TABRA: A PROJECTIVE TEST FOR COMPUTER PERSONNEL RESEARCH. PRELIMINARY REPORT.

Dr. Lucy Zabarenko
University of Pittsburgh, School of Medicine

George F. Badger, Jr.
University of Pittsburgh, Computer Center

Ellen B. Williams
University of Pittsburgh

In their summary of research on the selection and evaluation of computer personnel, Mayer and Stalnaker (1968) cited five areas in which substantive research was needed. One of these was development of "new observational techniques especially into the personality characteristics" of computer personnel. They also stated: "Two relatively large studies have been made into the personality of programmers, but neither has been published. Both, however, tend to indicate that the programmer has a differentiable personality just as he was found to have a universal interest pattern." These findings enhanced our interest in assessing the role of personality factors in computer work. We have been concerned with emotional responses to computers, and especially those feelings which affect the individual's ability to learn and deal with the machines. The goal of this inquiry was to gather information which might be of use not only in the selection of personnel for training, but also to assist computer users and those designing training programs.

A number of sources have supported the idea that emotional factors are important in this area. Confirmatory reports come from instructors who teach computer skills at many levels (Badger, Muscarro, and Ferguson, 1968-69), and the large number of jokes, cartoons, and protests about the machines are an informal indicator (Antioch Nemesis of...., 1969). Recently, Williams (1969) has surveyed a selected sample of ten books on computer science. By analyzing the author's introductory statements, she was able to document some of the underlying emotions involved in the man-machine interface. A search of the literature and inquiries to selected experts have thus far uncovered no published study systematically recording data on affective responses to computers.

The plan of this presentation is in five sections: (1) rationale; (2) instrumentation; (3) experimental design; (4) results of pilot work, and (5) serendipity control.

I. RATIONALE

In approaching this large arena, we have chosen three methods of data collection: a projective technique, a loosely structured interview and nonparticipant observation. In this report of the pilot work, we shall consider only the first of these, the projective technique. The test is called TABRA, from the first syllables of the phrase tabula rasa. This has seemed appropriate since it is so often assumed that people come to work with computers with a clean slate while we believe instead that there are many predispositions toward feelings and attitudes. We shall say more about this later.

We may begin by making explicit the assumptions involved in choosing this kind of measure instead, for example, of more straightforward means such as an attitude inventory. The first of the assumptions is that it is often difficult and painful for people to talk about their feelings, and frequently, they may not themselves be fully aware of some of these feelings. However, when a neutral or unstructured stimulus is presented, responses can be obtained which give a better picture of a person's feelings and his state of mind. Tests which utilize this method have been part of the arsenal of psychological measurement for many years. The best known representatives are the Rorschach Test, in which the respondent is invited to tell what he sees in a group of inkblots, and the Thematic Apperception Test in which he is asked to tell a story about each of a group of pictures.

It is important to remember that responses in these tests are extremely varied and sometimes lengthy, and it is the job of the test developer to be able to decipher the meaning of the responses and make maximum use of them. In the choice of a projective test for this investigation, then, we have assumed that (1) partially structured stimuli can be constructed so as to elicit information about how people feel about computers, and (2) we will be able to decode these responses.

Those unfamiliar with such an endeavor may well wonder where one begins. We began with hypotheses about the emotions involved in computer work. These hypotheses had evolved from three basic sources: the scientific literature on affective responses to machines in general, clinical experience with computer personnel, and observations. As an example, let us illustrate the latter. It is a common experience to have a terminal cease functioning before a procedure can be completed. Even experienced programmers have told us that they

often have a brief feeling of guilt when this happens. One student expressed this dramatically saying, "The terminal is dead. Did I kill it?" This led to the hypothesis that many people, perhaps unknowingly or unwittingly, think of the machines as being vulnerable, living beings, and in turn this hypothesis led us to a set of overlapping items on the test.

Before beginning work we sought to deal with the issues of privacy and confidentiality which are of urgent priority in this kind of research. The official guidelines of the American Psychological Association (Sawyer and Schechter, 1968) were followed in all phases of the work. No names have been recorded and all respondents were given instructions something like this: "In order to protect you, no names are being recorded in connection with this research. Would you please select a five-digit number for identification purposes. In case you would like to have any feedback from the test, you might want to make it a number you can remember, because that is the only way we will be able to get the data." Parenthetically, the same method has been used in interview data.

II. INSTRUMENTATION

The test is in three parts. The first is a semantic differential test in which the subject is asked to mark a continuous line (Figure 1). The second is a sentence completion (Figure 2), and the third is a ranking in which the respondent is shown a list of familiar mechanical devices and three Rorschach inkblots and asked to judge their similarity to computers (Figure 3).

Respondents were encouraged to fill out the test as rapidly as possible and to avoid pondering over items. Most people finished the initial task in under 10 minutes. An inquiry followed, in which each item was reviewed and the respondent was asked to say more about his answer. The total test rarely took more than 30 minutes. At the end of the testing, respondents were asked for comments on and suggestions for the test.

III. EXPERIMENTAL DESIGN

The experimental design reflects our interest in the study of feelings as they affect learning of computer skills. We studied two groups of individuals, those who seemed to "take to" the work, Group I, and those who, in the absence of any intellectual deficiency, found it difficult or impossible to master, Group II.

TABRA - Form 1 10/16/68

Code: Age: Sex: Group:

Part A

Mark the solid line at the place which shows your opinion.

COMPUTERS ARE:

1. Alive——————————————————Dead
2. Crowded——————————————————Empty
3. Awake——————————————————Asleep
4. Straight——————————————————Tangled
5. Unfriendly——————————————————Friendly
6. Human——————————————————Mechanical
7. Full——————————————————Hungry
8. Wild——————————————————Tame
9. Harmless——————————————————Destructive
10. Dirty——————————————————Clean
11. Free——————————————————Contained
12. Democratic——————————————————Dictatorial

Figure 1

TABRA - Form 1 10/16/68

Part B

Complete these sentences to express your feelings. Do every one and make a complete sentence.

1. A mother_____

2. Inside computers_____

3. In general machines should_____

4. The color of computers_____

5. Terminals_____

6. I feel_____

7. Dreams and computers_____

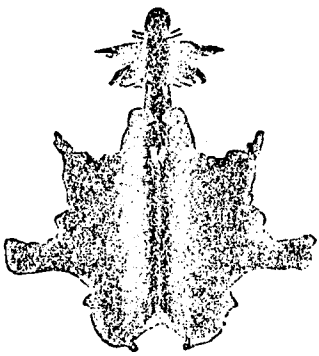
Figure 2

Part C.

Below are listed a number of mechanical devices. Beginning with 1 rank them according to their similarity to computers:

1. _____ Television
2. _____ Vacuum Cleaner
3. _____ Bicycle
4. _____ Car
5. _____ Radio
6. _____ Typewriter

Below are three pictures, A B and C. Beginning with 1 rank them in the order in which they remind you of computers.



A._____

B._____

C._____

Figure 3

Three criteria were used to decide group assignments: academic performance, professional competence, and personal statements. Academic performance was judged by work in a beginning computer science course. If a student did well and seemed to the instructor to be apt, he was assigned to Group I. If he withdrew or failed, he was assigned to Group II. If an individual was working with computers successfully as part of his daily professional tasks, he was classified as Group I. This subgroup included graduate students in computer science and in educational research. When an individual was definite in his aversion to or discomfort with computers, he was assigned to Group II. An example of this was a graduate student in a field allied to computer science who, after a 13-week course in computer science, stated, "I know that I will never get near one of those things again if I can help it. I hate it." Table 1 shows the data on group assignments by criteria and by sex.

Table 1

CRITERIA FOR GROUP ASSIGNMENT

Criterion Used	Group I		Group II	
	Male	Female	Male	Female
Academic	4	2	9	1
Professional	8	1		
Personal			3	2
Totals	12	3	12	3

The choice of groups at the extremes of aptitude rests on the assumptions that these groups might show most clearly the affective facilitation or interference we were curious about, and also might enable us to begin our work of correlating data from the test with specific, observable learning behaviors.

IV. RESULTS

The 30 protocols to be discussed here were collected over a period of 20 months; 17 were administered by the senior author (LMZ); 13 were administered by the junior author (EBW).

In a preliminary study like this one, the most valuable results for the researcher are those which can lead to improvements in instrumentation and design. The practitioner, however, is likely to be interested in data which might aid him in his daily work, even though the evidence is sparse and the conclusions tentative. We shall try to satisfy both of these informational needs in the material that follows. But we cannot state our caveat too loudly. The infant test whose genesis we chronicle today is very far from being ready for the rough and tumble of limited use, much less the gladiatorial combat of selection procedures. It will be some time before either of these is imminent and besides, we all know that infant mortality is always high in this field.

We shall present for each subtest, corrections in the test and design to be tried in future research and some preliminary conclusions.

GENERAL CONSIDERATIONS

In general TABRA was well received. A few students refused to participate because they were opposed to testing in general. For one graduate student the idea of a "personality test" induced frank panic. But for most people, rapport developed easily and the interview after the test was granted readily. However, the unstructured nature of the test was troublesome to those people for whom a test requires definite answers. Yet, apparently the test was just structured enough to make the elicitation of spontaneous thoughts and fantasies difficult for other respondents. We are hoping that altered instructions may alleviate some of these problems. Two specific tactics are to encourage respondents to work rapidly at first and tell us more in the inquiry, and to give specific sanction for the free play of imagination.

The use of the three criteria for group classification (academic, professional, and personal) has convinced us that these categories cannot be properly considered together. We look forward to a research design in which each category (academic, professional, and personal) could be compared within Groups I and II.

Part A

The semantic differential is a measurement technique which began in the study of meaning. It has, according to the editors of a recent source book (Snider and Osgood, 1969), "come to play an important role in almost all of the behavioral sciences". In clinical investigations it has been used in conjunction with other projective techniques (Friedman, Johnson, and Fode, 1964; Otten and Van Castle, 1963), in the study of multiple personality (Osgood and Luria, 1954), psychotherapy (Endler, 1961) and dreams (Moss, 1961). We could find no published account of a study using the semantic differential itself as a projective technique with an inquiry to probe responses.

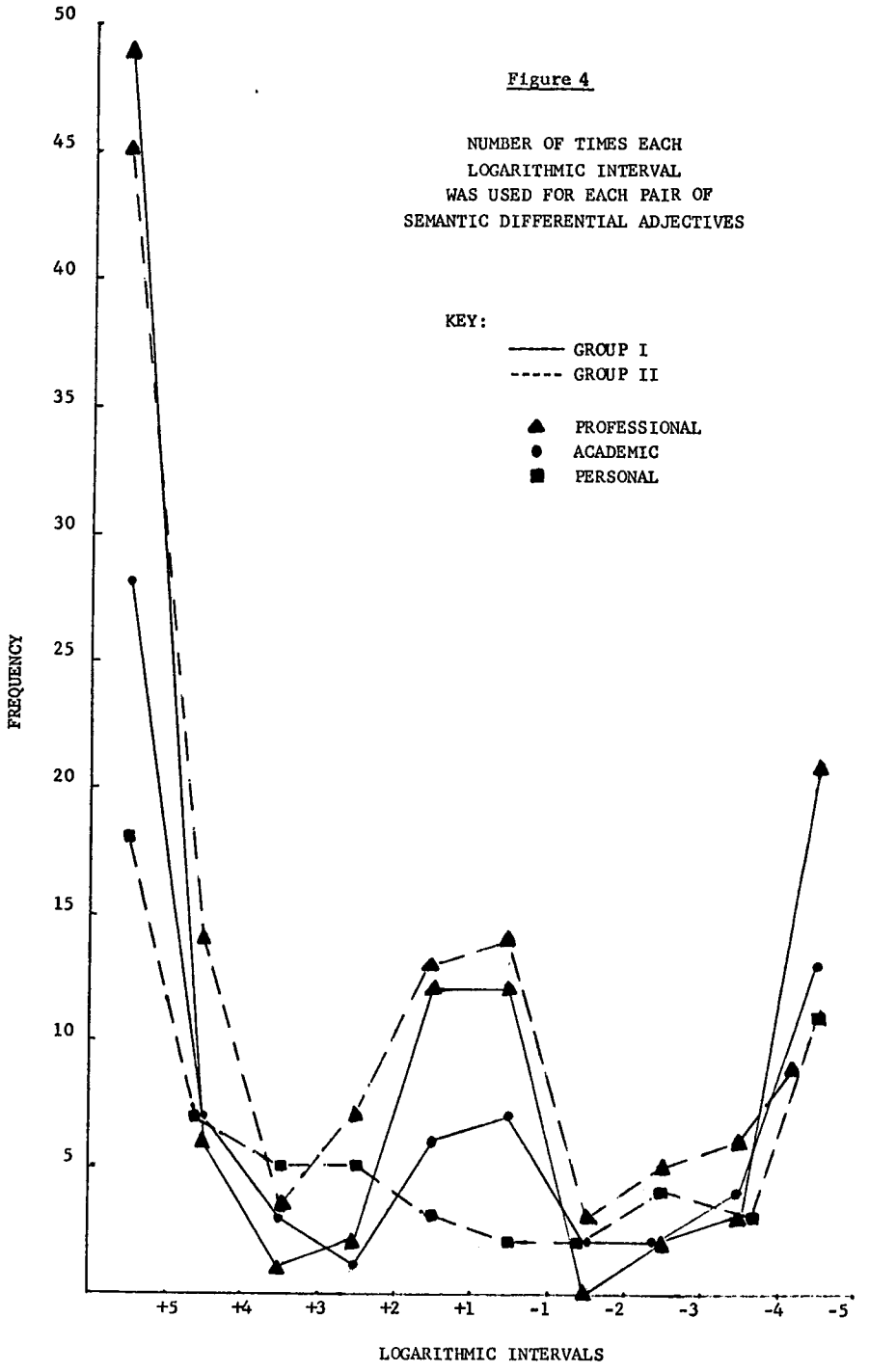
In TABRA Part A we used two main scoring schemes. In both, the reference point was the mid-point in the line between the two adjectives. One scale was marked off in the line between the two adjectives. One scale was marked off in equal intervals from this point; the second was marked off in a logarithmic progression. The rationale for this dual approach was straightforward. Equal intervals are the most frequently used in scoring the semantic differential; on the other hand there is reason to believe (Stevens, 1961, 1962) that human perception, even of complex entities, is better fitted to a more complex function. Also, when we considered the familiar tendency of ratings and judgments to cluster at mid-range, the log scale seemed to offer additional precision in this area.

All adjectives tending to picture computers in a favorable and/or realistic light were grouped at the positive end of the scale. These included: alive, crowded, awake, straight, friendly, mechanical, full, tame, harmless, clean, contained, and democratic. At the negative poles were: dead, empty, asleep, tangled, unfriendly, human, hungry, wild, destructive, dirty, free, and dictatorial.

The log scale helped to clarify some of the data (see Figure 4). Bearing in mind the different numbers of respondents in each classification group, and the fact that raw data had been plotted, the similarity between Groups I and II is striking: the largest numbers of responses are at the extreme positive end of the scale. Second highest frequencies are at the negative end of the scale, and in the center range. The only subgroup category which does not follow this pattern is composed of those five people classified in Group II because of their overt feelings of antipathy toward computers. They show no bunching in the center ranges and have

Figure 4

NUMBER OF TIMES EACH
LOGARITHMIC INTERVAL
WAS USED FOR EACH PAIR OF
SEMANTIC DIFFERENTIAL ADJECTIVES



almost equal tallies at both positive and negative extremes.

As we hoped, the inquiry data often contained clues about the reasons for the preferred adjective and the inquiry also revealed how respondents were thinking and feeling about computers. At first inspection there seems to be little difference between the two groups, but there are indications that the distinctions, when they emerge, will be complex and subtle. As an example, let us take two adjective pairs: hungry--full, and wild--tame.

For the pair hungry--full, four subjects in Group I saw the machine as hungry; for example, "I picture it s as just waiting to be fed". Two additional respondents saw the machine as hungry but passive. "They don't go after anything. You have to feed them information." Two subjects recalled an earlier adjective pair, crowded--empty, and attempted some synthesis. "They are crowded and yet empty. It's full of users and storage space." Four subjects rejected the item. Examples of what we define as rejection are, "I can't find any relevance", or "Does not apply, is irrelevant". Three respondents gave neutral statements to the full--hungry item, and two of these three statements stressed the computer's passivity. "It doesn't really care. He's neutral, he'd take anything he's given. He doesn't care." and, "When you're hungry you want something, but they can't get it themselves."

Of 15 Group II respondents, eight felt the machine was hungry. The reasons for this were its receptivity: "It's accessible to whatever you might want to put in"; its devouring characteristics, "They devour information", and its insatiability, "It's always hungry... never will be full". Four respondents indicated they thought the machine was full of various entities, parts, information, thoughts, programs, and even users. Three respondents were neutral because, as one of them said, "It's full of facts (but) there's a lot more they can have put into them."

Group II individuals seemed to be more able to express their feelings about the computer's appetitive condition, while Group I people were uncertain or unable to. One striking trend was that Group II people saw the machine as passive, unable to fend for itself; to Group I's it was active, devouring, and insatiable.

With the adjective pair wild--tame, seven subjects in Group I saw this as depending upon the user. "It's neutral. It can go either way depending on the user.

It's tame in the sense of restricting, wild in the sense of freedom." Four of these seven subjects saw this dependence in the light of some sort of struggle for control. "They're not tame. Tameness means responsive to a whim. The computer is like an animal that is carefully trained and cared for; not a dog, rather like a spirited horse." One of these four saw the machine-man interface as a battle. "I can get it to to do what I want in the end. It wins because it wants commas. But I always outsmart it." One subject saw the machine as tame, two saw it as wild. "Things are always going on, lights flashing, they are always moving." Three subjects rejected the item and two more were neutral.

On the other hand, six respondents in Group II stated emphatically that the computer was tame by virtue of its subservient position and its logical systematic nature. "They don't do anything you don't tell them to do." Four subjects recalled occasions on which a computer had been "wild", indicating the importance to them of the fact that the machine could get out of control. "A couple of times one went haywire. For one error I made it returned four pages of error messages. The machine messed up." Two respondents felt that the computer was wild: "It never worked well for me." There was one rejection of the item and two neutral responses. "Neutral. It was wild at times. I see it as frantic, but not all the time. Sometimes it's dormant."

This item seemed to tap feelings about control, not only who shall run things--the operator or the computer--but also the more troublesome idea of the machine being out of control.

Part B

In the *Handbook of Projective Techniques* (1965) the editors say of sentence completion methods that they are "a valid test, generally speaking, and probably the most valid of all the projective techniques reported in the literature". After an extensive review of the literature, Goldberg (1965) stated, "The sentence completion is a valuable instrument in the assessment of personality. It compares favorably to other standard instruments." He also noted that, "Critics...while citing its psychometric advantages, have tended to regard it as undynamic and of relatively little use for general clinical purposes." He suggested this relative unpopularity may have to do with the fact that, "much research has suggested that subjects can exercise considerable conscious control over their responses."

We believe this feature, a disadvantage in clinical psychology, may be an asset in this study. Since much learning takes place at conscious levels, the way in which affects are coped with at these levels is of great interest.

Pilot work has suggested changes in some of the sentence completion items. Four items, those dealing with mother, color, dreams, and directly with feelings, need to be focused more exactly on computers. For example, Item 1, instead of reading "A mother_____" might be changed to read "A computer's mother_____." Or Item 6, "I feel_____" might be changed to read "A computer's feelings_____." The alterations in the items will also be consonant with the hypothesis that the more abstract, i.e., less personal, the sentence completion item, the more it is likely to elicit a projective response.* Another change will be in Item 5, which uses the word "terminals". In the computer dialect most familiar to us, this is the term applied to the apparatus employed to transmit messages from some distance directly to the computer. To some respondents it meant railway stations, fatal illness, the attachment point for an electrical circuit, or simply, the end.

An examination of the responses to specific items in Part B showed some striking similarities as well as leads to differences. Item 3, "In general machines should_____," elicited almost identical responses. Thirteen individuals from Group I and 12 from Group II answered that what a machine should do is behave properly. Two of those responding this way in Group II used exactly the same words, "In general machines should be seen and not heard." One person in each group felt that machines should be looked after.

On the other hand, Item 5, the one concerning terminals, drew out somewhat different patterns. Eleven of the 15 respondents in Group I took our meaning from the item, and five of these responded in a manner which was full of feeling. Three responses were negative--terminals were untrustworthy, noisy, always being used by somebody

---

*"The projection hypothesis, which generally serves as the theoretical rationale underlying the use of projective tests, was made most explicit by Frank. Presented simply, the projection hypothesis states: When an individual is forced to impose meaning or order on an ambiguous stimulus complex, his response is a 'projection' of his 'feelings, urges, beliefs, attitudes, and desires...'" (Murstein, 1965, p. 781)

else; two responses were positive--terminals were a godsend, very convenient. In Group II only eight people responded to the word terminal as having to do with computers and none of these responses revealed much affect.

Item 6 revealed even clearer differences. In response to "I feel_____," only two people in Group I made a link to computers, and these responses were rather coolly descriptive; for example, "I feel I've had a limited exposure to computer utilization." Eight people from Group II connected feelings and computers, and four of the eight were frankly negative. For example, I feel biased against machinery, some hostility toward computers, completely alienated by computers, that the general processes and concepts of computers will never be clear to me.

Using a more quantitative approach, we found that Group I individuals used fewer words to complete sentences. In Group I the mean number of words per item per person was approximately five. Item 4, the one dealing with color, averaged 3.36 words. It often elicited a one-word answer. Group II respondents used over six words per item. Their most tersely answered question was number 7, the one on dreams, which for the most part they saw as unrelated to computers.

Part C

The mechanical devices were ranked by criteria having almost entirely to do with electronic or mechanical complexity, and the two groups showed no manifest differences. The bicycle was clearly least like a computer, the vacuum cleaner next and then the car. Typewriter, TV, and radio were in closer contest for the three machines most similar to the computer. Those who worked well with computers ranked the typewriter as being most similar to the computer, whereas those who had difficulty working with computers considered the TV to be most similar.

Reactions to the inkblots had a good bit of patterning, but not sufficient to allow reporting here. One of the improvements we are most anxious to introduce is a more extensive inquiry in the next version of the test. This endeavor may pose some problems, as many people were threatened by the frankly unstructured nature of the stimuli. Also, many respondents know that inkblots are used to "probe the personality", and consequently their guard is up.

V. SERENDIPITY CONTROL

Having avowed at the outset our inability to present conclusions, we should like to end this presentation with a brief statement of what we have learned and how we should like to learn more. In planning future research in an area as unmined as this, one can predict that serendipity will be rampant and thus must be thoughtfully provided for.

The underlying psychological problems in dealing with the computer seem to hinge on the issues of whether the machines are seen as active or passive, capable of being controlled or wild, and whether they are essentially benign or evil. Group II people seem a little more able to express these conflicts than those in Group I, but essentially the difficulties are the same for both groups. What is different are the ways in which these disturbing ideas are managed. This is one area we hope to study next.

There are two more research areas which seem promising as sources of data. One is a study of the computer center as a subculture or moiety or tribal alliance, whichever it turns out to be. The second is an investigation of the specific work pathology, the syndrome often known as "the computer bum".

If one sets out to measure the gait of an elephant, it would be better not to use a 12-inch rule, especially if this 12-inch rule is printed on a strip of elastic. On the other hand, it is better to start, even with imperfect implementation, than to be overwhelmed by the task. We feel the greatest hazard in research in this area is a tendency toward oversimplification. Our hope is that eventually we may supply more adequate tools for the work.

BIBLIOGRAPHY

"Antioch Nemesis of Witch Hunt Era Heads Drive to Frustrate Computers", *The Antiochian*, 40, No. 6, 1969.

Badger, G., J. Muscarro, and R. Ferguson, *Personal Communications*, 1968-1969.

Endler, Norman S., "Changes in Meaning during Psychotherapy as Measured by the Semantic Differential", *Journal of Counseling Psychology*, 8, 1961, pp. 105-111.

Friedman, C. Jack, Carleton A. Johnson, and Kermit Fode, "Subjects' Descriptions of Selected TAT Cards via the Semantic Differential", *Journal of Consulting Psychology*, 28, 1964, pp. 317-325.

Goldberg, Philip A., "A Review of Sentence Completion Methods in Personality Assessment". In, Murstein, Bernard I. (Ed.), *Handbook of Projective Techniques*, Basic Books, Inc., New York, 1965, pp. 777-821.

Mayer, D. B. and A. W. Stalnaker, "Selection and Evaluation of Computer Personnel--The Research History of SIG/CPR", *Proceedings, 1968 ACM National Conference*, 1968, pp. 657-670.

Moss, C. Scott, "Experimental Paradigms for the Hypnotic Investigation of Dream Symbolism", *International Journal of Clinical and Experimental Hypnosis*, 9, 1961, pp. 105-117.

Murstein, Bernard I. (Ed.), *Handbook of Projective Techniques*, Basic Books, Inc., New York, 1965.

Osgood, Charles E. and Zella Luria, "A Blind Analysis of a Case of Multiple Personality using the Semantic Differential", *Journal of Abnormal and Social Psychology*, 49, 1954, pp. 579-591.

Otten, Mark W. and R. L. Van Castle, "A Comparison of Set "A" of the Holtzman Inkblots with the Rorschach by Means of the Semantic Differential", *Journal of Projective Techniques and Personality Assessment*, 27, 1963, pp. 472-460.

Sawyer, J. and H. Schechter, "Computers, Privacy, and the National Data Center: The Responsibility of Social Scientists", *American Psychologist*, 23, No. 11, 1968, pp. 810-818.

Snider, James G. and Charles E. Osgood (Eds.), *Semantic Differential Technique. A Source Book*, Aldine Publishing Company, Chicago, 1969.

Stevens, S. S., "To Honor Fechner and Repeal His Law", *Science*, 133, 1961, pp. 80-86.

Stevens, S. S., "The Surprising Simplicity of Sensory Metrics", *American Psychologist*, 17, 1962, pp. 29-39.

Williams, E. ., *An Interpretation of the Opinions of Selected Authors of Computer Books about the Relationship between Man and the Computer*, Mss. 1969.