1990

Presence/Absence of an Observer in a Word Association Test

Morton Isaacs

Rochester Institute of Technology

Kathleen Chen

Rochester Institute of Technology

Follow this and additional works at: <http://scholarworks.rit.edu/article>

Recommended Citation

Isaacs, M., & Chen, K. (1990). Presence/Absence of an Observer in a Word Association Test. *Journal of Personality Assessment*, 55(1-2), 41-51. <https://doi.org/10.1080/00223891.1990.9674045>

This Article is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Articles by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

PRESENCE/ABSENCE OF AN OBSERVER IN A WORD ASSOCIATION TEST

Morton Isaacs and Kathleen Chen
Professors, Rochester Institute of Technology
Journal of Personality Assessment, 1990, 55 (1& 2) 41-51. Copyright
1990. Lawrence Erlbaum Associates Inc.

ACKNOWLEDGMENT Research for this article was partially sponsored
by a grant from the College of Liberal Arts, Rochester Institute of
Technology, Rochester, NY.

- - - - -

ABSTRACT Forty-eight subjects (24 men, 24 women) responded to 20
word stimuli presented by a microcomputer program that recorded
each subject's response latency and response to each stimulus word.

Emotionality of stimulus word and presence/absence of an observer
were independently varied. Two-way analysis of variance (ANOVA)
showed significant difference in latencies attributable to the
emotionality of the stimulus word, whereas neither the
presence/absence of the observer nor the interaction effect was
significant. The results suggest the possible use of latency analysis as
a tool in situations where an objective, statistical, and easily
obtainable evaluative measurement of the level of emotionality of
various verbal stimuli is desired.

INTRODUCTION

The clinical application of word association tests was stimulated
by the psychoanalytic movement in a technique that antedated the
flood of projective tests by more than half a century. Jung
(1918/1919) was the major contributor to the systematic
development of association tests. He selected stimulus words to
represent emotional complexes and analyzed the responses in terms
of reaction time (RT), content, and physical expression of emotional
tension by the subject. The use of word association techniques as a lie
detectors was also initiated by Jung and subsequently subjected to
extensive research (Lindsley, 1955). The rationale of its use based on
the aforementioned studies is supposedly to uncover areas of
emotional conflict often associated with the conscious telling of an
untruth.

A different approach to using word association tests was shown in the work of Kent and Rosanoff (1910), whose test used objective scoring and developed statistical norms. Rather than choosing words expected to be emotionally laden, based on the experimenter's knowledge of the subject, they chose common, neutral stimulus words which could be expected to evoke similar associations from many people. In scoring the test, the median frequency value of the responses the subject gives to the 100 stimulus words is used as an index of commonality. Any responses not found in the normative tables were considered as unique.

The diagnostic use of word association tests as a tool declined with the knowledge that response frequency varies with age, socioeconomic and educational level, sex, cultural background, and other personality factors (Kurcz, 1966; Palermo & Senkins, 1965; Victor, Kostyla, & Jones, 1985), because it appeared that proper interpretation of scores required norms on many subgroups and additional information about the subject.

Nevertheless, the Kent-Rosanoff test has retained its position as a standard laboratory technique. Additional norms have been gathered and the technique has been extensively employed in research on verbal learning, verbal behavior, and personality (Bastian, 1961; Blank, 1980; Goldfarb & Halpern, 1984; Hormuth, 1982; Jenkins & Russell, 1952; O'Neill, 1953; Palermo & Senkins, 1965; Postman & Keppel, 1970; Rattenbury, 1983).

In all of the studies just mentioned, however, it is the response word that has been recorded and analyzed in detail whereas the response latency has not been rigidly researched (e.g., the standard set of directions presents the subjects with a list of words with one or more blank spaces following each word, and the subject is instructed: "... After each word write the first word that it makes you think of. Start with the first word; look at it; write the word it makes you think of; then go on to the next word.... Work rapidly until you finish all 100 words."- Postman & Keppel, 1970, p. 2). How rapidly the subject responds to any one word, or for that matter whether the subject glances up to a previous response before writing down subsequent ones, is not controlled. Relatively few studies strictly record latencies, and those that do tend to use these latencies as a tool to test assumptions about connectivity among response words (Pollio, 1966; Tecce & Glassco, 1965; Wiggins, 1957) rather than focus on the distribution and analysis of the latencies themselves.

Small differences in response latencies can be confounded with differences in the recording process (e.g., imprecision of the timepiece, errors in recording the latencies) and possible distraction of the subject during the stimuli presentation by the experimenter or the recording process itself.

Using a microcomputer, however, it is possible to record both the latency of response and the response itself with great accuracy and without the intervention of another human, thus permitting strict statistical comparisons of unconfounded response latencies. There is, however, some question as to whether human-gathered responses differ from those collected by computer. Canoune and Leyhe (1985) found that subjects that were administered the Gordon Survey of Interpersonal Values showed a significantly higher level of conformity and lower level of self-recognition than subjects who had the same test administered by a computer. McCoullough and Wenck (1984) found that people are more open and honest with a computer than with a human. More directly to the issue, Nakamura and Wright (1965) reported that RT increased to word association stimuli in the presence of an experimenter. On the other hand, Koson, Kitchen, Koshen, and Stodolosky (1970) found no significant difference between responses to humans and those to computer interviewers for high- or low anxiety test items, nor did Rezmovic (1977) for moderate anxiety test items.

Two reasons might be posited to account for any differences found when using computer versus human data gathering. It might be that the mere presence of a human actively auditing the incoming information leads the subject to modify this information in an Evaluation Apprehensions effect (Baumeister, 1982), or it may be that the subjects are aroused by the presence of another person ("social facilitation"; Zajonc, 1965) and this general arousal heightens individual differences in response latencies.

To obtain a comparison of differences in subjects' responses due to the presence or absence of a human, it would be necessary to hold all other factors constant. It was therefore decided to set up two conditions. In the absence of observer condition, the assistant set up the experiment and left the room, and in the presence of observer condition the assistant remained in the room purportedly to write down the subject's responses to the stimuli, actually collected by the computer program in both cases. It was believed that having an observer record the subject's responses would be a close analog of

the usual clinical use of the word association presentation, but with the dual advantages of accuracy in data collection and avoidance of the possible confounding of subject's typing speed against verbal response speed in the response latency measure.

METHOD

Subjects

Subjects were 24 male and 24 female student volunteers at the Rochester Institute of Technology, with a mean age of 21.3 (SD = 2.1). The 10 most emotional and 10 least emotional words were chosen from a list based on galvanic skin response (GSR) measurements performed by Smith and Brown (1922). A program was created in BASIC for the Apple II+ microcomputer which collected some subject background data, presented the stimulus words in a previously determined random order sequence, and recorded the subject's response to each stimulus and its latency in hundredths of a second. The subject's name was also requested and temporarily saved to serve as one of the word stimuli presented to the subject during the experiment, but for confidentiality reasons was not permanently recorded.

Procedure

Each subject was assigned sequentially either to the "Presence" or "Absence" of observer condition. The following instructions were then displayed on the computer screen:

DIRECTIONS

This will be a word association test. We wish to see how long it takes you to think up a response to a word and what that response is. The word "ready" will appear on the screen and the computer will beep. Then a stimulus word will show on the screen replacing the word "ready." Think of a response. As soon as you have the first one, type it into the computer. Enter it by pressing the "Return" key. You can correct any typing errors by using the back-arrow key to go back to the error and then retyping from there. When finished with your response, press the return key for the next word. The first three words are practice ones; the experimenter will stay with you to help with any questions. After that, the experimenter will . . .

(Presence subjects) be recording your responses but is not permitted to talk with you during the experiment.

(Absence subjects) leave the room while you respond to the actual words.

After three practice words, the screen displays the following:

(Presence subjects) Press the space bar to continue the experiment. Remember, the researcher may not talk to you once the experiment has begun.

(Absence subjects) The researcher will now leave the room; then press the space bar to continue with the experiment.

Response latencies were calculated by the program in centiseconds from the time the stimulus word was first displayed on the screen to the press of the first response letter by the subject. At the conclusion of the last stimulus word, the program saved the response latencies and the responses on a disk file.

RESULTS

The results of the ANOVA on the latencies for the two conditions (stimulus emotionality and presence of observer) and their interaction indicated that neither the presence/absence of observer condition nor the interaction effect was significant ($F = 0.287$ and $F = 0.489$, $p > .05$); emotionality of stimuli, however, was significant, with high-emotional words displaying significantly longer mean latencies than low-emotional words ($F = 25.19$, $p < .001$).

Table 1 lists the stimulus words in order of the mean response latencies from longest to shortest in one hundredths of a second, the order of presentation of the stimulus words, their word frequencies as listed in Kucera and Francis (1967, pp. 139-263), and the response commonality (how many of the 48 subjects gave the most common response).

The mean response latencies were correlated with the word presentation order to ascertain whether subjects might have responded faster as they gained more experience with the computer presentation of the stimuli. The Pearson correlation was not significant ($T = 0.28$, $df = 18$, $p > .05$), indicating that the presentation order of

the 20 stimulus words did not predict their mean latencies well nor did the correlation of mean latencies with word frequencies ($r = -0.17$, $df = 17$, $p > .05$) or with response commonality ($r = -0.155$, $df = 18$, $p > .05$)

TABLE 1

Stimuli, Mean Latency of Response, Stimulus Presentation Order, Word Frequency, Response Commonality, and Number of Extreme Responses

Stimulus Word	Mean Latency*	Present Order		Word Freq**	Response Commonality	Number of Extremes
Proud	2.35	20	50	7	17	
Divorce	1.86	8	29	6	16	
Marry	1.71	7	18	5	10	
Name***	1.62	10	—	39	7	
Flower	1.38	13	23	7	5	
Afraid	1.37	19	57	26	7	
Dance	1.34	16	90	7	6	
Glass	1.27	9	99	13	5	
Swim	1.19	18	15	14	4	
Love	1.18	6	232	24	5	
Give	1.14	11	391	29	4	
Woman	1.13	12	224	25	7	
Wound	1.13	14	28	14	2	
Berry	1.10	2	9	15	2	
Kiss	1.08	4	17	9	5	
Carrot	1.07	1	1	12	2	
White	1.07	5	365	29	4	
Hunger	1.04	3	17	12	1	
Pond	0.97	15	25	16	0	
Pencil	0.90	17	34	34	0	

* in seconds.

** Based on data reported by Kucsa and Francis (1960).

*** Each subject was presented with his or her name as the stimulus word. Frequency data are not applicable and all calculations were based on $N = 19$.

EXTREME SCORES

Examination of the subjects' raw scores showed that some subjects responded with an unusually long latency to one or more stimulus words when compared to their latencies to the balance of the stimuli. This was not unexpected, as a long response latency is one of

the cues therapists use to identify emotionally arousing stimuli, and a neutral stimulus to the majority of subjects can be idiosyncratically emotional to a few (Colby, 1960). These extreme scores were operationally defined as any score whose latency fell beyond 2.33 standard deviation units above the mean in the subject's distribution (one-tailed test, .99 significance level). The highest extreme score, if one existed, was identified for a subject and excluded from the subject's distribution. A revised mean and standard deviation for the subject's remaining latencies were then calculated, and the procedure performed recursively for the subject until none of the subject's remaining response latencies exceeded the criterion. The median of extreme scores for all subjects was 1.73, with a range from 0 to 8.

Identifying extreme scores per subject offered a completely different method for checking the rank of emotionality of each of the stimulus words. Because more emotional words should result in longer latencies for most subjects, the higher the emotionality of the word the more likely it should be to cause a delay beyond the criterion for the subject; therefore, after identifying extreme words by their latency of response in the subject's distribution, the number of extreme scores excluded per word should differ for high- and low-emotional words.

Table 2 lists the stimulus words in descending order by extremity scores derived from the recursion of the subjects' individual distributions (ordered within ties according to the mean latency scores), and the original and revised latency means and standard deviations.

TABLE 2

Stimulus Words, Number of Extreme Scores per Stimulus Word, Emotional Category, and Means and Standard Deviations for Both the Original and Recursively Revised Mean Latency Scores

Stim Words	Extreme Scores	GSR Emot Category	Original		Recursive	
			Mean	SD	Mean	SD
Proud	17	High	2.35	2.61	1.30	0.49
Divorce	16	High	1.86	1.50	1.37	0.39
Marry	10	High	1.71	1.18	1.35	0.55
Name'	7	High	1.62	1.96	1.11	0.48
Afraid	7	High	1.37	1.10	1.01	0.41
Woman	7	High	1.13	0.94	0.83	0.46

Dance	6	High	1.34	0.96	1.10	0.36
Flower	5	Low	1.38	0.97	1.15	0.45
Glass	5	Low	1.27	0.80	1.07	0.40
Love	5	High	1.18	0.60	1.07	0.47
Kiss	5	High	1.08	0.66	0.90	0.31
Swim	4	Low	1.19	0.65	1.04	0.29
Give	4	Low	1.14	1.15	0.88	0.39
White	4	Low	1.07	1.04	0.82	0.26
Wound	2	High	1.13	0.52	1.07	0.44
Berry	2	Low	1.10	0.49	1.04	0.40
Carrot	2	Low	1.07	0.40	1.03	0.35
Hunger	1	Low	1.04	0.54	1.00	0.45
Pond	0	Low	0.97	0.37	0.97	0.37

*Each subject was presented with his or her name as the stimulus word.

A one-way ANOVA of the number of extreme scores per word, based upon the prior division into high- and low-emotion words, yielded $F(1, 18) = 12.08$, $p = .003$ (one-tailed), with those stimuli classified as highly emotional based on the GSR showing significantly more extreme scores than those classified as low.

DISCUSSION

Previous studies of the effect of stimulus emotionality on response latencies in a word association test using timed discrete responses in an oral modality have consistently reported that emotional words are responded to more slowly than neutral words (Bodin & Geer, 1965; Brown, 1965; Epstein & Fenz, 1962; Jacobs, 1955; Powell, 1955; Smock & Thompson, 1954; Weintraub, Silverstein, & Klee, 1960). In this study, where the stimuli were presented visually and the responses collected by typing on a computer keyboard, response latencies were also significantly longer for the higher emotional words than for the lower ones, an important indication that measuring response latencies by computer without human intervention yields valid information.

One interpretation of these findings is that emotional arousal derives from social pressure, or evaluation apprehension. Geer and Mollenauer (1964) suggested, on the basis of their research, that the slower RT found for traumatic stimuli in their research may be due to the subject's seeking for a response word that is socially desirably and hence from a different class of responses than the negative ones usually associated with it.

A different possibility is that lengthened latencies are due to heightened internal arousal due to anxiety-laden stimuli rather than to external pressure. Smock and Thompson (1954), for examples using the difference between RTs to affective and to neutral stimuli as a measure, found the greatest response latency in students for culturally taboo stimuli, with sibling rivalry, oedipal, anal, Ad oral stimuli following. That sibling rivalrous stimuli took longer to respond to than Oedipal, anal, and oral stimuli seem to indicate that it was internal rather than the social desirability of the response that lengthened the time.

Compromise position is that of Pollio and Lore (1965) who attributed this lengthening to the concept of Generalized conditioned avoidance responses." They believed that this increased RT to emotional words is a result of that part of the term (Osgood, Suci, & Tannenbaum's, 1957, concept), that consists of an implicit avoidance response, part of the response originally evoked by the significance for which the sign (word) stands. Implicit avoidance responses are responsible for the disruption of the usual response set, and the resulting necessity to adopt a different response strategy, such as searching for a socially acceptable response, adds to the RT.

If the fear of social disapproval was a significant force in lengthening the latencies, then our study should have found either a direct effect or an interaction effect of emotionality of stimuli and the presence/absence of an observer. No such effect was found. Our data therefore supports the second explanation, that the lengthened latencies are due to internal conflict and indicate that the measurement of response latency is potentially much more useful than would be the case if there were either a simple or an interactive administrator factor influencing a subject's response delays. Latencies in responding to emotionally arousing words appear to be situationally independent: for example, the characteristics of the test administrator or the style of test administration may be irrelevant.

A possible clinical use might then be established by calculating both the number of extremes generated by the individual to a set of words and which words were responded to extremely in terms of the respondent's own mean latency and standard deviation for the set, avoiding the difficulties inherent in trying to generate group means and standard deviations of emotional words that might differ from subgroup to subgroup within a culture.

The categories that assume a negative or avoidance characteristic appear to be determined, at least in part, by the developmental level of the subject. Erikson (1950) hypothesized that RT to stimuli of different conflict areas should be related to changing stages of emotional development. Supporting this hypothesis is Powell's (1955) study showing that changes in the conflict areas are age related, with religion and vocational outlook as the most conflicted area for preadolescents and adults respectively. The results of our study are consistent with Erikson's hypothesis. Neutral stimuli (e.g., hanger, carrot, berry) may be rank ordered at the oral development level and hence responded to quickly. Some of the words responded to with the greatest latency in this study (e.g., proud and the subject's own name) are reflective of the identity stage of later adolescence with the other high-emotional words (e.g., Move, divorce, kiss, and dance) falling into the next stage in Erikson's sequence, that of intimacy.

Indeed, the one major exception to higher emotional stimuli being correlated with greater numbers of subjects waiting extremely long times to respond is the word "wound", classified in the high-emotional category but with only two subjects giving extreme latencies to it. It may be that this is due to its being a homophonic word (two pronunciations, each with a different meaning), or it may be that the word wound has lost its emotionality over the intervening time. Further research would be needed to distinguish these hypotheses.

The finding that the presence/absence of the human observer did not produce any significant difference also offers support to those who have found no difference in results from computer versus human administration of tests (e.g., Koson et al., 1970; Rezmovic, 1977). It appears in our study that subjects concentrated on the task rather than on the presence or absence of a human observer during the answering process; qualitative reports from the observer of the subject reactions bore this out, because the observer reported that none of the subjects asked any questions nor made any remarks during the actual presentation of the stimuli, but rather remained focused on the task.

The independence of response latencies from word frequency and word presentation order is particularly interesting. Postman (1970) noted that word frequency is, of course, a major variable that is manipulated or must be controlled in most investigations using verbal materials" (p. 24), and indeed, before responses to stimulus

words in a word association test can be used, it is necessary to establish the word frequency for each subgroup. If response latency is independent of word frequency but the actual response word is not, then a valuable aid in the therapeutic identification of emotionally charged words for a patient, regardless of subgroup, may be available.

REFERENCES

Bastian, J. R. (1961). Associative factors in verbal transfer. *Journal of Experimental Psychology*, 49, 229-240.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, 91, 3-26.

Blank, T. O. (1980). Observer and incentive effects on word association responding. *Personality and Social Psychology Bulletin*, 6, 267-272.

Bodin, A. M., & Geer, J. H. (1965). Association responses of depressed and non-depressed patients to words of three hostility levels. *Journal of Personality*, 33, 392-408.

Brown, W. P. (1965). Emotional indicators in word association. *British Journal of Psychology*, 56, 401-412.

Canoune, H. L., & Leyhe, E. W. (1985). Human versus computer interviewing. *Journal of Personality Assessment*, 49, 103-106.

Colby, K. M. (1960). Experiment on the effects of an observer's presence on the image system during psychoanalytic free association. *Behavioral Science*, 5, 216-232.

Epstein, S., & Fenz, W. D. (1962). Theory and experiment on the measurement of approach-avoidance conflict. *Journal of Abnormal and Social Psychology*, 47, 230-235.

Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

Geer, J. H., & Mollenauer, S. O. (1964). Meaning class and affective content in word association. *Psychological Reports*, 15, 900.

Goldfarb, R., & Halpern, H. (1984). Word association responses in normal adult subjects. *Journal of Psycholinguistic Research*, 13, 37-55.

Hormuth, S. E. (1982). Self-awareness and drive theory: Comparing internal standards and dominant responses. *European Journal of Social Psychology*, 12, 31-45.

Jacobs, A. (1955). Formation of new associations to words selected on the basis of reactiontime-GSR combinations. *Journal of Abnormal and Social Psychology*, 51, 371-377.

Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology* 47, 818-821.

Jung, C. G. (1919). *Studies in Word association* (F. Eder, Trans.). London: Heinemann, Moffat Yard. (Original work published 1918)

Kent, G. H., & Rosanoff, A. J. (1910). A study of association in insanity. *American Journal of Insanity*, 67, 37-96, 317-390.

Koson, D., Kitchen, C., Koshen, M., & Stodolosky, D. (1970). Psychological testing by computer: Effect on response bias. *Educational and Psychological Measurement*, 30, 803-810.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, R.I: Brown University Press.

Kurcz, I. (1966). Inter-language comparison of word-association responses. *International Journal of Psychology*, 1, 151-161.

Lindsley, D. B. (1955). The psychology of lie detection. In G. T. Dudyha (Ed.), *Psychology for law enforcement officers* (pp. 89-125). Springfield, IL: Thomas.

McCullough, C. S., & Wenck, L. S. (1984). Current microcomputer applications in school psychology. *School Psychology Review*, 13(4), 429-439.

Nakamura, C. Y., & Wright, H. D. (1965). Effects of experimentally induced low drive, response mode, and social cues on

word association and response speed. *Journal of Experimental Research in Personality*, 1, 122-131.

ONeill, W. M. (1953). The effect of verbal association of tachistoscopic recognition. *Australian Journal of Psychology*, 5, 42-45.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Palermo, D., & Jenkins, J. J. (1965). Sex differences in word associations. *Journal of General Psychology*, 72, 77-84.

Pollio, H. R. (1966). *The structural basis of word association behavior*. The Hague, Netherlands: Mouton.

Pollio, H. R., & Lore, R. K. (1965). The effect of a semantically congruent context on word association behavior. *Journal of Psychology*, 61, 17-26.

Postman, L. (1970). The California norms: Association as a function of word frequency. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 241-320). New York: Academic.

Postman, L., & Keppel, G. (Eds.). (1970). *Norms of word association*. New York: Academic.

Powell, M. (1955). Age and sex differences in degree of conflict within certain areas of psychological adjustment. *Psychological Monographs*, 69(2, Whole No. 387).

Rattenbury, F. R. (1983). Associative disturbances in schizophrenia, schizoaffective disorders, and major affective disorders: Comparisons between hospitalization and one-year follow-up. *Journal of Consulting and Clinical Psychology* 51, 621- 623.

Rezmovic, V. (1977). The effects of computerized experimentation on response variance. *Behavior Research Methods and Instrumentation*, 9, 144-147.

Smith, W., & Brown, W. (1922). *The measurement of emotion*. New York: Harcourt, Brace.

Smock, C. D., & Thompson, G. G. (1954). An inferred relationship between early childhood conflicts and anxiety response in later life. *Journal of Personality*, 23, 88-98.

Tecce, J. J., & Glassco, J. A. (1965). Word association time as a function of anxiety (drive) and response competition. *Psychological Reports*, 16, 40.

Victor, C. I., Kostyla, S., & Jones, R. (1985). Conservatism as a factor of sexual responding to a word-association test. *Journal of Social Psychology*, 125(2), 275-276.

Weintraub, W., Silverstein, A., & Klee, G. D. (1960). The corrections of deviant responses on a word association test. *Archives of General Psychiatry*, 3, 17-20.

Wiggins, S. S. (1957). Two determinants of associative reaction time. *Journal of Experimental Psychology*, 54, 144-147.

Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 269- 274.