

SESSION X CLINICAL APPLICATIONS

GEOFFREY LOFTUS, *University of Washington, President*

Psychological systems questionnaire: An objective personality test designed for on-line computer presentation, scoring, and interpretation

JAMES H. JOHNSON, RONALD A. GIANNETTI, and THOMAS A. WILLIAMS
Eastern Virginia Medical School, Norfolk, Virginia 23501

Previous research efforts to use on-line computer systems for personality assessment are briefly reviewed. Shortcomings in the conversion of paper-and-pencil testing forms to computer media are noted. Construction of a new instrument, specifically designed for use with an on-line computer system, is described. Advantages of this approach are noted.

In the early 1960s, psychologists began to use computers to score and interpret psychological tests. Computer programs were developed for the Minnesota Multiphasic Personality Inventory (MMPI) (Pearson, Swenson, Rome, Mataya, & Brannick, 1965), the Rorschach (Piotrowski, 1964), the California Psychological Inventory (Finney, 1966), the 16 Personality Factor Questionnaire (Karson & O'Dell, 1975), the Sentence Completion Test (Veldman, 1967), the Holtzman Ink Blot Test (Gorham, 1967), the Kelly Role Construct Repertory Test (Space, 1975), the Slosson (Hedl, O'Neil, & Hansen, 1973), and the Wechsler Adult Intelligence Scale (Elwood & Griffin, 1972).

These early computerized assessment approaches relied on batch processing techniques. A standardized test was administered in the traditional manner. Responses were then translated into computer-readable media and input to a computer system for scoring, interpretation, and printing of reports. Interpretative output was based primarily on "cookbook" formulations or algorithms developed to mimic a clinician.

Following the introduction of relatively inexpensive on-line computer technology in the early 1970s, research was undertaken to administer, score, and interpret psychological tests totally by computer. Test items were presented on cathode-ray tube terminals (CRTs), the responses being collected at a keyboard and transmitted directly for computer processing. Examples include MMPI systems developed by Johnson and Williams (1975), Kleinmuntz and McLean (1968), and Lushene, O'Neill, and Dunn (1974).

On-line administration of standardized psychological

tests is more efficient than off-line systems approaches, but researchers have begun to understand that the ultimate potential of on-line assessment systems cannot be realized using instruments originally designed for paper-and-pencil administration. As Giannetti, Klingler, Johnson, and Williams (1976) noted, on-line computer technology offers a fundamentally different medium from the traditional paper-and-pencil medium. On-line technology makes available the computer's decision capabilities, computing power, and data bank throughout the testing process. These features provide test constructors with potential lacking in all previous forms of assessment media.

Several examples describe the important differences in these media. On-line assessment systems can collect data in a flexible manner. By using a logical interrogatory branching sequence rather than a fixed linear series of questions, a large pool of items can be available to obtain information, while the individual being assessed need answer only a subset of items that are relevant to him/her. For example, if a respondent indicates that he/she has never married, items detailing marriage can be skipped. Moreover, the text of specific items can be altered during the administration process to suit the specific respondent. For example, an item beginning with, "Did your brother . . . ?" can be replaced with, "Did your older brother . . . ?" Flexible data gathering schemes can also be used where an obvious rational method does not exist. By using response-contingent testing approaches (Wood, 1973), item branching can be based on previous test performance. For example, validity items might be presented at the beginning of a questionnaire and scored during the course of adminis-

tration. If a respondent scores above a previously determined cutoff point, then further testing is terminated.

Additionally, the speed with which on-line administered psychological instruments can be scored and interpreted allows new applications for testing. Since results can be available within seconds following the completion of test administration, on-line assessment approaches are more likely to be useful for real-time clinical decision making. For example, a psychological test administered at the time of application for care provides data that are used by intake evaluators for triage decision making. On-line systems can, thus, be used for interventionally relevant purposes (Johnson, Williams, Klingler, & Giannetti, 1977) rather than for descriptive purposes.

The full potential offered by the on-line medium for psychological assessment cannot be tapped using currently validated instruments developed for the paper-and-pencil medium. The range of questions available on any one paper-and-pencil assessment device is limited. Altering the item presentation sequence and the items themselves would threaten the previously determined norms, reliabilities, and validities. Therefore, to make full use of the potential of on-line assessment approaches, new psychological assessment instruments need to be developed expressly for the new media.

The only previous work in this area has been limited and preliminary in nature. Greist and his colleagues at Wisconsin have maintained an active research program aimed at developing new assessment devices expressly for the on-line medium (cf. Greist, Gustafson, Stauss, Rowse, Laughren, & Chiles, 1973). They constructed a series of assessment devices with personalized questions and branching techniques. However, these devices are intended to mimic a clinical interview and are based on a face valid model of objective testing. That is, questions are asked and responses are collected that are assumed to be correct and individually valid. The Greist computer interviews and resulting interpretations make use of none of the psychometric technologies that have been developed during the last 50 years of research on paper-and-pencil objective personality test construction. While this work represents an important advance in the history of on-line assessment, it is not dissimilar to the approach taken by Woodworth (1917) in the construction of the *Personal Data Sheet*. Behavioral scientists have long been aware of the fact that such face valid assessment methods are fraught with problems due to limitations in a respondent's honesty, ability to read, insight, memory, and idiosyncratic interpretations of words and phrases (e.g., Landis, Zubin, & Katz, 1935).

Assessment devices that are required not only take advantage of the unique capabilities of the on-line computer medium but also take into account the advances in psychometric theory and practice that occurred within the last 50 years. The purpose of this

paper is to describe the current status of the Psychological Systems Questionnaire (PSQ), an instrument for objective personality assessment that is being developed specifically for the on-line computer medium and is psychometrically up to date. It is our hope that this exposition will be of use to those interested in the problems of on-line assessment and will help to stimulate other developmental efforts in this new area of research.

PSYCHOLOGICAL SYSTEMS QUESTIONNAIRE

Initial work on the Psychological Systems Questionnaire grew out of our experience in the development of an on-line computer-assisted psychological assessment system (PAU) at the Veteran's Hospital in Salt Lake City (Williams, Johnson, & Bliss, 1975). The aim of the PAU is to collect comprehensive background data on individual applicants for care in a timely fashion so results can be used by intake clinicians in making disposition decisions. Patients requesting care are administered a large battery of standardized psychological tests on CRTs. Interpretative results are printed in time to assist the clinician in triage decision making. The PAU data base consists of approximately 2,500 questions about the patient's psychological, social, and medical functioning, and the entire evaluation process requires 5-7 h. For many reasons this process has been viewed as too lengthy. However, the time requirement is fixed by the use of currently available assessment instruments designed for the paper-and-pencil medium. In order to reduce the length of time required for intake data gathering, we began work on the Psychological Systems Questionnaire.

Initially, five senior clinicians were asked to rationally categorize the items within the data base. As a result of this categorization, information was available about the exact number of content areas covered and the item-etric representation within each of the categories. The clinicians felt that some important content areas were excluded from the data base; they also thought that other areas were overrepresented. Experts in clinical care and psychopathology throughout the country were contacted for suggestions regarding an improved item pool. After obtaining their opinions, it was determined that an adequate item pool for treatment decision questions should cover domains such as: somatic history, personality, psychopathological symptoms, social history, quality of living environment, vocational background, family background, readiness to accept treatment, genetic background, and dissimulation. Accordingly, items from our original data base were distributed into each of these categories. Experts in each area were then consulted to construct additional items to provide full coverage for each of the relevant domains. A construction test of approximately 300 items was completed for each of the recognized domains. Each construction test was administered to samples ranging

form 600 to 900 normals and psychiatric patients. Subjects were obtained from community mental health centers in California, Minnesota, Nebraska, and Utah; from psychiatric hospitals in Utah, Washington, and Canada; and from the general population in the State of Utah.

Results for each construction test were factor analyzed to determine the empirical components of item response variance within the rationally determined domain of content. For each subgroup the item correlation matrix was factored with the PA-II subprogram of SPSS using squared multiple correlations as estimates of commonalities. Factors with eigenvalues greater than unity were rotated by the varimax method. Items with varimax factor loadings greater than .4 were retained in the final item pool.

Certain of the domain item pools were subjected to further developmental procedures. For example, the dissimulation items were subjected to an empirical development procedure as well (Johnson et al., 1977). In addition, a set of rational multiple-choice questions was developed to precede domain testing in order to provide demographic information and rational information for test administration branching.

The present PSQ consists of 24 multiple-choice demographic and branching questions and 729 true-false questions. The first 18 true-false questions make up a dissimulation index. Based on the responses to these questions, a determination is made whether or not the remaining 711 questions are to be presented.

A preliminary factor analysis of the PSQ item pool has been completed using 178 subjects (64 females and 114 males). Half of the subjects were drawn from psychiatric settings and half were drawn from normal settings. The 120 items loading most heavily on construction test factors were factor analyzed using the PA-II subprogram of SPSS. Factors with eigenvalues greater than unity were rotated using the varimax method. These factors were then correlated with items from the entire pool. Potential scales were developed from the items correlating significantly ($p < .05$) with each particular factor and where the items correlated higher with the specific factor than with any of the other factors. These potential scales were refined individually through additional factor analysis. The principal factors for each potential scale were used as the basis for construction scales. This procedure resulted in a number of construction scales: family background, genetic background, adolescent socialization, work interest, readiness for treatment, extraversion, socialization, somatic symptoms, acting-out, depression, anxiety, obsessional thinking, manic behavior, schizoid thinking, alcohol abuse, drug abuse, use of social agencies, paranoid symptoms, life changes, social desirability, dissimulation, etc.

At present the computer program for the administration, scoring, and interpretation of the PSQ is under

construction. The program has several unique features. The first items presented are from the validity scales, dissimulation and social desirability. If a subject appears to be in the process of obtaining a score suggesting invalidity on these scales, he/she is reminded to consider his/her answers more carefully. If they continue to answer so that they obtain an invalid score, further testing is discontinued once a cutoff point is reached.

Item presentation is limited to those questions relevant to the subject (i.e., there are no questions about work if the subject is unemployed.) All scales are scored during the administration process. When a respondent is found to have scored high on a particular scale, he/she is asked for comments. For example, the computer might say, "Mr. Jones, your test responses seem to indicate that you are very depressed. Would you care to make any comments about this?" These comments are then taken free-form from the typewritten keyboard.

Interpretative output includes scale scores, a scale profile, and a narrative interpretation that includes item data, scale data, and respondent comments. An example of this integrated approach follows:

Mr. Jones reports that he has previously received outpatient treatment and has been hospitalized once for emotional problems. He indicates that his friends and family have not been supportive in assisting him to get better. Results on the Readiness for Treatment Scale indicate that he is not well motivated to follow a regular treatment program. When questioned about this he commented, "That may have been true in the past, but I think I am ready to work hard now."

We are currently in the process of collecting 2,400 PSQs for final scale development. At the conclusion of data collection, the entire item pool will be simultaneously factor analyzed using a modified version of the BMD factor analysis program. The eventual interpretative output from the computer-administered PSQ will then be augmented with a series of predictive statements about the likely outcome of various forms of treatment. All predictive statements will be determined from linear combinations of the factorially derived scales as based on our previous research in this area (Giannetti, Johnson, Klingler, & Williams, 1978).

DISCUSSION

The present paper describes the rationale and method of construction for the Psychological Systems Questionnaire. We believe that the Psychological Systems Questionnaire represents a major advance in the area of on-line computerized assessment. It is the first questionnaire designed specifically for on-line computer administration that also takes into account contemporary

psychometric theory and practice. It is an interventionally relevant assessment device aimed exclusively at providing information for clinical decision making. While further validation research needs to be done with this instrument, work to date heralds an important breakthrough in the field of on-line psychological assessment.

REFERENCES

- ELWOOD, D. L., & GRIFFIN, H. R. Individual intelligence testing without the examiner: Reliability of an automated method. *Journal of Consulting and Clinical Psychology*, 1972, **38**, 9-14.
- FINNEY, J. C. Programmed interpretation of MMPI and CPI. *Archives of General Psychiatry*, 1966, **15**, 75-81.
- GIANNETTI, R. A., JOHNSON, J. H., KLINGLER, D. E., & WILLIAMS, T. A. A comparison of linear and configural MMPI diagnostic methods with an uncontaminated criterion. *Journal of Consulting and Clinical Psychology*, 1978, **46**, 1046-1051.
- GIANNETTI, R. A., KLINGLER, D. E., JOHNSON, J. H., & WILLIAMS, T. A. The potential for dynamic assessment systems using on-line computer technology. *Behavior Research Methods & Instrumentation*, 1976, **8**, 101-103.
- GREIST, H. H., GUSTAFSON, D. H., STAUSS, F. F., ROWSE, G. L., LAUGHREN, T. P., & CHILES, J. A. A computer interview for suicide-risk prediction. *American Journal of Psychiatry*, 1973, **130**, 1327-1332.
- GORHAM, D. R. Validity and reliability studies of computer-based scoring system for inkblot responses. *Journal of Consulting Psychology*, 1967, **31**, 65-70.
- HEDL, J. H., O'NEIL, H. F., & HANSEN, D. N. Affective reactions toward computer-based intelligence testing. *Journal of Consulting and Clinical Psychology*, 1973, **40**, 217-222.
- JOHNSON, J. H., & WILLIAMS, T. A. The use of on-line computer technology in a mental health admitting system. *American Psychologist*, 1975, **30**, 388-390.
- JOHNSON, J. H., WILLIAMS, T. A., KLINGLER, D. E., & GIANNETTI, R. A. Interventional relevance and retrofit programming: Concepts for the improvement of clinician acceptance of computer-generated assessment reports. *Behavior Research Methods & Instrumentation*, 1977, **9**, 123-132.
- KARSON, S., & O'DELL, J. W. A new automated interpretation system for the 16PF. *Journal of Personality Assessments*, 1975, **39**, 256-260.
- KLEINMUNTZ, B., & MCLEAN, R. S. Diagnostic interviewing by digital computer. *Behavioral Science*, 1968, **13**, 75-80.
- LANDIS, C., ZUBIN, J., & KATZ, S. E. Empirical evaluation of three personality adjustment inventories. *Journal of Educational Psychology*, 1935, **26**, 321-330.
- LUSHENE, R. E., O'NEILL, H. F., JR., & DUNN, T. Equivalent validity of a completely computerized MMPI. *Journal of Personality Assessment*, 1974, **38**, 353-361.
- PEARSON, J. S., SWENSON, W. M., ROME, H. P., MATAYA, P., & BRANNICK, T. L. Development of a computer system for scoring and interpretation of Minnesota Multiphasic Personality Inventories in a medical clinic. *Annals of the New York Academy of Science*, 1965, **126**, 682-692.
- PIOTROWSKI, Z. A digital-computer interpretation of inkblot test data. *Psychiatric Quarterly*, 1964, **38**, 1-26.
- SPACE, L. G. A console for the interactive on-line administration of psychological tests. *Behavior Research Methods & Instrumentation*, 1975, **7**, 191-193.
- VELDMAN, D. J. Computer-based sentence completion interviews. *Journal of Consulting Psychology*, 1967, **14**, 153-157.
- WILLIAMS, T. A., JOHNSON, J. H., & BLISS, E. L. A computer-assisted psychiatric assessment unit. *American Journal of Psychiatry*, 1975, **132**, 1074-1076.
- WOOD, R. Response-contingent testing. *Review of Educational Research*, 1973, **43**, 529-544.
- WOODWORTH, R. S. *Personal data sheet*. Chicago: Stoelting, 1917.