Commentary

Human Development 2000;43:51–52

Human Development

Statistical Analysis, Experimental Method, and Causal Inference in Developmental Behavioral Genetics

Eric Turkheimer Mary Waldron

Department of Psychology, University of Virginia, Charlottesville, Va., USA

Key Words

Analysis of variance · Behavioral genetics · Research methods

We have had our say on similar topics in these pages before [Turkheimer & Gottesman, 1991; Turkheimer, Goldsmith, & Gottesman, 1995], so we will be very brief now. We agree with much of what Vreeke has to say about the difficulty of building causal developmental models from what he calls 'analysis of variance', but we think he has mislocated his target. If the problem is really with analysis of variance, it is much more widespread than Vreeke suggests, because in one form or another everyone in the behavioral sciences uses ANOVA to analyze data, equally so the experimental ethologist conducting t-tests or Wilcoxons on experimental and control groups of rats, the socioenvironmentalist estimating correlations among family variables, the structural equation modeler fitting complex path diagrams to covariance matrices, and the here beleaguered behavioral geneticist conducting biometric analyses of twins or adoptees.

In fact, what really bothers Vreeke has nothing to do with statistical analysis. He is dubious of efforts to use correlational (in the nonstatistical sense of the word, meaning 'nonexperimental') research designs to make causal inferences about development. We agree, except to the extent that he locates the problem in behavioral genetics. If ANOVA is the universal statistical method in the behavioral sciences, playing fast and loose with causal inference is the universal bad habit, one that socioenvironmental developmentalists are at least as guilty of as behavioral geneticists. The 1999 issues of *Developmental Psychology*, for example, comprise 128 articles, including two commentaries and 126 empirical papers that use ANOVA in some form to analyze data. Out of these 126, 22 use random assignment to groups as a basis for causal inference, but the other 104 – only 5 of them involving behavioral genetics – fit statistical models to correlational data.

Editor's note: The action editor for this article was Claes von Hofsten.

KARGER

Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0018–716X/00/0431–0051$17.50/0

Accessible online at:
www.karger.com/journals/hde

Eric Turkheimer
Department of Psychology, University of Virginia
102 Gilmer Hall, PO Box 400400
Charlottesville, VA 22904-4400 (USA), Tel. +1 804 982 4732
Fax +1 804 982 4766, E-Mail turkheimer@virginia.edu

Perhaps we can find some common ground by working an example. What does Vreeke make of the following study, which we have selected more or less at random from the many available candidates? The NICHD Early Childcare Research Network [1999] recently published a report of child functioning through 36 months in children of depressed and nondepressed mothers. Results, as described in the abstract, were as follows:

> Children whose mothers reported feeling depressed performed more poorly on measures of cognitive-linguistic functioning and were rated as less cooperative and more problematic at 36 months. Depression-group differences in school readiness and verbal comprehension were accounted for by maternal sensitivity. Depressive-group differences in expressive language and ratings of cooperation were moderated by maternal sensitivity: maternal sensitivity predicted better outcomes more noticeably among children whose mothers reported feeling depressed. (p. 1297)

The discussion section begins 'Using a large, diverse sample and their young children... we examined the impact of maternal reports of chronic depression and co-occurring risk factors on mother-child interaction and child functioning' (p. 1305). *Impact of?* The paper concludes with recommendations for intervention programs for depressed mothers. Note that we are not voicing the usual behavioral genetic complaint that the authors are ignoring obvious genetic pathways between mothers and children, although of course they are. What is at issue here, is that the authors are using ANOVA (in this case, regression) to make inferences about complex developmental processes using strictly correlational data.

So what are you going to do? By and large, researchers do not resort to correlational data out of ignorance or laziness. Rather they are faced with the age-old choice between using clones-in-cages designs to study much simplified versions of interesting developmental problems in the laboratory, or studying problems in all their complexity in the field, using methods that are, if one allows oneself to admit it, often difficult to take seriously. We do not have the answer to this dilemma and neither does Vreeke, so in the meantime it will be best if everyone refrains from pointing fingers.

References

NICHD Early Child Care Research Network (1999). Chronicity of maternal depressive symptoms, maternal sensitivity, and child functioning at 36 months. *Developmental Psychology, 35*:51, 1297–1310.

Turkheimer, E., & Gottesman, I.I. (1991). Individual differences and the canalization of human behavior. *Developmental Psychology, 27*, 18–22.

Turkheimer, E., Goldsmith, H.H. & Gottesman, I.I. (1995). Commentary. *Human Development, 38*, 142–153.