Are California Psychological Inventory Items Differently Heritable?

John C. Loehlin

Received 5 Dec. 1985—Final 2 July 1986

The procedure for classifying items from the California Psychological Inventory into "genetic" and "environmental" groups, as carried out by Horn, Plomin, and Rosenman (Behav. Genet. 6:17–30, 1976) with the Veterans twin sample, was repeated with male pairs from the National Merit twin sample. There was no significant consistency in the items selected into the two categories in the two samples. It was concluded that individual personality questionnaire items may be inherently too ambiguous for profitable behavior genetic analysis.

KEY WORDS: heritability; items; twins; personality; California Psychological Inventory.

INTRODUCTION

The question whether some personality traits are low and others high in heritability has usually been addressed at the level of personality trait scales (e.g., Carey et al., 1978; Carey and Rice, 1983; Loehlin, 1982, 1985, 1986; Loehlin and Nichols, 1976; Vandenberg, 1967; Young et al., 1980; Zonderman, 1982). The consensus in recent studies seems to be that such scale differences in heritability exist, although there is not yet agreement on just which traits are high and which low.

But might we be making things too difficult by looking at the matter at the wrong level of aggregation? Personality traits, as measured by typical personality inventories, tend to be broad averages over a wide array of specific behaviors. It is at least logically possible that more narrowly...
defined behavior patterns, of the sort tapped by individual questionnaire items, might show sharper differences in the genetic or environmental influences upon them, differences that get blurred together when such items are lumped into scales.

Loehlin and Nichols (1976) examined this hypothesis using data from the California Psychological Inventory (CPI) on 490 pairs of monozygotic (MZ) and 317 pairs of dizygotic (DZ) twins from the National Merit twin sample. They selected 179 CPI items which had a high individual test-retest reliability in Goldberg and Rorer's (1964) study and sorted them into "high"- and "low"-heritability groups separately in two random halves of the total twin sample. High heritability was defined as an MZ intraclass correlation at least 0.20 higher than the DZ intraclass correlation, and low heritability as an MZ correlation at most 0.02 higher than the DZ correlation. There were 55 high- and 38 low-heritability items, by this criterion, in the first half-sample and 54 and 31, respectively, in the second.

Of course, the critical question is, Were they the same items? Loehlin and Nichols concluded that the high-heritability items showed only chance agreement between the two subsamples but that there was a modest degree of consistency at the low-heritability end, which they speculated was due to CPI items which reflected social attitudes—e.g., "A person who doesn't vote is not a good citizen," "I believe women should have as much sexual freedom as men." Attitudes on such items might plausibly be a function of family environment, for these high-school-age twins.

Horn et al. (1976) carried out a similar division of CPI items into high- and low-heritability groups, using data from the Veterans twin sample. This sample consisted of middle-aged male twins who were World War II veterans—99 MZ and 99 DZ pairs. The procedure of Horn et al. differed in a few details. They, too, picked out individually reliable items based on the Goldberg and Rorer study but used a different index of reliability, Goldberg and Rorer's "Ambdex" statistic, which is based on a combination of retest stability and endorsement frequency. Horn et al. also used a somewhat different item selection procedure. They split their sample randomly in half and calculated MZ and DZ intraclass correlations in both half-samples, but they used cross-validation as part of their item selection process rather than as a check on the stability of their results. They defined a "genetic" item as one for which the MZ correlation in each half-sample exceeded the DZ correlation in its own and in the other half-sample by 0.10 or more and for which the average DZ correlation was not more negative than -0.10. An "environmental" item was one for which the average MZ and DZ correlations were both 0.10 or greater,