A REVIEW OF ACADEMIC PERMANENT-PRODUCT DATA
COLLECTION AND RELIABILITY PROCEDURES IN
APPLIED BEHAVIOR ANALYSIS RESEARCH¹

MICHAEL BRYAN KELLY²

UNIVERSITY OF KANSAS

Three publication sources were reviewed to determine the recent conventions for collecting, and assessing the reliability of, academic permanent-product data (handwriting, examination papers, etc.) in applied behavior analysis. The primary source was the *Journal of Applied Behavior Analysis* (1968-1974). Secondary sources included conference proceedings titled, *A new direction in education: behavior analysis* (E. Ramp and B. L. Hopkins, Eds., Lawrence, Kansas: Support and Development Center for Follow Through, Department of Human Development, University of Kansas, 1971), and *Behavior Analysis and Education* (G. Semb, Ed., Lawrence, Kansas: Support and Development Center for Follow Through, Department of Human Development, University of Kansas, 1972). Finally, as a test of the generality of the findings in the two applied behavior analysis sources, the current issue of each of 14 psychological and/or educational journals was reviewed. Thirty JABA studies reported academic permanent-product data, but only 14 reported reliability. Increasingly more product data through 1973 have been reported along with a greater proportion of authors reporting reliability. The review of the two conference proceeding publications revealed the same trend. In 1971, only three studies reported academic product data, none with reliability, while in 1972, 15 reported academic data, nine including reliability assessment. The review of 14 current education/psychology journal issues revealed four studies reporting academic data, none with reliability. Across all sources, about one-half of the studies reported reliability. Most of the studies reporting reliability described the frequency of reliability assessment, with approximately equal numbers of JABA studies reporting reliability for each paper or reliability for each session. The use of uninformed observers was reported in only three JABA studies and one conference study. Marks made on subjects' papers by either the teacher or the primary observer were reported masked for reliability purposes by only two JABA and two conference studies. Reliability was calculated on a session total basis in two JABA studies. Point-by-point agreement was given in nine JABA and three conference studies. Perfect reliability (mean agreement of 100%) was reported in only six JABA and three conference studies. Scores between 90 and 100% were reported in nine JABA and four conference studies. Scores below 80% were reported in three JABA studies. No other percentage agreement scores were reported, although one JABA study reported correlational reliability (Pearson r). In summary, recently more studies have dealt with academic data and, until 1974, a greater proportion of these studies reported reliability assessment, and relatively few studies reported either replicable methods, 100% agreement, or controls for maintaining rater independence.

DESCRIPTORS: permanent-product data, education, behavioral recording, reliability, academic response

¹The full text of this manuscript is deposited with the National Auxiliary Publications Service (NAPS). For copies, order NAPS Document #02702. Order from ASIS/NAPS Co., C/O Microfiche Publications, 440 Park Avenue South, New York, New York 10016. Remit with order for each copy \$3.00 for microfiche or \$5.00 for photocopies. Make checks payable to Microfiche Publications. The author would like to thank Eugene A. Ramp and the Follow Through graduate students for their support during

the development of this review. The preparation of this manuscript was supported in part by a grant (OEG-0-8-522422-4433) to the University of Kansas Support and Development Center for Follow Through, Department of Human Development, University of Kansas, Lawrence, Kansas 66045.

²Department of Human Development and Family Life, University of Kansas, Lawrence, Kansas 66045.

Received 21 February 1975.
(Final acceptance 1 July 1975.)