Using HyperCard to administer a figural test on the Apple Macintosh

DAVID SCHULDBERG

University of Montana, Missoula, Montana

and

WILLIAM G. NICHOLS

Harvard, Massachusetts

To date, applications of automated assessment techniques in personality testing have largely been limited to objective personality instruments with text stimuli; few assessment applications have involved graphic stimuli. Although projective personality instruments generally include ambiguous graphic or pictorial stimuli, computer applications with these procedures have been limited to automated scoring and interpretation, administration of sentence completion devices employing text stimuli, and the use of mechanical methods rather than computer graphics to display visual stimuli. In the present report, we describe a Macintosh HyperCard application for administering an objective personality test with visual stimuli, the Barron-Welsh Revised Art Scale of the Welsh Figure Preference Test. This test consists of a series of figural stimuli and a binary "like"/"dislike" response format, and it thus represents an administration procedure between standard objective self-report inventories involving text stimuli and a "true"/"false" response or variant, and tests such as the Rorschach or TAT that are both figural and free-response. The HyperCard language provides a variety of promising techniques useful for microcomputer test administration.

To date, computer applications in psychological assessment have included pictorial stimuli only to a limited extent. Although computer graphics displays are used in psychological experimentation and clinical neuropsychological testing, pictorial stimuli have met with limited application in automated personality assessment. It is likely that, in the future, increased use will be made of analogue test procedures or simulations, and demonstrations have already been developed in which CD-ROM storage devices have been employed in interactive analogue assessment and training.

The first author gratefully acknowledges the support of a NIMH Postdoctoral Fellowship in Clinical Research from the Department of Psychiatry, Yale University, the assistance of the Yale University Clinical Science Research Center in Psychiatry (Grant 5 P50 MH 30929), the University of Montana Office of Research Administration, and a grant from Montanans on a New Track for Science. He would also like to thank Donald M. Quinlan and Will Cate for their assistance. Requests for reprints or correspondence regarding this paper should be sent to David Schuldberg, Department of Psychology, University of Montana, Missoula, MT 59812-1041. Instructions for obtaining a diskette with a sample HyperCard stack and scripts are also available.

Microcomputer technology is useful in the administration, scoring, and interpretation of psychological tests (Ben-Porath & Butcher, 1986; Butcher, 1987; Fowler, 1980; Moreland, 1987; Space, 1981). The computer has special advantages if adaptive or tailored methods of administration are used (e.g., Ben-Porath, Slutske, & Butcher, 1989). In this type of testing, an algorithm is used to select highly informative next items to assess a particular attribute and to terminate the test when an abbreviated set of items has provided sufficient information. Microcomputers are also extremely useful for scoring objective personality instruments, especially when many items are scored for a large number of scales or when item weightings are used. Finally, attention and criticism have been directed toward the use of microcomputers to generate automated narrative interpretations on the basis of a subject's test profile (Matarazzo, 1985).

In earlier work with computers and projective tests, Veldman (1967) administered a simplified sentence completion device with a computer. This test involved a text stimulus and allowed only a single word response, representing a rudimentary but important demonstration of computer applications to projective test methodology. Miller (1986) developed a BASIC program to guide subjects through what amounted to a self-administration of the traditional Rorschach cards; prompted by the computer, each subject handled the actual cards and entered the responses on the computer keyboard. This procedure allowed the free responses characteristic of the Rorschach and of most traditional projective devices (albeit using a keyboard), but it did not utilize the graphics capabilities now available for presenting test stimuli on a monitor screen.

In the present paper, we focus on the use of a microcomputer as a tool for administering figural test items, and demonstrate an application for presenting a non-animated figural stimulus in a personality assessment device. The subject responds to each item with "like" or "dislike." The use of a pictorial stimulus in this application, along with a constrained binary response format, places it between the Rorschach or TAT, which involve pictorial stimuli and a free verbal response format, and commonly automated tests such as the MMPI, which employ a text stimulus and a binary ("true"/"false") response format.

Description

A computer-administered version of the Barron-Welsh Revised Art Scale of the Welsh Figure Preference Test was developed as part of a research project being done to investigate format and retest effects on test occasion equivalence in automated testing (Schuldberg, 1988). The Barron-Welsh Revised Art scale (Welsh & Barron, 1963) consists of 60 line drawings, designs, or figures, to which

the subject responds with either "like" or "dislike." Validation research has indicated that the scale can serve as a measure of creativity or creative potential that is relevant to a variety of endeavors (Welsh, 1975).

In the conventional administration of the test, the designs are presented in a test booklet that contains a total of either 400 or 86 items, and the subject responds on a separate conventional answer sheet. Half of the 60 items on the Revised Art scale (a subset of the entire Welsh Figure Preference Test) are keyed "like," and half are keyed "dislike."

The present application was created with HyperCard for the Apple Macintosh and its associated script language, HyperTalk (Apple Computer Inc., 1987, 1988). The test stimuli are arranged as cards in a HyperCard stack, with 1 figure per card. The presentation of the figures one at a time differs from the format of the test booklets, where each set of two pages contains an array of up to 16 figures. The 26 figures in the Art Scale test booklet not scored on the Revised Art Scale were omitted to conserve disk space.

A single-card background is used for each of the stimulus cards, with three buttons that allow the subject to respond with "like" or "dislike," or to back up to the previous card (see Figure 1). The program "scripts" associated with the buttons are part of the card background. The background also includes a text field for the item number and a gray background similar to that used in the booklet forms of the test. This gray background contains a blank rectangular "hole" into which the different individual stimulus items were pasted. The proportions of this rectangle match those in the conventional test booklet. To achieve the maximum amount of resolution and detail possible in reproducing and displaying the figures, some of which are low-contrast images, the figures are reproduced at approximately 150% size on the Macintosh

screen. The subject responds to each stimulus by using the mouse to "click" a button on the card.

The items of the Barron-Welsh Revised Art Scale were digitized with the Thunderscan scanner and software (Thunderware, Inc., 1987); the images were scanned in enlarged form, the actual size presented on the screen. After the raw scanned images were saved for future use, editing, and possible contrast enhancement, each individual figure was "cut" from the gray background digitized from the original test booklet (which did not scan well), "pasted" into the Macintosh scrapbook, and then "pasted" onto a new card in the stack of stimulus items. The images are reproduced on the Macintosh screen with moderate faithfulness; the most image degradation occurs for some line images, which appear considerably more fuzzy than in the printed original, and for images including a low-contrast gray wash. Some images were touched up slightly with the HyperCard paint tool when errors in the digitizing process added distracting elements or resulted in omissions of important portions of the figures.

The stimulus cards, as well as a title card, instructions for using the mouse and for taking the test, and other introductory and concluding cards, are arranged in a HyperCard stack. The subject's progress through the test is controlled by brief HyperTalk scripts associated with the stack and with individual cards and buttons. In research situations, an examiner is available to explain the use of the mouse at the beginning of the session. On a later card, the subject is asked for a research identification number or name and gender; these are stored in HyperCard "containers." Progress through these introductory cards is controlled by the test-taker's clicking a "Continue" button.

When the subject has finished providing the introductory information, has run through the instruction cards, and arrives at the first test item, the date and time for the beginning of the test are stored. Additional scripts could be written to compute and store response latencies for each card. The response to each figure is stored as an item in a "container" (answerPrefs in the script below). At the end of the test, the subject's demographic information, the date, the times of starting and finishing the test, and the "container" of item responses are all written to a disk file. At the close of the session, this file is closed, the data disk is ejected, and the computer is shut down.

As a demonstration of the simplicity and flexibility of the HyperCard script language (HyperTalk) for assessment applications, two sample scripts are reproduced here.

The "Like" button has the following short script:

```
on mouseUp
  recordLikes "L"
  go to next card
end mouseUp
```

The second script is a stack script that contains the procedure for storing the item number from the card, and an "L" or "D" corresponding to "like" or "dislike" for each figure:

Figure 1. Sample HyperCard card for an item of the Barron-Welsh Revised Art Scale. Portions of the Barron-Welsh Art Scale of the Welsh Figure Preference Test are reproduced by special permission of the publisher, Consulting Psychologists Press, Inc., Palo Alto, CA 94306, from the Barron-Welsh Art Scale, by George S. Welsh, Ph.D. and Frank Barron, Ph.D. Copyright 1949. Further reproduction is prohibited without the Publisher's consent.

```

on recordLikes preference
  global answerPrefs
  put first word of field "ItemNumber" & tab & preference
  into line (number of this card) of answerPrefs
end recordLikes

```

A script on the final card opens a uniquely named disk file and writes the demographic information and data to the disk. Additional scripts handle such tasks as returning to a previous card if the subject pushes this button. A script is available for randomizing the order of the cards at the opening of the stack and presenting the figures in this new order. The entire procedure for administering a test of arbitrary length and saving the data can be set up with approximately 50 lines of HyperTalk code.

For the HyperCard application described here, copies of the stack and HyperCard software are kept on a floppy disk for each subject. Subjects are tested in a Macintosh teaching laboratory, and the application is started separately for each subject. Responses are saved on each subject's disk and collated for later analysis.

Discussion

We expect that with some coaching on the use of the mouse, even computer novices will be able to use this testing procedure. This type of stack should be very useful in other psychological assessment procedures and tests for educational use, especially devices that access large databases of text, still pictures, or animated materials stored on CD-ROM media. HyperCard interfaces for CD-ROM devices are already available, and this software represents a powerful and convenient method for interactive access to databases that include complex stimuli, pictures, or moving video images. Given such large stimulus databases, interactive assessment procedures or simulations become practical where the "items" are presented adaptively, depending on the subject's responses.

Additional considerations and possible additions to the present stack. The present software was developed primarily to administer the items of the Barron-Welsh Revised Art Scale and to demonstrate the usefulness of the HyperCard script language for figural test administration. Since the subject's responses are stored on the same floppy disk as the testing stack, the safeguarding of the integrity of the test data is accomplished through the use of subject code numbers, ejecting the disk at the end of the program, and the physical locking up of the diskettes after use. Procedures could be used to lock and encrypt the data files, but these are of limited practical usefulness in protecting data from users familiar with the system.

Since the software was developed to administer only one scale, the scoring of each subject's responses is accomplished with a separate program that counts the endorsement of the 30 indicative and 30 nonindicative items. Overall rates of endorsement of "like" or "dislike" are also tallied, to provide an index of possibly invalid

responding. In research with the computerized version of the test, this scoring is accomplished in a separate pass after the data from a group of subjects are collated. HyperCard scripts could easily be developed to score the test "on the fly." It would also be possible to administer the test to many subjects, and to collate, store, and analyze responses over a local area network; this, however, presents additional problems of data security.

The test is primarily a research instrument, and interpretation of the Revised Art Scale is done with reference to means from a wide range of samples reported in the manual (Welsh, 1980). It would also be possible to write a set of HyperTalk scripts that would compare a subject's score to the normative data and print or display a brief paragraph reporting results and providing interpretive information.

Strengths and weaknesses of HyperCard as a tool for developing psychological assessment applications.

This stack and associated scripts can easily be adapted to any style of objective personality test, not only ones with a figural stimulus; the buttons may be relabeled "true" and "false," and text items can be placed on the cards either in text fields or as pasted images. Similarly, free-format responses typed in by the subject could be stored as text fields associated with each card.

It is worthwhile to note that the use of scripts in this testing application to constrain and limit the subject's passage through the testing stack represents a violation of the basic Hypercard "philosophy," which emphasizes the user's freedom in moving between cards and among stacks. HyperCard represents a powerful environment for creating and using psychological assessment tools (Geisler-Brenstein & Brenstein, 1989; Martin & Wilcox, 1989). With new additions to HyperCard software to include color, improvements in the software's speed of execution and use of disk space, enhancements in the script language, and the increasing availability of programming linkages to other languages, HyperCard will become an increasingly friendly and powerful development tool for applications in psychological assessment and experimentation.

Possible future applications. An extension of this application would be the development of a HyperCard stack with a more classically "projective" personality test. As an example, with the color capabilities of the Macintosh II, with the availability of a new version of the HyperCard software with the capability to handle color images (which older HyperCard versions could not do), or with the use of HyperCard-like MS-DOS software to exploit the improved quality of IBM-compatible EGA and VGA displays, one could envision a completely computer-administered Rorschach, with the inkblots presented on the color monitor and the subject's entering free-format responses at the keyboard, dispensing with both the traditional Rorschach plates and the tester's verbatim transcription of responses. Digital sound playback and sampling may some day provide a way of giving the test instructions and inputting the subject's responses to the stimuli,

providing a task that would be similar in many respects to a traditionally administered projective test, save for the (possibly crucial) absence of an actual human examiner.

Simulations are likely to represent the future direction of automated psychological assessment. Work has already been done on skills assessment and training packages (for both social and business skills) that include interactive video techniques. The stimuli in these applications consist of moving video scenarios stored on CD-ROM devices. The future may hold a confluence of projective and cognitive behavioral or "behavior analytic" (Goldfried & D'Zurilla, 1969) assessment techniques, methods involving complex, digitized stimuli that are either ambiguous, in the case of projective assessment, or highly realistic video "movies" ("videos"), in analogue behavioral assessment. Increasingly, subjects will be allowed to make free-form linguistic responses or to engage in actual task-relevant behaviors such as moving a joystick in a flight trainer. Assessment will also become increasingly interactive or adaptive, including provisions for branching to different situations or scenarios depending on the subject's response to preceding items. The sophistication of portable and laptop computers makes them potentially very powerful platforms for psychological assessment.

Research is being conducted on the equivalence of the computerized and pencil and paper forms of the Barron-Welsh Revised Art Scale; a computerized version of a test is not necessarily the "same test" as its paper and pencil counterpart, and a variety of tests of format equivalence need to be made when evaluating a computerized version of an instrument (Honaker, 1988; Wilson, Genco, & Yager, 1985). The present application demonstrates that the HyperTalk script language is sufficiently powerful and flexible to allow the HyperCard environment to be adapted to fulfill a number of different functions in psychological assessment. HyperCard's particular strengths lie in its provisions for presenting a variety of forms of information to the subject, including text, figural, and even animated stimuli, as well as in its flexible capability for providing control over the presentation of stimuli.

REFERENCES

- APPLE COMPUTER, INC. (1987). *Macintosh HyperCard user's guide*. Cupertino, CA: Author.
- APPLE COMPUTER, INC. (1988). *Macintosh HyperCard script language guide: The HyperTalk language*. Reading, MA: Addison-Wesley.
- BEN-PORATH, Y. S., & BUTCHER, J. N. (1986). Computers in personality assessment: A brief past, an ebullient present, and an expanding future. *Computers in Human Behavior*, 2, 167-182.
- BEN-PORATH, Y. S., SLUTSKE, W. S., & BUTCHER, J. N. (1989). A real-data simulation of computerized adaptive administration of the MMPI. *Psychological Assessment: A Journal of Consulting & Clinical Psychology*, 1, 18-22.
- BUTCHER, J. N. (1987). The use of computers in psychological assessment: An overview of practices and issues. In J. N. Butcher (Ed.), *Computerized psychological assessment: A practitioner's guide* (pp. 3-14). New York: Basic Books.
- FOWLER, R. D. (1980). The automated MMPI. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 69-84). Norwood, NJ: Ablex.
- GEISLER-BRENSTEIN, E., & BRENSTEIN, R. J. (1989). The potential of HyperCard for psychological research and instruction: A general discussion and description of two research applications. *Behavior Research Methods, Instruments, & Computers*, 21, 307-311.
- GOLDFRIED, M. R., & D'ZURILLA, T. J. (1969). A behavior-analytic model for assessing competence. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 1, pp. 151-196). NY: Academic Press.
- HONAKER, L. M. (1988). The equivalency of computerized and conventional MMPI administration: A critical review. *Clinical Psychology Review*, 8, 561-577.
- MARTIN, T. A., & WILCOX, K. L. (1989). HyperCard administration of a block-design task. *Behavior Research Methods, Instruments, & Computers*, 21, 312-315.
- MATARAZZO, J. D. (1985). Clinical psychological test interpretations by computer: Hardware outpaces software. *Computers in Human Behavior*, 1, 235-255.
- MILLER, F. E. (1986). The development and evaluation of an online computer-assisted Rorschach inkblot test. *Journal of Personality Assessment*, 50, 222-228.
- MORELAND, K. L. (1987). Computerized psychological assessment: What's available. In J. N. Butcher (Ed.), *Computerized psychological assessment: A practitioner's guide* (pp. 26-49). NY: Basic Books.
- SCHULDBERG, D. (1988). The MMPI is less sensitive to the automated testing format than it is to repeated testing: Item and scale effects. *Computers in Human Behavior*, 4, 285-298.
- SPACE, L. G. (1981). The computer as psychometrician. *Behavior Research Methods & Instrumentation*, 13, 595-606.
- THUNDERWARE, INC. (1987). *ThunderScan user's guide*. Orinda, CA: Author.
- VELDMAN, D. J. (1967). Computer-based sentence-completion interviews. *Journal of Counseling Psychology*, 14, 153-157.
- WELSH, G. S. (1975). *Creativity and intelligence. A personality approach*. Chapel Hill, NC: Institute for Research in Social Science.
- WELSH, G. S. (1980). *Manual: Welsh Figure Preference Test*. Palo Alto, CA: Consulting Psychologists Press.
- WELSH, G. S., & BARRON, F. (1963). *Barron-Welsh Art Scale*. Palo Alto, CA: Consulting Psychologists Press.
- WILSON, F. R., GENCO, K., & YAGER, G. G. (1985). Assessing the equivalence of paper and pencil vs. computerized tests: Demonstration of a promising methodology. *Computers in Human Behavior*, 1, 265-275.

(Manuscript received February 2, 1990;
revision accepted for publication May 21, 1990.)