Development of a Computerized Sexual Assessment Laboratory

George W. Barnard, MD; Lynn Robbins, BS; David Tingle, MD, MPH; Theodore Shaw, BA; and Gustave Newman, MD

Computer-assisted psychological assessment has been operational for 25 years. It has been well received by patients and shows a degree of reliability that is comparable to that of conventional testing. The authors report on the development of a computerized psychosexual assessment laboratory in a forensic facility housing convicted sex offenders who volunteered for treatment. The development of the laboratory was in response to the need for uniform psychosocial data on each resident that can be used for making better diagnoses, developing personalized treatment programs, and assessing treatment outcome.

The process of selecting the psychological tests and programming them for interactive administration is described, as are the testing procedures using the penile plethysmograph with stimuli based on the Tanner developmental stages. A computer-assisted psychosocial assessment that produces a psychosocial history was developed. The laboratory's overall value in the treatment program is assessed.

Computer-assisted psychological assessment has been operational for about 25 years. Fowler,¹ in a review of its history and development, stated that psychometricians were among the first faculty members to make use of computers when they began to be established on university campuses in the mid 1950s, although the post World War II enthusiasm for psychological assessment had begun to wane by the late 1950s. The establishment of community mental health clinic systems lead to a resurgence in the demand for psychological testing. The Minnesota Multiphasic Personality Inventory (MMPI) was a natural choice for experimentation in the area of computerized interpretation. Marks and Seeman's² handbook, "The Actuarial Description of the Abnormal Personality," published in 1963, was the first large-scale application of the actuarial method to clinical prediction. This work, together with other "actuarial cookbooks" for the MMPI, paved the way for MMPI administration and evaluation by computer technology. The first such program became operational in the early 1960s in the Mayo Clinic in Rochester, Minnesota. Shortly after their system became operational, three

Dr. Barnard, Ms. Robbins, Dr. Tingle, and Dr. Newman are from the Department of Psychiatry, University of Florida, Gainesville, FL. Mr. Shaw is director, Sex Offender Treatment Unit, North Florida Evaluation and Treatment Center, Gainesville, FL. Address correspondence and reprint requests to Dr. Barnard, Department of Psychiatry, University of Florida, Box J-256, JHMHC, Gainesville, FL 32610.

Based on a paper presented at the Seventeenth Annual Meeting of the American Academy of Psychiatry and the Law in Philadelphia, Pennsylvania on October 18, 1986.

other computer-based test interpretation (CBTI) systems appeared of which perhaps the best known was that developed by Fowler.³ This was made commercially available by Roche in 1965 and was the first national MMPI mail-in CBTI service for psychologists and psychiatrists.

The first direct computer interview was performed in 1966 at the University of Wisconsin, and by the 1970s developments in computer technology had made it possible to administer items on the screen of a computer terminal and to permit the subject to respond on the keyboard. By now, MMPIs could be administered in relatively remote locations and subjected to on-line testing and interpretation. In 1971 Johnson and Williams⁴ developed computer programs to administer, score, and interpret several psychometric and social history instruments at the Veterans Administration Hospital in Salt Lake City, Utah. By 1973 they had established a computer-based psychiatric unit (PAU) in the hospital. Their test battery included an MMPI, an intelligence test, a social history and problem checklist, a depression inventory, and a structured mental status examination conducted by an interviewer and recorded at the terminal. Several studies^{5,6} comparing the PAU with the traditional approach at the hospital were reported. The results of these studies suggested that the PAU assessments were superior and more internally consistent in such parameters as diagnostic accuracy, decision making, patient and staff acceptance, and cost efficiency as compared with traditional

methods and, additionally, provided reports in half the time. Predictably, the availability of a wide variety of computerized psychological testing to various mental health disciplines has lead to territorial disputes.

Much of the benefit of computer interviews as opposed to computerized assessment derives from their inherent structure and specificity. For example. Weitzel et al.⁷ in 1973 compared freeform and structured checklists examination reports on 49 patients for 15 mental status items and found a surprisingly high tendency for the examiners to omit items on the free-form method. In contrast to humanly administered interviews, computer interviews are 100 percent reliable as they never forget to ask a question and, given the same pattern of responses by a patient, the computer will always ask the same question in the same way. They additionally have the potential for being less uncomfortable or embarrassing to the patient especially when sensitive information such as thoughts of suicide, sexual difficulties, or other psychological problems are being covered. Greist and Klein⁸ showed that volunteer subjects were significantly more likely to reveal their sexual problems to the computer than to a psychiatric interviewer, even one of the same sex. To judge from the studies reported to date, patients do not seem to feel illused by computer administered testing. White⁹ in 1983 found that 80 percent of college students preferred taking the MMPI by computer while none preferred pencil and paper administration. Johnson and Williams¹⁰ reported that

Computerized Sexual Assessment Laboratory

46 percent of subjects said that they were more truthful when responding to the computer than to the clinician. Overall. they reported that patients participating in their PAU program were strongly favorable to computerized testing. Although some claim that computer interviews are impersonal and inhumane. this concern is voiced much more frequently by the professionals than by their patients. Other criticisms of computerized interviews include their difficulty with anything other than structured verbal information and their relative inability to tailor the wording of questions. These objections have been met to some degree by modern-day techniques such as "branching," where the answer to a single critical question can determine which series of questions the individual should be exposed to next.

Validity and Other Considerations

In a review of validity studies Moreland¹¹ concluded that most of the studies have supported the validity of computerized testing as compared with testing administered and interpreted by mental health professionals. He felt that the evaluation of validity in this burgeoning area of computerized psychological testing was sufficiencly important that he developed 14 desirable characteristics for such studies in the future.

In their critique of computerized personality assessment, Butcher, Keller, and Bacon¹² reviewed some of the recent advances and suggested future directions for which the versatility and flexibility of computerized personality assessment can be improved. They emphasized that,

although computer administration of standard tests offers a potential savings in testing time, a reduction of examinee mistakes, and an increase in the number of valid protocols as the computerized procedure forces the examinee to answer all the questions, it cannot be assumed that computer-administrated versions of personality instruments are parallel forms to the paper and pencil versions. Factors such as test instructions, difficulty of response requirements, and response latency of the computer may all possibly interact with personality or psychopathological states of the examinees to make particular error patterns more likely.

Hofer and Green¹³ addressed some of these issues with a somewhat different perspective. They pointed out that irrelevant or extraneous factors incidental to the computerized administration may adversely affect test performance so that people would not receive the same score if tested by computer that they would have received if they had been tested conventionally. People accustomed to working with computers might have an advantage taking computerized tests, particularly if the procedures were complicated. Unfamiliarity with computers is probably correlated with ethnicity. gender, age, and socioeconomic status. A nonequivalence due to unfamiliarity might appear statistically as poor performance by many groups. Further differences between a pencil and paper test and computer administration include the need to push a button once a response has been given, preventing the possibility of changing one's reply later and leading to frustration in some test takers. Despite these considerations. Hofer and Green¹³ concluded that the computer presentation is probably psychometrically superior in that the first answer given is probably most likely to be truly representative of what the patient feels at that time. The inability to retrace one's steps and change may still lead to differences between the test as administered conventionally as opposed to computerized. Test performance is particularly vulnerable to surroundings. especially when dealing with an unfamiliar computer. The authors emphasize the need for a comfortable, quiet room, adequate rest periods, and a clean computer display with adequate resolution, absence of glare, clear response devices, and short and uniform time delays between items to avoid extraneous factors that may affect response. A more cautious approach was taken by Matarazzo,¹⁴ who in a recent article expressed concern over the possibility of misuse of the readily available computerized testing and questioned the credibility of current validity tests.

Development of the Laboratory

From the foregoing discussion, it can be seen that a computer has the potential of playing an invaluable role in a sexual assessment laboratory. The North Florida Evaluation and Treatment Center (NFETC) in Gainesville, Florida, includes a 60-bed residential treatment program for men who have been convicted of and sentenced for sex offenses. While in prison, they volunteer to participate in the treatment program at NFETC. They are not there under sentence and can elect to leave the program and return to prison at any time. Although Florida spends 3.5 million per year on its sex offender treatment programs, little is known at the moment about their effectiveness. Up until recently, the assessment of the offenders has often been global, treatment relatively nonspecific, and follow-up evaluation minimal. Despite these shortcomings, obtaining approval from the state to initiate and develop the laboratory proved to be a difficult, uphill struggle.

In the past year, the staff at NFETC has initiated more intensive and innovative treatment techniques for their sex offenders including both behavioral and psychotherapeutically oriented modalities. The laboratory was conceived, therefore, to meet four goals: (1) to define the nature and extent of individual sexual deviancy, ideally, during the patient's first week of treatment; (2) to determine which treatment modalities would have the most beneficial impact; (3) to repeat testing to assess patient progress during the treatment period; (4) to gather statistical data to better determine and assess the profile of sexual deviants and the impact of treatment upon them. From the beginning we attempted to computerize our assessment laboratory as much as possible to help to ensure that our assessment procedures would be standardized from resident to resident and that we could rapidly obtain data permitting us to alter the treatment program to meet the individual needs of the resident if necessary. Additionally, maximal computerization would facilitate the measurement of the resident's changes in the program

using initial data as a baseline, and we could readily compare data between offender groups such as rapists and pedophiles.

Psychological Tests

In the process of selecting the psychological tests, we consulted with directors of laboratories engaged in research with the sex offender and reviewed the sex offender assessment literature.^{15, 16} We were unable to secure permission from several of the publishers to write the software to test interactively at the computer so we had to substitute other tests for which we could secure permission. There were three tests, however, for which we could not find adequate substitutes. These tests are given in the paper and pencil form. The technician scores them by hand and enters the data into the computer.

Because we needed short tests that would provide good estimates of the resident's intelligence and reading level, we selected the Shipley Institute of Living Scale,^{17,18} from which we derive an estimated verbal IQ, and the Wide Range Achievement Test,¹⁹ which provides an estimate of grade reading level.

We included two measures of general personality. The Personality Research Form- E^{20} gives 20 different personality measures that are of relevance in understanding a population of sex offenders. The Interpersonal Behavior Survey²¹ measures the presence of aggressive and assertive behaviors and assesses their effects on interpersonal interactions. Because so many of our sex offender population report problems with alcohol, we included the Michigan Alcoholism

Screening Test,²² an instrument devised to detect alcoholism.

Several inventories were included that document the resident's sexual attitudes. beliefs, and experiences. The Clarke Sex History Ouestionnaire²³ is an instrument that investigates sexual experiences through the life span. It requires detailed responses of the nature, frequency, and diversity of sexual activity. Cognitive distortions related to pedophilia are detected by the Abel Pedophilia Cognition Scale.²⁴ The Burt Scales²⁵ measure cognitive acceptance of interpersonal violence and rape myth, sex role stereotyping, sex role satisfaction, adversarial sexual beliefs, and sexual conservatism. Abel and Becker's Sexual Interest Cardsort²⁶ measures the degree of sexual arousal or repulsion to scenarios describing the variants of sexual behavior. The Multiphasic Sex Inventory²⁷ assesses a wide range of psychosocial characteristics. It has proved to have real value in further delineating the resident's deviant sexual interests and behavior. It has a child molest scale, a rape scale, an exhibitionism scale, and an atypical sexual outlet scale.

In pretesting these computerized tests, we saw the need for a simplified keyboard. We made up double-key covers with the question responses, "true," "false," "yes," "no," and the numbers 1 through 10. We blanked out all the other keys on the keyboard. Because our program allows the resident to go backward to previous questions and change his answers, we made key covers that read "forward" and "backward." We also allowed for the resident to stop the testing procedure at any point and pick up at a later time, so we had a key cover that read "stop." Finally, arrow keys were added to indicate clearly the manner in which the cursor is moved around on the screen. We found this simplified keyboard to be nonthreatening to the user. We "locked" the other keys, i.e., made them nonusable, so there is very little the resident can do to disrupt the testing procedure through improper response at the keyboard.

In addition to these psychological tests, we developed a computer-assisted psychosocial assessment (CAPSA) program that each resident takes interactively at the computer following his battery of psychological tests. CAPSA provides a comprehensive data base of the subject's psychosocial history.

To be sure that our subjects understood our computerized battery we obtained educational consultation. Readability is the objective measure of the difficulty of written materials. It is generally reported in terms of grade levels. Word difficulty and sentence length are two major factors that influence readability. Using Fry's Graph for Estimating Readability,^{28,29} we found that one half of our items required a fifth- to sixthgrade reading level and the remainder required a seventh-grade reading level. This goal was easily met by the vast majority of our subjects.

Physiological Tests

The physiological portion of our assessment program comprised two systems: a biolab and penile plethysmograph. The biolab consists of a threechannel computerized biolab by Autogenics-Cyborg that measures heart rate. pulse wave velocity, and galvanic skin response. It can be used as a biofeedback instrument or to measure the physiological responses of the individual as he is presented with auditory or visual sexual stimuli. The computerized penile plethvsmograph by Technicraft measures the sexual arousal experienced by the resident in response to these erotic stimuli. During this test, the resident sits in a private, sound-attenuated room with a mercury strain gauge fitted on the shaft of his penis. The computer runs the projector or the tape recorder, records the baseline level, measures the percentage of full erection and the length of time the erection is maintained, converts from analog to digital data, determines when detumescence has occurred, and then presents the next stimulus.

At the present time we are using slides, originally confiscated by the Federal Bureau of Investigation, that we obtained from another research laboratory. We have three slides each in nine categories of age, sex, and violence.

Because there has been little attempt to standardize the stimuli, we are in the process of developing our own audiotape and stimulus slides.³⁰ Both the slides and the audio will be based on the Tanner³¹ developmental stages, which are more definitive criteria than age, and the audio will have additional controls for suggestibility stimuli.

The plethysmograph reports have had significant value insofar as some residents who deny their paraphilias in the psychological testing, admit to them when confronted with these data. The plethysmograph data, while extremely valuable, are not "foolproof" in that some individuals have the ability to suppress physiological responses to the stimuli.

Procedure for Testing

Each incoming resident is tested in the first week of admission. During the several days delegated to the laboratory assessment, the resident is kept relatively isolated from the other residents in the sex offender unit. We instituted this procedure to control the passing of information from long-term residents as to the nature of the assessment and "proper" ways to respond.

A detailed and informed consent form is read. understood, and signed by the prospective patient before embarking on the initial evaluation at the laboratory. Evaluation in the laboratory is now a requirement for treatment in the sex offender program at NFETC. A trained laboratory technician is present at all times during the testing procedures and strict adherence to professional standards is maintained. Should the individual have any emotional problems during the testing procedure, professional help is readily available.

The psychologists administer the MMPI, Wechsler Adult Intelligence Scale-Revised, Rorschach, Draw-A-Person, and Thematic Apperception Test during the second week after admission.

Discussion and Future Developments

The assessment laboratory has only been in full operation for a few months,

so we have not had the opportunity to develop a retesting procedure. As the treatment program at NFETC lasts approximately 18 months, we plan to retest at the halfway mark and upon completion of the program. One concern we had when we selected our psychological test battery was that the instruments should be designed to show change over time. Whether they will measure those changes brought about by the treatment program remains to be seen.

Intuitively, it would seem that, unless all deviant sexual arousal choices can be addressed in treatment rather thanmerely the one for which the individual is convicted, recidivism would be likely. Abel *et al.*³² found that 70 percent of their 24 cases showed paraphilic arousal not reported during the initial clinical interview, and Freund and Costell³³ performed psychophysiological measurements on nondeviant men who responded not only to adult women but also to adolescents and young girls in decreasing order indicating that deviant arousal may be a quantitative and not a qualitative difference. Although penile tumescence measurements may be inconclusive in one third of admitting and two thirds of nonadmitting pedophiles, they do remain the most reliable indication we have of sexual preference and in our laboratory lend themselves readilv to correlation with the items admitted to in the sexual history and questionnaire.

Future data gathered by the laboratory might throw further light on the childhood and adolescent characteristics of paraphiliacs and should enable us to continue our earlier work in this important and little understood area.³⁴ Although our psychological consultants feel that we have "a good mix" of standardized psychological tests it is possible in the future that we may add, omit, or in other ways change the existing protocol as experience is gained. From our immediate perspective the rapidity with which an initial evaluation can be performed greatly expedites the development of a personalized treatment plan and the initiation of treatment.

References

- Fowler RD: Landmarks in computer-assisted psychological assessment. J Consult Clin Psychol 53:748–59, 1985
- 2. Marks PA, Seeman W: The Actuarial Description of Abnormal Personality: An Atlas for Use with the MMPI. Baltimore, Williams & Wilkins, 1963
- 3. Fowler RD: Computer processing and reporting of personality test data. Paper presented at the meeting of the American Psychological Association. Los Angeles, September 1964
- Johnson JH, Williams TA: The use of online computer technology in a mental health admitting system. Am Psychol 30:388-90, 1975
- 5. Klinger DE, Miller DA, Johnson JH, et al: Strategies in the evaluation of an on-line computer-assisted unit for intake assessment of mental health patients. Behav Res Methods Instrum 8:95-100, 1976
- 6. Klinger DE, Miller DA, Johnson JH, *et al*: Process evaluation of an on-line computerassisted unit for intake assessment of mental health patients. Behav Res Methods Instrum 9:110-16, 1977
- Weitzel WD, Morgan DW, Guyden TE: Toward a more efficient mental status examination. Arch Gen Psychiatry 28:215–18, 1973
- Greist JH, Klein MH: Computer programs for patients, clinicians and researchers in psychiatry, in Technology in Mental Health Care Delivery Systems. Edited by Sidowski JB, Johnson JH, Williams TA. Norwood, NJ, Ablex, 1980, pp 161–82

- White DM: An assessment of the comparability of a computer administration and standard administration of the MMPI. Unpublished master's thesis, University of Alabama, Tuscaloosa 1983
- Johnson JH, Williams TA: Using on-line computer technology in a mental health admitting system, in Technology in Mental Health Care Delivery Systems. Edited by Sidowski JB, Johnson JH, Williams TA. Norwood, NJ, Ablex, 1980, pp 237-52
- 11. Moreland KL: Validation of computer-based interpretations: Problems and aspects. J Consult Clin Psychol 53:816-25, 1985
- Butcher JN, Keller LS, Bacon SF: Current developments and future directions in computerized personality assessment. J Consult Clin Psychol 53:803–15, 1985
- 13. Hofer PJ, Green BF: The challenge of competence and creativity in computerized psychological testing. J Consult Clin Psychol 53:826-38, 1985
- 14. Matarazzo JD: Computerized clinical psychological test interpretation. Am Psychol 41:14-24, 1986
- Barlow DH: Assessment of sexual behavior, in Handbook of Behavioral Assessment. Edited by Ciminero AR, Calhoun KS, Adams HE. New York, Wiley, 1977, pp 461–508
- 16. Laws DR, Osborn CA: How to build and operate a behavioral laboratory to evaluate and treat sexual deviance, in The Sexual Aggressor: Current Perspectives on Treatment. Edited by Greer JG, Stuart IR. New York, Van Nostrand Reinhold Company, 1983, pp 293-335
- Shipley WC: Shipley Institute of Living Scale for Measuring Intellectual Impairment: Manual of Directions and Scoring Key. Hartford, CT, The Institute of Living, 1939
- Heinemann AW. Harper RG, Friedman LC, et al: The relative utility of the Shipley-Hartford Scale: Prediction of Wais-R IQ. J Clin Psychol 41:547–51, 1985
- Jastak S, Wilkinson GS: The Wide Range Achievement Test-Revised. Wilmington, DE. Jastak Associates, 1984
- Jackson DN: Personality Research Form Manual. Port Huron, MI. Research Psychologists Press, 1984
- Mauger PA, Adkinson DR: Interpersonal Behavior Survey (IBS) Manual. Los Angeles, Western Psychological Services, 1980
- Selzer ML: The Michigan Alcoholism Screening Test: the quest for a new diagnostic instrument. Am J Psychiatry 127:89–94, 1971
- 23. Langevin R, Handy L, Paitich D, et al: Ap-

Bull Am Acad Psychiatry Law, Vol. 15, No. 4, 1987

Computerized Sexual Assessment Laboratory

pendix A: A new version of the Clarke Sex History Questionnaire for Males, Erotic Preference, Gender Identity, and Aggression in Men: New Research Studies. Edited by Langevin R. Hillsdale, NJ, Erlbaum, 1985, pp 287–305

- Abel GG, Becker JV, Cunningham-Rathner J, et al: The Treatment of Child Molesters. Atlanta, Behavioral Medicine Laboratory, Emory University, 1984
- 25. Burt MR: Cultural myths and supports for rape. J Pers Soc Psychol 38:217–30, 1980
- Abel GG, Becker JV: Sexual Interest Cardsort. Atlanta, Behavioral Medicine Laboratory, Emory University, 1985
- 27. Nichols HR, Molinder I: Multiphasic Sex Inventory Manual. Tacoma, WA, Nichols & Molinder, 1984
- 28. Fry E: Readability formula that saves time. J Reading 11:513-6, 575-578, 1968
- 29. Fry E: The Readability graph validated at

primary levels. Reading Teacher 22:534-8, 1969

- 30. Fuller AK, Barnard GW, Robbins L, *et al*: Sexual maturity as a criterion for classification of phallometric stimulus slides. Arch Behav, in press
- Tanner JM: Foetus into Man: Physical Growth from Conception to Maturity. Cambridge, Harvard University Press, 1962
- 32. Abel GG, Mittelman MS, Becker JV: Sexual offenders: results of assessment and recommendations for treatment, in Clinical Criminology: Current Concepts. Toronto, M & M Graphics, 1985
- Freund K, Costell R: The structure of erotic preference in the nondeviant male. Behav Res Ther 8:15-20, 1970
- 34. Tingle D, Barnard GW, Robbins L, et al: Childhood and adolescent characteristics of pedophiles and rapists. Int J Law Psychiatry 9:103-16, 1986