
that all gene effects were additive, and cited the previous paper (39) as
evidence to support that assumption.

The same test battery was given to many of the same twins three years
later in a second wave of cognitive testing (43), so longitudinal FPC data
were available for 39 MZT, 33 DZT, 19 MZA, and 54 DZA pairs. The FPC r's for
these twin pairs on the two occasions are given in Table 3, along with data
for the full original sample.

INSERT TABLE 3 HERE

The MZ r's are quite stable across the full and longitudinal samples,
but the DZ r's fluctuate. Compared with the original sample, the DZT r_2 has
diminished to the vanishing point, while the DZA r_4 has increased
substantially. Doubling (r_1-r_2) would give absurd heritability estimates of
156% and 170% on the two testing occasions; the evidence for nonadditive
variance is apparently overwhelming. But no hint of this nonadditive variance
appears in the (r_3-r_4) of the longitudinal sample, where the DZA r's are

considerably more than half the MZA r's. SATSA's modeling now ruled out

nonadditivity, and produced heritability estimates of 82% and 80% for the two

testing occasions. Despite the essentially zero r of the longitudinal DZT's,

the model fitted to the data dropped the d^2 term. The authors gave two reasons

for ruling nonadditive variance out of court: the exclusively additive model

is more parsimonious, and -- typical SATSA practice notwithstanding -- "twin

analyses have little power to distinguish between the two components of

genetic variance."

CONCLUSION

We have described what seem to us to be a number of serious problems in

the design, reporting, and analyses by the psychologists engaged in the MISTRA

and SATSA projects under the rubric of behavior genetics. We recognize that

they are engaged in psychological rather than genetical research. Indeed it

might be said that the only genetical theory involved in their analyses are

the numbers 1, 1/2, and 1/4 representing the genotypic correlation for

identical twins, and the additive and nonadditive genotypic correlations for

fraternal twins. Nevertheless they represent a face of genetics that may be

most familiar to behavioral and social scientists.

SATSA's authors wrote as long ago as 1992, "Indeed it is a legitimate

argument that the ubiquitous evidence for genetic influence on personality

questionnaires makes it no longer interesting to document heritability for yet

another personality trait" (45). The argument evidently failed to persuade

them. In the intervening years SATSA research reports have estimated

heritabilities for traits such as openness to experience, agreeableness, and

conscientiousness (1), perceptions of organizational climate (19), stressful

life events (25), and indirect aggression, verbal aggression, suspicion, and

guilt (17). SATSA has also proceeded to multivariate analyses, in which the

genetic and environmental components of the covariances, as well as the

variances, of two or more traits are estimated (40, 46).

It is not apparent what scientific purposes are served by the sustained flow of heritability numbers for psychological characteristics. Perhaps molecular geneticists need those numbers to guide their search for the underlying genes? Perhaps clinical psychologists need those numbers to guide their selection of therapies that work? Or perhaps educators need those numbers to guide their choice of teaching interventions that will be successful? We have seen no indication of the usefulness of the heritability numbers for any of those purposes. Indeed, it has been widely recognized that malleability is not the opposite of heritability; see e.g. Goldberger (16), Maccoby (28).

A case in point is provided by the recent study of regular tobacco use among SATSA's twins (24). Heritability was estimated as 60% for men, only 20% for women. Separate analyses were then performed for three distinct age cohorts. For men, the heritability estimates were nearly identical for each cohort. But for women, heritability increased from zero for those born between 1910 and 1924, to 21% for those in the 1925-39 birth cohort, to 64% for the 1940-58 cohort. The authors suggested that the most plausible explanation for this finding was that "a reduction in the social restrictions on smoking in women in Sweden as the 20th century progressed permitted genetic factors increasing the risk for regular tobacco use to express themselves." If purportedly genetic factors can be so readily suppressed by social restrictions, one must ask the question, "For what conceivable purpose is the phenotypic variance being allocated?" This question is not addressed seriously by MISTRA or SATSA. The numbers, and the associated modeling, appear to be ends in themselves.

Literature Cited

1. Bergeman CS, Chipuer HM, Plomin R, Pedersen NL, McClearn GE, Nesselroade JR, Costa PT, McCrae, RR. 1993. Genetic and environmental effects on openness to experience, agreeableness, and conscientiousness: an adoption/twin study. *J. Pers*. 61:159-79

2. Bouchard TJ. 1997. Whenever the twain shall meet. *The Sciences*. 37 (Sept./Oct.):52-7

3. Bouchard TJ, Hur YM. 1998. Genetic and environmental influences on the continuous scales of the Myers-Briggs Type Indicator: an analysis based on twins reared apart. *J. Pers*. 66:135-49

4. Bouchard TJ, Lykken DT, McGue M, Segal NL, Tellegen, A. 1990. Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science*. 250:223-28

5. Bouchard TJ, Lykken D, Segal N, Wilcox K. 1986. Development in twins reared apart: a test of the chronogenetic hypothesis. In *Human Growth: A Multidisciplinary Review*, ed. A Demirjian, pp. 299-310. London: Taylor & Francis

6. Bouchard TJ, McGue M. 1990. Genetic and rearing environmental influences on adult personality. *J. Pers*. 58:263-92

7. Bouchard TJ, Pedersen N. 1999. Twins reared apart: Nature's double experiment. In *On the Way to Individuality: Current Methodological Issues in Behavior Genetics*, ed. MC LaBuda, EL Grigorenko, pp. 71-93. Commack, NY: Nova Science Press

8. Bouchard TJ, Segal NL, Lykken DT. 1990. Genetic and environmental influences on special mental abilities in a sample of twins reared apart. *Acta Genet. Med. Gemellol*. 39:193-206

9. Cederlof R, Friberg L, Jonsson E, Kaij L. 1961. Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Acta Genet. Med. Gemellol*. 11:338-62

10.   Cederlof R, Lorich U. 1978. The Swedish twin registry. In Twin *Research: Part B. Biology and Epidemiology*, ed. WE Nance, G Allen, P Parisi, pp.189-95. New York: A. R. Liss

11.   deCastro JM. 1993. A twin study of genetic and environmental influences on medical and psychological traits. *Physiol. Behav*. 54:677-87.

12.   DiLalla DL, Carey G, Gottesman II, Bouchard TJ. 1996. Heritability of MMPI personality indicators of psychopathology in twins reared apart. *J. Abnorm. Psychol*. 105:491-99.

13.   Eckert ED, Bouchard TJ, Bohlen J, Heston LL. 1986. Homosexuality in twins reared apart. *Br. J. Psychiatry*. 148:421-25

14.   Eckert ED, Heston L, Bouchard TJ. 1981. MZ twins reared apart: preliminary findings of psychiatric disturbances and traits. In *Twin Research*, *Vol. 3, Part B*, ed. L Gedda, P Parisi, WE Nance, pp. 179-86. New York: A. R. Liss

15.   Finkel D, Pedersen NL. 2000. Contribution of age, genes, and environment to the relationship between perceptual speed and cognitive ability. *Psychol. Aging*. 15:56-64

16.   Goldberger AS. 1979. Heritability. *Economica*. 46:327-47

17.   Gustavsson JP, Pedersen NL, Asberg M, Schalling D. 1996. Exploration into the sources of individual differences in aggression-, hostility- and anger-related (AHA) personality traits. *Pers. Individ. Diff*. 21:1067-71

18.   Harris JR, Pedersen NL, McClearn GE, Plomin R, Nesselroade JR. 1992. Age differences in genetic and environmental influences for health from the Swedish Adoption/Twin Study of Aging. *J. Geront.: Psychol. Sci*. 47:P213-22

19.   Hershberger SL, Lichtenstein P, Knox, SS. 1994. Genetic and environmental influences on perceptions of organizational climate. *J. App. Psych*. 79: 24-33

20.   Hur YM, Bouchard TJ. 1995. Genetic influences on perceptions of childhood family environment: a reared apart twin study. *Child Dev*. 66:330-45

21.   Hur YM, Bouchard TJ, Lykken DT. 1998. Genetic and environmental influence on morningness-eveningness. *Pers. Individ. Diff*. 25:917-25

22.   Joseph, J. 1998. The equal environment assumption of the classical twin method: a critical analysis. *J. Mind Behav*. 19:325-58

23.   Kaprio J, Koskenvuo M, Artimo M, Sarna S, Rantasalo I. 1979. *Baseline characteristics of the Finnish Twin Registry: Section I: Materials, methods, representativeness, and results for variables special to twin studies*. Publications of the Department of Public Health Science, M47, Univ. Helsinki, Finland.

24.   Kendler KS, Thornton LM, Pedersen NL. 2000. Tobacco consumption in Swedish twins reared apart and reared together. *Arch. Gen. Psychiatry*. 57:886-92

25.   Lichtenstein P, Pedersen NL. 1995. Social relationships, stressful life events, and self-reported physical health: genetic and environmental influences. *Psychol. Health*. 10:295-319

26.   Lichtenstein P, Pedersen NL, McClearn GE. 1992. The origins of individual differences in occupational status and educational level: a study of twins reared apart and together. *Acta Sociol*. 35:13-31

27.   Lykken DT. 1982. Research with twins: The concept of emergenesis. *Psychophysiology*. 19:361-73

28.   Maccoby EE. 2000. Parenting and its effects on children: on reading and misreading behavior genetics. *Annu. Rev. Psychol*. 51:1-27

29.   McClearn GE, Pedersen NL, Plomin R, Nesselroade JR. 1991. *Age and gender effects for individual differences in behavioral aging: the Swedish Adoption/Twin Study of Aging*. Manuscript Report, Center for Devel. Health Genet., Penn. St. Univer.

30.   McCourt K, Bouchard TJ, Lykken DT, Tellegen A, Keyes M. 1999. Authoritarianism revisited: genetic and environmental influences examined in twins reared apart and together. *Pers. Individ. Diff*. 27:985-1014.

31.   McGue M, Bouchard TJ. 1989. Genetic and environmental determinants of information processing and special mental abilities: a twin analysis. In *Advances in the Psychology of Human Intelligence*, *Vol. 5*, ed. RJ Sternberg, pp. 7-45. Hillsdale, NJ: Erlbaum

32.   Moloney DP, Bouchard TJ, Segal NL. 1991. A genetic and environmental analysis of the vocational interests of monozygotic and dizygotic twins reared apart. *J. Vocat. Behav*. 39:76-109.

33.   Newman DL, Tellegen A, Bouchard TJ. 1998. Individual differences in adult ego development: sources of influence in twins reared apart. *J. Pers. Soc. Psychol*. 74:985-95.

34.   Pedersen NL, Friberg L, Floderus-Myrhed B, McClearn GE, Plomin R. 1984. Swedish early separated twins: identification and characterization. *Acta Genet. Med. Gemellol*. 33:243-50

35.   Pedersen NL, Lichtenstein P, Plomin R, DeFaire U, McClearn GE, Matthews KA. 1989. Genetic and environmental influences for Type A-like measures and related traits: a study of twins reared apart and twins reared together. *Psychosom. Med*. 51:428-40

36.   Pedersen NL, McClearn GE, Plomin R, Nesselroade JR. 1992. Effects of early rearing environment on twin similarity in the last half of the life span. *Br. J. Devel. Psychol*. 10:255-67

37.   Pedersen NL, McClearn GE, Plomin R, Nesselroade JR, Berg S, de Faire U. 1991. The Swedish Adoption/Twin Study of Aging: an update. *Acta Genet. Med. Gemellol*. 40:7-20

38.   Pedersen NL, Plomin R, McClearn GE, Friberg L. 1988. Neuroticism, extraversion, and related traits in adult twins reared apart and reared together. *J. Pers. Soc. Psychol*. 55:950-57

39.   Pedersen NL, Plomin R, Nesselroade JT, McClearn GE. 1992. A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychol. Sci*. 3:346-53

40.   Pedersen NL, Reynolds CA, Gatz M. 1996. Sources of covariation among Mini-Mental State Examination scores, education, and cognitive abilities. *J. Geront.: Psychol. Sci*. 51B:P55-63

41.   Plomin R, Lichtenstein P, Pedersen NL, McClearn GE, Nesselroade JR. 1990. Genetic influence on life events during the last half of the life span. *Psychol. Aging*. 5:25-30

42.   Plomin R, McClearn GE, Pedersen NL, Nesselroade JR, Bergeman CS. 1988. Genetic influence on childhood family environment perceived retrospectively from the last half of the life span. *Devel. Psychol*. 24:738-45

43.   Plomin R, Pedersen NL, Lichtenstein P, McClearn GE. 1994. Variability and stability in cognitive abilities are largely genetic later in life. *Behav. Genet*. 24:207-15

44.   Plomin R, Pedersen NL, McClearn GE, Nesselroade JR, Bergeman CS. 1988. EAS temperaments during the last half of the life span: twins reared apart and twins reared together. *Psychol. Aging*. 3:43-50

45.   Plomin R, Scheier MF, Bergeman CS, Pedersen NL, Nesselroade JR, McClearn GE. 1992. Optimism, pessimism and mental health: a twin/adoption analysis. *Pers. Individ. Diff*. 13:921-30.

46.   Saudino KJ, Pedersen NL, Lichtenstein P, McClearn GE, Plomin R. 1997. Can personality explain genetic influences on life events? *J. Pers. Soc. Psychol*. 72:196-206

47.   Tellegen A, Lykken DT, Bouchard TJ, Wilcox KJ, Segal NL, Rich S. 1988. Personality similarity in twins reared apart and together. *J. Pers. Soc. Psychol*. 54:1031-39

48.   Watson P 1982. *Twins: An Uncanny Relationship*. New York: Viking. 207 pp.

TABLE 1   Some heritabilities estimated by MISTRA

_____

Trait	rMZA	rDZA	h^2
Femininity (6)	.33	.50	44%
Morningness-Eveningness (21)	.47	.45	54%
Extraversion-Introversion (3)	.60	.02	57%
Vocational Interests (32)	.50	.47	45-50%
Right-Wing Authoritarianism (30)	.69	.00	50-65%

_____

TABLE 2   Heritabilities of Life Events estimated by SATSA

Sub-scale	Genders Pooled	Males	Females
Undesirable Events	.36	.09	.41
Desirable Events	.31	.08	.50
Uncontrollable Events	.18	.00	.22
Controllable Events	.43	.14	.53

TABLE 3   Correlations for general cognitive ability reported by SATSA
_____
_____

Data Set	MZT	DZT	MZA	DZA
Full original sample	.80	.22	.78	.32
Longitudinal sample, 1st testing	.84	.06	.84	.50
Longitudinal sample, 2nd testing	.88	.03	.70	.48

_____