1998

Effect Of Item Clarity And Probability Of Item Endorsement On Response Latencies On Personality Test Items

James Robert Kinney
Indiana State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

CURRICULUM VITAE

James Robert Kinney was born on January 7, 1959 in Rochester, New York. He grew up in neighboring Greece, New York, and graduated from Greece Arcadia High School in 1977. He attended the State University of New York at Buffalo from 1985 to 1990. Mr. Kinney received a Bachelor of Arts degree from the State University of New York at Buffalo in 1990. He came to Indiana State University in the Spring of 1991 and began graduate studies in school psychology. Mr. Kinney received a Master of Education degree in 1994. He completed a school-based internship at the Covered Bridge Special Education District in 1994. In the fall of 1997, Mr. Kinney completed a final clinical internship at Crestwood Children's Center in Rochester, New York. Mr. Kinney is currently employed as a school psychologist in the Rochester City School District.

EFFECT OF ITEM CLARITY AND PROBABILITY OF ITEM ENDORSEMENT ON RESPONSE LATENCIES ON PERSONALITY TEST ITEMS

A Dissertation

Presented to

The School of Graduate Studies

Department of Educational

and School Psychology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

James R. Kinney

May, 1998

APPROVAL SHEET

The dissertation of James R. Kinney, Contribution to the School of Graduate Studies, Indiana State University Series III, Number 729, under the title *Effect of Item Clarity and Probability of Item Endorsement on Response Latencies on Personality Test Items* is approved as partial fulfillment of the requirements for the Doctor of Philosophy Degree.

April 20, 1998
Date

Director of Dissertation

Committee Member

Committee Member

Committee Member

Committee Member

April 23, 1998
Date

For the School of Graduate Studies

ABSTRACT

In this study, regression analysis was used to examine the affects of item clarity and probability of endorsement on response latencies of 60 undergraduates responding on personality test items from the Sixteen Personality Factor Questionnaire (16 Personality Factor test) (Cattell, Eber, & Tatsuoka, 1970, 5th Edition). Response latencies to personality test items, though frequently studied, have yet to be operationally utilized in the interpretation of personality tests. Forty-four items from the 16 Personality Factor test were selected. Each item from the original test was rated for clarity and matched with an emotionally neutral statement with an equivalent number of words. The standard test items and reconstituted items were also matched for linguistic complexity in terms of word frequency and syntactic complexity. For example:
Standard test item: "I consider myself a very socially bold, outgoing person."
Reconstituted test item: "Books and magazines can be found in a library."
Original test items were administered to 60 undergraduates in standard, pencil-and-paper format and

computer format. In the computer format each standard test item was followed by a linguistically matched, emotionally neutral reconstituted item and all response latencies were recorded. The data were analyzed with response latency being the dependent variable and item clarity and endorsement probability as independent variables. It was found that item clarity and endorsement probability did not relate to response latency on standard test items, neutral questions matched to standard test items, or adjusted test items. This finding was attributed to the likelihood that subject responses to items reflect a binary decision-making process which requires relatively simple and consistent responses.

16 Personality Factor test item response latencies adjusted by subtracting latencies of linguistically neutral items were also not affected by item clarity or probability of endorsement.

Consistent with the finding of both Van Merrienboer et al. (1989) and Rattan (1992), it is suggested that the amount of time required for the successful completion of a task depends, in part, on the task's psychological complexity and specific nature. It is also suggested, consistent with Sternberg (1989), that the amount of time required for various tasks does not

operate as a consistent function (of intelligence), but rather as a function of the interaction between the task and the individual's psychological and intellectual make up.

ACKNOWLEDGMENTS

I would like to begin by thanking the members of my doctoral committee at Indiana State University who made this dissertation possible through their leadership, expertise, guidance and support. I sincerely thank Dr. P.G. Aaron for acting as chairperson for my doctoral committee. I also wish to thank Dr. Richard Antes, Dr. David Andrews, Dr. Liam Grimley, and Dr. Walter Sullins for their editorial assistance.

In addition to my committee members, several other people have provided instrumental assistance in completion of this work. I thank Dr. Kenneth Walker, whose support during the proposal writing of this study was appreciated. I also extend thanks to Dr. William Littlejohn, who generously volunteered Blumberg Center computers to collect data for this study. Finally, I thank Dr. Joel Morse, internship director at Crestwood Children's Center, for his interest in this research and my professional development.

I would also like to express my appreciation for several friends and family members who have supported me in this endeavor. I thank my mother and father, Nell and Myron Kinney, who have always supported my academic

endeavors and have been full of empathy about graduate school. I thank my parents-in-law, Roger and Helen Hoelz, for providing me with warmth and encouragement through the years. I thank my two sons, Mitchell and Alexander, who have been a continual source of joy. Finally, I thank my wife, Jill Kinney, who throughout this project has had the rare ability to give perspective without minimizing the enormity of this task. Her insight and sense of humor have made graduate school, and the rest of my life, infinitely more enjoyable.

TABLE OF CONTENTS

LIST OF TABLES

Table

Chapter 1

INTRODUCTION AND LITERATURE REVIEW

This study examines the extent to which characteristics such as semantic clarity of test items in a personality assessment instrument and people's general tendency to endorse these test items affect the time it takes to respond to these test items. The growing body of research regarding computerized administration of personality tests influenced the selection of this problem.

Traditionally, responses to items in personality test inventories have been evaluated in terms of forced choice answers such as "yes"/"no," "true"/"false," or "agree"/"disagree." Introduction of computer administered testing has resulted in yet another potential dependent variable -- response latency. Computer administered tests make it possible to assess response latency more precisely than it is possible with paper-and-pencil tests. One of the intriguing questions that is often raised is whether differences in response

latency provide a measure of yet another personality factor that is not tapped by paper-and-pencil tests. In order to answer this question confidently, it becomes necessary to eliminate or control the contribution of extraneous factors such as length of test items, their linguistic complexity and the reading speed of the test-taker to response latency. For example, if a client takes significantly more time than the group mean to respond to the statement "I abhor sex-related entertainment which is exploitative in nature," should the long response latency be attributed to the unusual words such as "abhor" and "exploitative," or the unusual length of the sentence, or to the client's emotional sensitivity to sex? If the client's response latency is smaller to a neutral sentence that contains the same number of words and words of the same degree of familiarity, but makes no reference to sex, then an emotional reaction to "sex" could be considered. Consequently, elimination of confounding factors such as length and linguistic complexity of test items and the reader's processing speed, all of which are likely to contribute to differences in response latency, may be helpful in improving the validity of the interpretation of computer-administered personality tests. The present study examines the degree of contribution of certain

selected item characteristics and reader characteristics to response latency.

During the past decade, controversy over factors that could possibly contribute to response latencies has increased. In this dissertation, research that has examined the relationship between item and test-taker characteristics and response latencies is reviewed. As Park, Hong, Lee, Cho, and Im (1991), point out, reading speed, verbal ability and motor speed of the individual being tested influence the amount of time required to read, comprehend and respond to questions. Item length also affects the amount of time required to read (Park et al., 1991). Item clarity and the likelihood of endorsing a test item have additionally been suggested as factors that may interact with response latency (Park et al., 1991; Fekken & Holden, 1994). Literature reviewed in subsequent sections suggests that until further understanding of the factors that affect response latency is achieved, response latency cannot be used as an additional criterion in interpreting personality test scores.

Several methods have been proposed to control factors such as response time and item length. In the following section a method proposed by Holden, Fekken and Cotton (1991) will be described along with comments

by other researchers who have criticized this method. Then, a procedure that will introduce means of controlling factors associated with response latency will be introduced in a hypothetical form. The relevant literature addressing the development of reaction time and response latency research will also be reviewed in this section.

Historical Studies of Reaction and Response Time

Contemporary studies of response time were preceded by studies of reaction time, such as those conducted by Sir Francis Galton (1883) and James McKeen Cattell (1890). Their early research was directed toward establishing reaction time as a correlate of intelligence. While reaction time is now considered the time taken by an individual to react to a stimulus, Galton and Cattell interpreted reaction time as the measurement of neural transmission speed (Cattell, 1895). Breitwieser (1911), a researcher and student of Cattell's, attempted to distinguish between shorter and longer reaction times by comparing these times with the placement of a subject's attention. Cattell's and Breitwieser's early studies of reaction time were considered unsuccessful by many experts as they failed

to establish any connection between reaction time and psychological differences (Jensen & Monro, 1979).

Since 1950, reaction time has been studied in connection with information processing theory. Research by Hick (1952) and Hyman (1953) suggests the reaction time to a stimulus is related to the amount of information represented by the stimuli. Their studies consisted of presenting "bits" or "strings" of information to subjects. Hyman (1953) and Hick (1952) observed longer reaction times when subjects presented with longer strings of information. Additionally, Hick (1952) postulated that as speed of reaction increased, the accuracy of recall suffered.

Personality Testing and Response Latencies

While responses to personality inventories have been traditionally evaluated in terms of "yes" or "no" answers, a body of research seeks to use response latency also to identify response patterns (Rogers, 1974; Payne, 1974; Popham & Holden, 1990; Fekken & Holden, 1992; Holden, Kroner, Fekken, & Popham, 1992; Tryon & Mulloy, 1993; Fekken & Holden, 1994). It has been suggested that response latency patterns may also be used to detect invalid responses (Holden, Kroner, Fekken, & Popham, 1992). Researchers have concluded

that if response latency data are to be used to detect response patterns, then the latencies must represent only the amount of time required to mentally select or reject a personality item (Park, Hong, Lee, Cho, & Im, 1991; Holden, Kroner, Fekken, & Popham, 1992; Fekken & Holden, 1994). Thus, test item characteristics and reader characteristics should not be included in the time spent on the actual mental selection or rejection of items even though they may contribute to response latency. An example of a test item characteristic is item length. A longer item requires a longer response time because it takes longer to read. An example of an individual characteristic not related to personality is individual reading speed. A slow reader has a longer response time. In both examples these response latencies are longer because of the additional time the individual takes to read and comprehend the test item. For measures of response latencies to be valid representations of the endorsement/nonendorsement decision process involved in responding to personality test items, it is necessary that they are free of the item and individual factors affecting them. Researchers have shown that the following factors influence response latencies: the length of test items, the complexity of test items, how fast an individual reads and the

individual's overall verbal ability (Holden, Fekken, & Cotton, 1991; Park, Hong, Lee, Cho, & Im, 1991). All of these factors represent sources of error of response latencies because they are all unrelated to the individuals' decision to select or reject a personality test item. The present study is concerned with the effect of two additional factors that can affect response latency of personality test items: 1) item clarity and 2) probability of item selection.

Personality Testing and the Sixteen Personality Factor Questionnaire

Personality is commonly assessed by psychologists using questionnaires such as the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1967), the California Personality Inventory (CPI; Gough, 1987), the Eysenck Personality Inventory (EPI; Eysenck & Eysenck, 1968), the Personality Research Form-Form E (PRF; Jackson, 1984), the NEO Personality Inventory-Revised (NEO PI-R; Costa & McCrae, 1992a), the Myers-Briggs Type Indicator (MBTI; Myers & McCaulley, 1985) and the Sixteen Personality Factor Questionnaire (16 Personality Factor test; Cattell, Eber, & Tatsuoka, 1970). These questionnaires are composed of 'items'

which the individual selects or rejects as characteristic of himself or herself.

According to the Ninth Mental Measurements Yearbook, the 16 Personality Factor test is "most valuable as a personality measure in settings such as personnel selection, guidance counseling, or personality research, where assessment of 'normal range' personality traits is important" (p. 1392 in Mitchell, 1985). The 16 Personality Factor test represents a self-report measure designed to assess factors identified by Raymond Cattell as basic to all personalities. Originally published by Cattell in 1949, the test was revised in 1994 to create the 16 Personality Factor Questionnaire (5th Edition) (Cattell et al., 1994). The latest edition of the 16 Personality Factor test includes several changes. Previous versions of the 16 Personality Factor test included four different forms. Poor reliability and validity of two of the forms led to criticisms from Zuckerman and others (in Mitchell, 1985). The fifth edition of the 16 Personality Factor test uses one form and features enhanced reliability and validity data (Conn & Rieke, 1984). Internal consistency coefficient alphas for the fifth edition's 16 primary factor scales yielded

weighted averages ranging from .66 to .86 with a median of .75.

The 16 Personality Factor test includes 16 primary and five global factors. The test includes the following primary factors: 1) warmth, 2) reasoning, 3) emotional stability, 4) dominance, 5) liveliness, 6) rule-consciousness, 7) social boldness, 8) sensitivity, 9) vigilance, 10) abstractedness, 11) privateness, 12) apprehension, 13) openness to change, 14) self-reliance, 15) perfectionism and 16) tension. The five global factors include the following: 1) extroversion, 2) anxiety, 3) tough-mindedness, 4) independence and 5) self-control. Construct validity of the 16 Personality Factor test was demonstrated by relationships resulting from correlational, regression and principal component analyses of 16 Personality Factor test scores and scores of the CPI, the PRF, the NEO PI-R and the MBTI. Test-retest reliability of the 16 Personality test was established through correlational analyses. The test-retest reliability of the global scale for independence is .84 (N=204) with retesting at two weeks and .81 (N=159) with retesting at two months.

Background of Computerized Personality Testing

Response patterns on personality items have been examined by researchers for years (Rogers, 1974; Payne, 1974; Popham & Holden, 1990; Fekken & Holden, 1992; Holden, Kroner, Fekken, & Popham, 1992; Tryon & Mulloy, 1993; Fekken & Holden, 1994). Until 1989, it was hypothesized that computer administration of personality tests would reduce the number of socially desirable responses produced by test-takers because computer administered testing offers a certain degree of anonymity and is perceived by test-takers as being less judgmental (cited by Lautenschlager & Flaherty, 1990). However, research has failed to support this hypothesis. In fact, some research studies support the opposite conclusion. For instance, Davis and Cowles (1989) reported higher levels of socially desirable responses to personality items by subjects who were administered the Minnesota Multiphasic Personality Inventory (MMPI) through the computer. Lautenschlager and Flaherty (1990) point out that during computer administration of the MMPI, only one question may be seen at a time, whereas the traditional paper-and-pencil version allows respondents to access their previous answers. Furthermore, they suggest that participants register a high number of socially desirable responses for two

reasons. First, they wish to maintain a consistent pattern of responses. Second, because the computer-administered tests allow changing only one previous answer, the amount of control participants have over their answers is limited.

Researchers have found computer-administered versions of the MMPI to yield results that are highly correlated with the results of paper-and-pencil tests. However, higher standard scores (greater pathology) are also reported for the paper-and-pencil formats (Watson, Thomas & Anderson, 1992). Paper-and-pencil format yields higher scaled scores on the MMPI than scores obtained by computer administrations. In a meta-analysis of nine studies comparing traditional and computer-based administration, Watson, Thomas and Anderson (1992) reported profiles of each of the 13 original MMPI scale scores of subjects administered computer versions to be lower in elevation than the 13 scaled scores of paper-and-pencil protocols from the same subjects.

Factors Affecting Response Latencies

Administration of personality tests by computerized format allows for the collection of individual response latencies to test items. Popham

and Holden (1990) measured response latencies using a computer administered version of the MMPI and categorized the latencies in two ways. First, they examined response latencies according to each of the MMPI clinical scales. Second, they examined response latencies according to whether items were endorsed or rejected. They reported that individuals with higher scores (greater pathology) responded more quickly to items they endorsed on these scales than on items they did not endorse. Also, individuals with higher scores took longer to reject items from traditional MMPI clinical scales. Higher scoring subjects took less time to endorse items relevant to higher scores on content-based scales for the MMPI and took more time to reject items relevant to lower MMPI clinical scale scores.

Personality and Response Time

Popham and Holden's (1990) results support the concept that item response times are a product of an individual's personality as measured by personality test scale scores. If individual response latencies are related to individual self-concept, as suggested by Popham and Holden (1990) and others (Fekken & Holden, 1992 and Tryon & Mulloy, 1993), it is possible that

response latency data may be used to confirm or disconfirm the validity of personality test results.

Many personality tests such as the MMPI, the California Personality Inventory and the 16 Personality Factor test feature validity scales which are designed to help detect faking. Popham and Holden (1990) suggest that response latency measures may help them detect separate valid from invalid responses. Detecting invalid responses may help confirm or disconfirm the validity of self-report data provided by a personality test, thus enhancing the validity of the test.

Intelligence and Response Time

According to MacLennon, Jackson, and Bellantino (1988) a negative relationship exists between response latency to intelligence test questions and intelligence test scores. The rate of errors an individual commits decreases the more time the individual takes to respond.

Lally and Nettlebeck (1977) studied the performance of 48 subjects with IQ scores ranging from 57 to 138 on line-judgment tasks. Results revealed that accurately scoring retarded subjects had slower rates and longer response times than non-retarded subjects.

A second experiment (Lally & Nettlebeck, 1977) reported that the slower response times could not be attributed to gross movement difficulties.

Kranzler (1994) has investigated differences in response times on basic cognitive tasks using reading disabled and non-reading disabled children as subjects. Kranzler reported that evidence of a "speed-accuracy trade-off," or the effect of maximizing speed at the expense of accuracy, was negligible. In contrast to Lally and Nettlebeck's (1977) research regarding intelligence and response time, Kranzler reported that reading disabled children with higher IQ scores tended to take longer in solving cognitive tasks than did children with lower IQs.

Similar results were observed by Hoosain (1980) when native Chinese-speakers who spoke fluent English were asked to judge words based on a word's positive or negative connotations, subjects with lower intelligence test scores responded more quickly than subjects who scored higher on intelligence tests.

The apparent discrepancies between the studies of Van Merrienboer et al. (1989), Rattan (1992), Hoosain (1980), Lally & Nettlebeck (1977) and Kranzler (1994), may be explained by understanding the relationship between the task and the individual. The amount of time

required for the successful completion of a task may depend primarily on two factors: (a) task characteristics such as a task's complexity and specificity and (b) individual characteristics, such as those represented by global intelligence scores. Therefore, as suggested by Sternberg (1989), the amount of time required for various tasks may not operate as a consistent function of global intelligence, but rather as a function of the interaction between the task characteristics and the individual.

In a study of Matched Familiar Figures, Van Merrienboer, Jelsma, Timmermans and Sikken (1989), reported that the longer the response times of individuals, the lower are their error rates.

Research conducted on reaction time to personality test items has primarily focused on speed rates (Hunt, Lunneborg, & Lewis, 1975; Sternberg, 1977; Carver, 1990; Whitney, Kellas, & Ferraro, 1990). Rattan (1992) found that response time accounted for between 7.9% and 41% of the variability seen in intelligence test scores, with higher the intelligence test scores correlated with shorter the response times. Jensen (1980) arrived at a similar conclusion.

Other Factors that Influence Response Latencies

If response latencies are to be used in clinical interpretations, further exploration of the cognitive processes associated with response latencies is necessary. Several item and individual characteristics unrelated to personality can affect the amount of time an individual requires to respond to personality items. For instance, the reading speed and the verbal ability of the individual being tested affect response latencies (Park, Hong, Lee, Cho, & Im, 1991). The motor speed of individuals also affects response latencies (Park et al., 1991) and is reflected by the amount of time they take to motorically record their responses, either on paper or by computer keyboard. In addition to individual characteristics such as reading speed, verbal ability, and motoric speed, characteristics of test items also contribute to response latencies. Item length affects response latencies (Park et al., 1991) because a longer amount of time is required to read a longer sentence than a shorter one. Item length, along with the individual's reading speed, dictates how fast an individual will read the item. Item difficulty also affects response latencies. For instance, a longer amount of time is required to comprehend a difficult item than an item that is easily understood. The nature

of response (i.e., pressing a button, marking a square with an "X", etc.) also affects response latencies; a longer amount of time is required to respond by writing an answer than by pushing a button. The level of difficulty required for responding, along with the individual's motor speed, dictate how fast an individual will be able to physically respond to an item. Arriving at a valid assessment of the time involved in reaching a decision requires controlling for the effects of all of these factors as well as individual characteristics.

Controlling Response Latencies

Holden, Fekken, and Cotton (1991) have proposed a method of latency "preparation" that is designed to limit the effect of test-taker and the test item characteristics on response latencies. They claim that their method controls for the effects of item and individual characteristics on response latencies to personality test items. Their method of treating response latencies is called, "dual standardization," as it 'standardizes' latencies in two ways. As a preliminary step, in order to control for the effects of extremely low or high latency values, the extreme values are adjusted to fit within a range between 0.5 seconds to 40 seconds. Holden et al. (1991) report this

adjustment limits the effects extremely high and low response latency values have on the remaining steps of their latency preparation treatment [it may be noted that during their experiment on latencies and MMPI constructs (1991), fewer than 1% of response latencies required this initial adjustment]. Next, to control for the effects of individual characteristics such as reading speed, verbal ability, and motor speed, the response latencies are standardized for that person and used for adjusting his or her response latencies by their average latency value. For example, if an individual's average response latency were 1.74 seconds on all items, 1.74 seconds would be deducted from each single item latency, leading to positive and negative single item latencies for the individual. The result is a mean latency of zero. The presumption is that this procedure will lead to response latencies that are not affected by the individual's reading speed, verbal ability, or motor speed. Then, to control for the effects of item characteristics such as item length, item difficulty, or the level of difficulty required for responding, the response latencies are standardized item-by-item. In this procedure, each latency specific to an item is reduced by the average latency of the subject and this method is applied to all items in the

test. The assumption here is that this procedure will lead to response latencies that are not affected by item properties such as item length, item difficulty or the level of difficulty required for responding. This dual standardization method, with one adjustment for individual factors and another adjustment for item factors, has been used by several researchers who have examined response latencies associated with personality tests (Park et al., 1991; Holden, Kroner, Fekken, & Popham, 1992; Fekken & Holden, 1994).

However, not all response latency variation due to individual or item characteristics may be controlled for by applying the Holden et al. (1991) dual standardization procedure. Park et al. (1991) suggest the method can fail because it is too simplistic. Their research suggests that the effects of individual characteristics such as reading, verbal ability, motor speed, item characteristics and item endorsement probability affect latency data even after the adjustment of latencies using the dual-standardization procedures. Response latency research by Park et al. (1991) using the MMPI supports the idea that even after applying a dual standardization procedure, a number of item and individual characteristics continue to confound latencies.

Another way of controlling for these factors is to match each test item with a neutral sentence which is similar to all characteristics to the test item with the exception of emotional content. Item matching serves as a more precise way of dealing with the error factors associated with latency measurement than the Holden et al. (1991) method. It is based on matching for word frequency (how often words appear in print), a factor that influences individual reading speed and reading comprehension (Aaron & Joshi, 1992) and linguistic complexity (syntactic structure). The item matching method begins by creating sentences that are identical to the items in the personality test in terms of the number and frequency of the words used. To qualify as matched items, the sentences must first meet four criteria. First, the new sentence should contain the same number of words as the original test item. This limits the effect item length has on response latencies. Second, the matched items are constructed of words that have frequencies similar to the words in the test items. Matching is determined according to a frequency list such as the one by Carroll, Davies, & Richman (1971). This step ensures that the frequencies of original and matched items are similar. Third, the new sentence is constructed in such a way that its syntactic nature is

similar to the test sentence. If the test sentence is in the active voice, the new sentence is also in active voice; if the test sentence is in a passive negative interrogative form, the new sentence is also in the same syntactic form. Finally, matched items should be psychologically neutral, consisting of simple and unequivocal true-false statements (e.g., "A fish can swim"). This step ensures that matched items require a minimum amount of thinking to answer. Each linguistically neutral matched item is administered immediately after the personality test item to which it corresponds. Both latencies are recorded and the response latency on the linguistically neutral item is subtracted from the response latency on the personality item. The result is a latency free of effects due to factors such as sentence length and word familiarity.

The psycholinguistically-matched item adjustment method would be superior to the Holden et al. (1991) dual standardization procedure for several reasons. When the Holden et al. (1991) dual standardization procedure was proposed, the authors believed that responding to personality test items took place as a sequential process, as described by Holden's 1992 model. The sequential process model includes: (a) reading the item, (b) comprehending the meaning of the item, (c)

making the decision to select or reject the item and (d) making a physical response to the item. One problem with the model is that it fails to consider that interactions between item and individual characteristics may occur. As suggested by Sternberg (1989), the amount of time required for various tasks may not operate consistently as a function of global intelligence but rather as a function of the interaction between the specific task and individual. If an item is both long and complex, these two item characteristics will interact with individual reading speed and verbal ability thus lengthening the response latency. Also, the Holden et al. (1991) method of latency standardization fails to make adjustment for test fatigue. Because the Holden et al. (1991) standardization technique adjusts every recorded latency by the average latency, no attempt is made to differentiate between item responses recorded during periods of higher or lower levels of test fatigue. Test fatigue is also not acknowledged during the second step of the Holden et al. (1991) technique. Accounting for the effects that item characteristics, individual characteristics and test fatigue all have on response latencies is best achieved by the linguistically neutral item adjustment method.

Summary and Conclusions

Research by Park and colleagues (1991) suggests that before response latency analysis may be further studied for its contribution to personality test interpretation, the interaction between items and individual variables needs to be further examined.

Between 1991 and 1995 researchers examining response latency data have used latency dual-standardization procedures to control the effects of a number of item characteristics and person characteristics including: reading speed, verbal ability, motor speed, probability of item endorsement, item length and item complexity. All of these factors have been suggested as contributing to variation in response time (Park et al., 1991; Fekken & Holden, 1994).

However, two sources of error in the measurement of response latency not controlled for by using linguistically matched neutral items still remaining are: (a) clarity or vagueness of items and (b) probability that an individual will endorse or reject a particular item. Item clarity and the likelihood of endorsing a test item have been suggested as factors that may interfere with the individual response latency

associated with personality test items (Park et al., 1991; Fekken & Holden, 1994).

Item endorsement probability represents the probability that a particular item will be endorsed by any given respondent. Endorsement probability, also referred to as item "controversiality" (Park et al., 1991), is lower for personality test items rarely endorsed and higher on items frequently endorsed.

It is possible that examination of the interaction between test item clarity and probability of item endorsement, while controlling for errors in measurement, may lead to enhancing the validity of personality tests.

Presentation of the Problem

The present study attempted to evaluate the confounding nature of two variables (viz., item clarity and probability of endorsement) by using neutral, psycholinguistically-matched personality items to control for individual factors and then investigates the effects of item clarity and probability of endorsement on response latencies on personality test items. The personality test items used comprise one of the global factor scales of the 16 Personality Factor Questionnaire (5th Edition, Cattell, Eber, & Tatsuoka, 1970).

Question 1: What effect does the clarity of a personality test item have on individual response latency on 16 Personality Factor test items?

Null Hypothesis 1: Item clarity will not significantly effect individual response latencies on 16 Personality Factor test items.

Question 2: What effect does the probability that an individual will endorse an item have on individual response latencies on 16 Personality Factor test items?

Null Hypothesis 2: Probability of item endorsement will not significantly effect response latencies on 16 Personality Factor test items.

Question 3: What interactive effects do item clarity and probability of endorsement have on individual response latencies on 16 Personality Factor test items?

Null Hypothesis 3: Item clarity and probability of endorsement will not have a significant interactive affect on individual response latencies on 16 Personality Factor test items.

Limitations of the Present Study

The objective of the present study is to examine if significant differences exist between mean response latencies on items from the independence-dependence

global scale of the 16 Personality Factor test due to the following factors: (a) clarity of the item and (b) degree of probability of the item being endorsed. Due to methodological design, the current study has limitations. There is also the possibility that factors unknown at this time might also affect response latencies.

1. The sample is limited to undergraduate students from a medium-sized midwestern university. Because the sample may not be representative of people of all backgrounds, results will generalize appropriately only to populations from similar backgrounds.

2. Only adults aged 17 to 25 served as participants. Consequently, generalization to other age groups may be inappropriate.

3. Only items from the global factor scale for independence of the 16 Personality Factor test were used to reject or support the null hypotheses. Consequently, generalizations to the entire 16 Personality Factor test or items from other tests may be inappropriate.

4. Norm development and other factors that might be addressed regarding computer administration of personality items will not be examined in this study.

Definition of Terms

As the present study is primarily designed to establish differences among response latency due to item clarity and probability of endorsement, the following terminology was used.

1. Item clarity of the 16 Personality Factor test items was defined by ratings by a group of ten graduate students in a school psychology program at a midwestern state university. Items from the global scale for independence of the 16 Personality Factor test were rated along a seven-point rating scale ranging from 1 ("very vague") to 7 ("very clear") in a manner similar to Rogers (1974). Clear was defined by how easily an item is understood. Vague items were defined as those which were rated as having an ambiguous or confusing meaning.

2. Probability of item endorsement was defined as the percentage probability of endorsement as determined according to norms (N=462) provided by the Institute for Personality and Ability Testing, publishers of the 16 Personality Factor Questionnaire. For example, based on the endorsement norms provided, there is a 27% chance that a respondent will endorse test item #139, "I suspect that people who seem friendly to me could be disloyal behind my back."

3. Individual response latencies were defined as the total time of reading, thinking, deciding and responding to personality test items calculated as the difference in time between presentation of item and the time when response was made and recorded by computer.

4. Adjusted response latencies were adjusted by subtracting matched neutral item latencies from corresponding personality test item latencies.

5. Neutral matched question items were identical to their corresponding personality test items in the number and frequency of words used. Neutral matched items were administered immediately following the administration of each corresponding 16 Personality Factor test item.

6. Psycholinguistic matching refers to the method used for matching the words in the matched question to the words in the original test item in terms of word frequency (Davies and Richman, 1971).

7. Neutral items refers to matched question items which are also identical to test items in their syntactical make up.

8. "Emotionally neutral items" refers to neutral items and have little or no emotional content (e.g., Fish can swim).

9. "Personality test items" refers to the 44 items that make up the independence-dependence global scale of the 16 Personality Factor Questionnaire.

Assumptions

The present study makes the following assumptions based on the findings of earlier research studies.

1. It is assumed that item length contributes to the production of response latencies. Item length was controlled by using the same individual's response latencies on linguistically matched items to adjust latencies collected to 16 Personality Factor test items.

2. It is assumed that variability in the reading speed of test respondents is associated with individual response latencies. Variability in the reading speed of test respondents was controlled by using the same individual's response latencies on linguistically matched items to adjust latencies collected for 16 Personality Factor test items.

3. It is assumed that variability in the gross motor ability of individuals is associated with individual response latencies. Variability in participants' gross motor ability was controlled by using the same individual's response latencies on

linguistically matched items to adjust latencies collected for 16 Personality Factor test items.

Chapter 2

METHOD

Participants

The participants for this study were 60 Indiana State University undergraduate students, ranging in age from 17 to 25 years. All of the participants volunteering for the study were compensated with five dollars. Participants were tested individually or in pairs in a quiet room. After being informed regarding informed consent and anonymity, participants were asked to take a computer test. Participants were given standard instructions for the 16 Personality Factor test. Additionally, participants were informed that their test performance would be timed. All general requirements for informed consent were met while this study was conducted.

Instrumentation

The 16 Personality Factor Questionnaire is a personality assessment instrument consisting of 185

items. For the purposes of this study, the 44 items comprising the "global scale for independence" of the 16 Personality Factor Questionnaire along with 44 linguistically neutral items, were used to test the validity of the null hypotheses.

Clarity Ratings

To determine clarity of the 44 test items used, a group of ten graduate students in the school psychology program at a midwestern state university were asked to rate these items according to their clarity on a 7-point scale ranging from 1 ("very vague") to 7 ("very clear") in a manner similar to Rogers (1974). Subsequently the 44 test items used were assigned a value for clarity based on mean ratings for each question.

Matched Items

To control for item characteristics such as length and respondent characteristics such as reading speed, verbal ability, and motor speed the 44 test items used were linguistically neutral with a neutral, matched equivalent question counterpart, identical in the sentence length and word frequency (Carroll, Davies, & Richman, 1971). Neutral matched questions were also

created in such a way they required as little thought as possible for responding (e.g., A fish can swim).

Adjusted Latencies

Adjusted latencies were determined by subtracting the response latencies on linguistically neutral items from response latencies for their corresponding personality test items.

Probability of Endorsement Statistics

Probability for item endorsement is not listed in the standard document of the 16 Personality Factor Questionnaire. However, these data were obtained from the Institute for Personality and Ability Testing, publishers of the 16 Personality Factor test by making a written request.

2.3 Procedure and measures

The computer test consisted of the 44 items from the 16 Personality Factor global scale for independence. The 44 matched neutral items were also presented with the aid of the computer. All of the items were presented using an Apple Macintosh microcomputer special-purpose program created with HyperCard and its associated script language HyperTalk (Apple Computer

Inc., 1987, 1988, 1993). The automated format for the test and neutral items allowed three choices for participant responding: endorsement, non-endorsement, and "do not know." Participants were given a series of 20 sample questions that helped them to become acquainted with the format of the computer program. After standard instructions for the 16 Personality Factor test were followed, the computer administered test began.

The computer program administered all 44 personality test items and the 44 neutral items. The test items and neutral items were presented in an alternating fashion. The computer program recorded the time elapsed between the initial presentation of items on the computer screen and the key stroke response was made by the participant. Each response latency was recorded to the nearest one-hundredth second. The available responses offered by the program were: (a) true; (b) cannot say; and (c) false. The format of the personality test items allows for responses other than true or false on some questions. Subjects had the option to press key "a" for the first response choice, key "b" for cannot say, and key "c" for the third response choice. The computer program recorded the type

of response along with the response latencies for all questions administered.

Finally, to compare the difference between computer presentation and conventional presentation, participants were asked to take the standard paper-and-pencil version of the 16 Personality Factor test, administered using standard instructions for the test.

Design

The analysis of data used a stepwise multiple regression design, where the two independent variables of item clarity and probability of endorsement are used as independent variables to test the null hypotheses. The dependent variable was individual response latencies on personality test items, linguistically neutral items, and adjusted latencies.

Analysis

The three research questions and their corresponding null hypotheses were examined statistically using multiple regression of response latency on the three factors of item variability. The analysis points out the degree of relationship that item clarity (X_1), probability of endorsement (X_2), and the

combined effects of item clarity and probability of item endorsement have to response latency to standard items (Y), response latency to linguistically neutral items (Y_2), and item latency that is adjusted for linguistic variables (Y_3). The data for this study were analyzed using the SPSS statistical program, version 6.1, designed for the Apple Macintosh microcomputer. The analyses designs are shown in Table 1.

The relationship between paper-and-pencil and computer administered 16 Personality Factor test questions was also evaluated by computing the Pearson correlation coefficient. The significance level used for statistical tests was .05.

Table 1.
Design for the Analysis of Data.

Independent variables	Dependent variables
1. Item clarity	1. Response latency to standard items
2. Probability of item endorsement	2. Response latency to matched neutral items
3. Item clarity and probability of endorsement	3. Response latency to linguistically adjusted items

Chapter 3

Results

As the purpose of the hypotheses was to identify the effects of item clarity and item endorsement probability on personality test item response latencies, linguistically neutral item latencies, and adjusted personality test response latencies, a series of stepwise multiple regression analyses were performed. Within these analyses, each of the latency factors (real items; linguistically neutral items; and adjusted items) were individually regressed on item clarity, probability of endorsement and item clarity along with the probability of endorsement in the following steps: step 1: item clarity; step 2: probability of endorsement; step 3: item clarity and the probability of item endorsement. As this experiment is concerned with the response latencies to each of the three types of questions, data were obtained by averaging the latencies of all participants to each question. After the average response latency for each of the 44 items was

determined, means, variances and ranges were tabulated for each independent and dependent variable (see Table 2).

Latencies for Standard Personality Test Items

Hypotheses 1, 2 and 3 address the extent to which item clarity and probability of endorsement affect response latencies. In examining the effect of these independent variables on response latencies to personality test items, no significant findings were observed. Item clarity ($F,(1,42)=0.27$, $p>.05$) and endorsement probability ($F,(2,41)=0.73$, $p>.05$) were not found to significantly predict response latencies to personality test items in their standard form.

Table 2.

Response Latencies and Descriptive Statistics for the Three Types of Dependent Variables (N=44).

	means	variance	range
Y_1 16PF Test items	5.64	2.67	3.46 - 9.80
Y_2 Matched items	6.51	2.68	3.79 - 11.93
Y_3 Adjusted items (Y_1 - Y_2)	-0.87	5.13	-6.19 - 4.53

Note. Values represent time in seconds.

As seen in Table 2, the observed variance of personality test item response latencies was 2.67 seconds and the observed variance in neutral matched items was 2.68 seconds. However, the observed variance in adjusted items was 5.13 seconds.

Latencies for Linguistically Neutral Items

In examining the effect of item clarity and probability of endorsement on response latencies to linguistically neutral items, no significant findings were observed. In a stepwise multiple regression analysis using response latencies on linguistically neutral items as the dependent variable, item clarity ($F,(1,42)=3.86$, $p>.05$) and endorsement probability ($F,(2,41)=4.33$, $p>.05$) were not found to significantly predict response latencies on neutral matched items.

Combined Response Latencies

In examining the effect of item clarity and probability of endorsement on adjusted personality test item response latencies, no significant findings were observed. A stepwise multiple regression analysis using adjusted personality test item response latencies showed that none of the independent variables were statistically significant as predictors of adjusted

response latencies. Item clarity ($F,(1,42)=3.22$, $p>.05$) and item endorsement probability ($F,(2,41)=1.80$, $p>.05$) were not found to significantly predict response latencies to adjusted response latencies.

Validity of Computer Test Responses

Of the 60 participants in this study, 51 completed both the paper-and-pencil version of the 16 Personality Factor test and the computer test. A correlation coefficient between the results of the two administrations was computed. The two-tailed test of probability correlation coefficient between global factor scale for independence on the two types of question item administrations was 0.826, $p<.001$ (N=51). This correlation is comparable to those typically cited in other studies that examined the concordance of responses between paper-and-pencil and computer administered personality tests (Lushene & O'Neil, 1974; Reardon & Loughead 1988; and Sanitioso & Reynolds, 1992).

Correlations Between Variables

Other Pearson product-moment correlational analyses suggest similarities between personality test items and linguistically neutral items. Participant

response latencies on linguistically neutral items were similar in length to response latencies on personality test items. The correlation observed between response latencies on personality test items and those on linguistically neutral items is $r = .786$, $p<.05$ (N=44). Additionally, as seen in Table 2, the variance between the two question item types was nearly identical, indicating little dispersion among the participant's latencies on the two question types. The linguistically neutral question items appear to have accomplished the goal of bearing strong linguistic resemblance to the personality test items (see appendices 6.3 and 6.4).

To examine similarities between independent variables, such as item clarity and probability of endorsement, additional correlational analyses were performed. Item clarity ratings and probability of endorsement figures were not significantly correlated, $r= .074$, $p>.05$ (NS) (N=44).

Chapter 4

DISCUSSION

The primary goal of this study was to examine the effect item clarity and item endorsement probability have upon response latencies on personality test items, linguistically neutral items, and latencies adjusted for reading speed and vocabulary. The present study evolved from recent research aimed at understanding variation in response latency with a goal of eventually utilizing latencies in validating personality tests and aiding in test interpretation (Park et al., 1991; Holden, Kroner, Fekken & Popham, 1992).

In this study endorsement probability and item clarity were not established as predictors of response latencies on standardized personality test items, response latencies on linguistically neutral items or adjusted response latencies. A high correlation between paper-and-pencil test responses and computer-administered test responses was obtained.

Main Results

Confirmed Hypotheses

Two significant findings emerged from the analysis of variables. Among these was that for all three types of latencies used as dependent variables, item clarity and endorsement probability had no significant effect. This suggests that item clarity and endorsement probability do not play an important role in consistently determining response latencies on linguistically adjusted items.

Additionally, the prediction that responses to computer administered personality questions would be strongly correlated with responses from the paper-and-pencil test was confirmed. Responses to the paper-and-pencil version of the 16 Personality Factor test revealed concordance with responses to those questions used in conjunction with the computer-administered 16 Personality Factor test items with $r = 0.826$, $p<.001$ (N=51).

Rejected Hypotheses

By contrast to the significant findings described above, several predicted effects did not occur. It was hypothesized that item clarity and item endorsement probability would affect response latencies on

personality test items. Of the item factors representing independent variables of the original test in this study, neither had significant predictive effect on response latencies to the personality test items.

It was hypothesized that item clarity and item endorsement probability would affect response latencies on linguistically neutral items. Item clarity and endorsement probability had no significant predictive influence on response latencies on linguistically neutral items.

It was hypothesized that item clarity and item endorsement probability would affect latencies adjusted for reading speed and vocabulary. Item endorsement probability and item clarity were not significant as predictors of latencies adjusted for reading speed and vocabulary.

Methodological Issues

As previously mentioned, consistent with findings of Lushene and O'Neil, (1974), Reardon and Loughead (1988), and Sanitioso and Reynolds, (1992), this study suggests there is no significant variation in the type of responses produced for the paper-and-pencil test and the computer presented personality test.

Use of linguistically neutral items in examining response latencies on personality test items represents a procedure not previously used in response latency research. The present study fails to support endorsement probability as a significant predictor of response latencies on personality test items This is contrary to the previous findings of Park et al. (1991). However, since Park et al. (1991) employed the Holden method of dual standardization procedure to reduce item variables in observed latencies (Holden, at al., 1991), a direct comparison may be overly presumptive. Similarly, comparisons of the present study to others which have employed Holden's (1991) method of latency preparation (Park et al., 1991; Holden, Kroner, Fekken, & Popham, 1992; and Fekken & Holden, 1994) may be inappropriate.

The employment of linguistically neutral items demonstrated apparent effectiveness in controlling for the individual and item variables involved in response latency production, as demonstrated by the variance and correlation between test item and linguistically neutral item latencies.

The average of response latencies on linguistically neutral items was greater than latencies on personality test items from the original test. This

finding was not expected. It was expected that the linguistically neutral items, designed to be easily answered, would require less time to answer, although the opposite was found. Linguistically neutral items typically required more response time than their personality test item counterparts from the original test. This result may stem from the nature of the two question types used in this study. The initial premise for using linguistically neutral items was to employ questions that could be processed easily as a method of controlling for individual differences such as reading speed and vocabulary. Since the latencies on the linguistically neutral items averaged longer than the latencies on personality test items from the original test, it can be assumed that these latencies reflect more than merely the time required for participants to read, comprehend, decide, and physically generate a response to an item. It is possible that the personal nature of the question dramatically affects the average length and overall consistency of response latencies across different items. In other words, items that require the respondent to engage in introspective thinking are inherently more inconsistent than items free of personal interference.

The findings of the present study have several implications for researchers attempting to understand cognition and what is represented by the time it takes to respond to different types of questions. Sternberg (1989) suggests that the amount of time required for various tasks may not operate as a function of intelligence consistently, but rather as a function of the interaction between the task characteristics and individual characteristics. In this study, the specific task is responding to questions, representing either a personality test item or a modified item. According to Sternberg's (1989) postulate, if a series of consistent tasks is presented to an individual, the interaction between tasks and respondent characteristics should be consistent and the response time would be predictable. In the present study, when the demands of the task remained consistent for all participants, the predictive effects of item clarity and endorsement probability were determined as nonsignificant.

Response latencies on personally neutral, linguistically modified items appear to require more processing time than when the question forces the respondent to answer from a personal frame of reference. Shorter latencies on personality test items suggest more rapid retrieval of personal information and slower

retrieval of general information. It may be that response latencies to the different types of questions varied in length as a result of the varying familiarity participants have with the type of information required for a response. To better understand an individual's response latencies to personality questions, it would seem important to also examine an individual's familiarity with self.

The findings of this study have several practical methodological implications. One of these is that item characteristics studied appear insufficient in predicting response latencies to personality test items. Thus, use of an additional variable, perhaps one related to participant's scores on specific personality traits, seems necessary. A participant-based independent variable may provide information concerning whether a relationship exists between response latencies and self-awareness. Another methodological implication of this study is that creating linguistically neutral response latencies that are most effective in appropriately adjusting personality test item response latencies may mean creating linguistically neutral items based on personally-relevant neutral content.

Limitations of the Present Study

This study has some limitations. First, although linguistically neutral items were designed to reduce or eliminate linguistic variability and emotional factors which could lead to longer response times, they tended to generate longer response latencies than personality test items. The presence of longer average response latencies on linguistically neutral items suggests the linguistic characteristics of the items led to longer latencies. Due to the different nature of linguistically neutral items, they may not have been matched to personality items as closely as possible. Linguistically neutral items might have been more effective if their content reflected a personal frame of reference, rather than being neutral. Using neutral items as a means of determining whether significant regressive linearity existed between the adjusted item latencies and the three independent variables employed may not be practical.

Though a rating system similar to Rogers (1974) was employed to generate ratings based on item clarity for personality test items, the distribution of item clarity ratings was narrow and ranged between 3.10 and 6.40 with a variance of .56. Because of this, statistical findings based on item clarity ratings may

have been the least robust of the findings in the present study. Item clarity might have been more statistically significant in predicting response latencies if the rating scale had been based on a wider ranging scale. Another limitation of the clarity ratings was that they were limited to personality test items. Clarity ratings for linguistically neutral items would have allowed a direct comparison of clarity between linguistically neutral items and personality test items in standard form.

Prior to conducting the regression analyses for the present study, all participant responses were averaged for each question item. The averaging of participant's response latencies eliminated or reduced individual differences in response patterns. As suggested by Holden & Fekken (1994), such individual patterns of response play a vital role in the production of response latencies. As no individual-by-individual comparison of response latency patterns was included in the present study, any relationship between individual response sets and response latencies was overlooked.

Conclusions

Historically, research efforts have been directed toward understanding how standardized personality test

response latencies might be employed as an additional measure of test validity (Popham & Holden, 1990; Holden & Hibbs, 1995). One of the methodological problems inherent in previous research has been a failure to take test item features into account as contributors of latency. The notion of controlling for features such as item clarity and endorsement probability which may be used to control for individual differences in response latencies may be unnecessary.

The findings of the present study highlight the possibility of controlling respondent characteristics, such as reading ability, as a means of limiting variability in individual's response latencies on personality test items. The findings also suggest that future studies may benefit by employing a linguistically neutral item adjustment, using emotionally neutral, linguistically controlled items to further verify the legitimacy of using item clarity and probability of item endorsement to control for individual differences in response latencies.

In conclusion, although item clarity and endorsement probability were not established as significant predictors of adjusted 16 Personality Factor test item latencies in this study, they may have significance if latencies are examined in relation to

whether or not an item is endorsed. The findings of the present study suggest the need for additional studies to further explore the effectiveness of using linguistically controlled and emotionally neutral items which incorporate content that is personally relevant yet neutral. By examining the effects of using improved linguistically neutral items, it may be possible to determine to what extent self-knowledge influences response time to personality test items.

REFERENCES

Aaron, P.G., & Joshi, R.M. (1992). Reading problems: Consultation and remediation. New York, NY: The Guilford Press.

Apple Computer, Inc. (1987). Macintosh HyperCard user's guide. Cupertino, CA: Author.

Apple Computer, Inc. (1988). Macintosh HyperCard script language: The HyperTalk language. Reading, MA: Addison-Wesley.

Apple Computer, Inc. (1993). HyperCard reference Manual. Cupertino, CA: Author.

Breitwieser, J.V. (1911). Attention and movement in reaction time. Columbia University contributions to philosophy and psychology, 19, no. 4.

Carver, R.P. (1990). Reading rate: A review of research and theory. New York, NY: Academic Press.

Carroll, J.B., Davies, P., & Richman, B. (1971). The American Heritage word frequency book. Boston: Houghton Mifflin.

Cattell, James McKeen (1895). On reaction-times and the velocity of the nervous impulse. Memoirs of the National Academy of Sciences, 7, 393-415.

Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970). Handbook for the Sixteen Personality Factor Questionnaire, Champaign, IL: IPAT.

Conn, S.R., & Rieke, M.L. (Eds.). (1994). 16PF Fifth Edition Technical Manual. Champaign, IL: IPAT.

Costa, P.T., Jr., & McCrae, R.R. (1992a), NEO Personality Inventory Revised. Odessa, FL: Psychological Assessment Resources.

Davis, C., & Cowles, M. (1989). Automated psychological testing: Methods of administration, need for approval, and measures of anxiety. Educational and Psychological Measurement, 49, 311-320.

Department of Health and Human Services (1983). Department of Health and Human Services administration of grants: Federal regulations Title 45. Washington: U.S. Department of Health and Human Services, Office of the Secretary, Office of Grants and Procurements.

Doll, B. (1986). The ethical computer: The ethics of expert systems. Computer and Technological Applications in School Psychology Newsletter, Winter, 13-16.

Eysenck, H.J. & Eysenck, S.B.G. (1968). Manual for the Eysenck Personality Inventory. San Diego, CA: Educational and Industrial Testing Service.

Fekken, C.G., & Holden, R.R. (1994). The construct validity of differential response latencies in structured personality tests. Canadian Journal of Behavioural Science, 26, 104-120.

Fekken, C.G., & Holden, R.R. (1992). Response latency evidence for viewing personality traits as schema indicators. Journal of Research in Personality, 26, 103-120.

Gough, H.G. (1957). The California Personality Inventory. Palo Alto, CA: Consulting Psychologists Press.

Hathaway, S. & McKinley, J. (1967). The Minnesota Multiphasic Personality Inventory (rev. ed.). New York, NY: The Psychological Corporation.

Hick, W.E. (1952). On the rate of gain of information. Quarterly Journal of Experimental Psychology, 4, 11-26.

Holden, R.R., Fekken, G.C., & Cotton, D.G.H., (1991). Assessing psychopathology using structured test-item response latencies. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3, 111-118.

Holden, R.R., & Hibbs, N. (1995). Incremental validity of response latencies for detecting fakers on a personality test. Journal of Research in Personality, 29, 362-372.

Holden, R.R., Kroner, D.G., Fekken, G.C., & Popham, S.M. (1992). A model of personality test item response dissimulation. Journal of Personality and Social Psychology, 63, 272-279.

Hoosain, R. (1980). Mental speed, response time, and intelligence. Perceptual and Motor Skills, 51, 799-803.

Hunt, E.B., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? Cognitive Psychologist, 7, 194-227.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. Journal of Experimental Psychology, 45, 188-196.

Jackson, D.N. (1984). Personality Research Form, Form E. Port Huron, MI: Sigma Assessment Systems.

Jensen, A.R. (1980). Bias in mental testing. New York, NY: Free Press.

Jensen, A.R., & Monro, E. (1979). Reaction time, movement time, and intelligence. Intelligence, 3, 121-126.

Kranzler, J.H. (1994). Application of the techniques of mental chronometry to the study of learning disabilities. Personality and Individual Differences, 16, 853-859.

Lally, M., & Nettlebeck, T. (1977). Intelligence, reaction time, and inspection time. American Journal of Mental Deficiency, 82, 273-281.

Lautenschlager, G.H., & Flaherty, V.L. (1990). Computer administration of questions: More desirable of more social desirability? Journal of Applied Psychology, 75, 310-314.

Lushene, R.E., & O'Neil, H.F., Jr. (1974). Equivalent validity of a completely computerized MMPI. Journal of Personality Assessment, 38, 353-361.

MacLennon, R.N., Jackson, D.N., & Bellantino, N. (1988). Response latencies and the computerized assessment of intelligence. Personality and Individual Differences, 9, 811-816.

Mitchell, J.V., Jr. (Ed.). (1985). The Ninth Mental Measurements Yearbook. Lincoln, NE: The University of Nebraska Press.

Myers, I.B., & McCaulley, M.G. (1985). Manual: A Guide to the development and use of the Myers-Briggs Type Indicator. Palo Alto, CA: Consulting Psychologists Press.

Park, Y., Hong, C., Lee, C., Cho, T., & Im, N. (1991). Effects of intelligence and item characteristics upon the MMPI response latency (1991). Mental Health

Research Institute publication, Hanyang University, Seoul, Korea.

Payne, F.D. (1974). Relationships between response stability and item endorsement, social desirability, and ambiguity in the MMPI and the CPT. Multivariate Behavioral Research, April, 127-148.

Popham, S.M., & Holden, R.R. (1990). Assessing MMPI constructs through the measurement of response latencies. Journal of Personality Assessment, 54, 469-478.

Rattan, G. (1992). Speed of information processing: A third factor of the WISC-R. Journal of School Psychology, 30, 83-90.

Reardon, R., & Loughead, T. (1988). A comparison of paper-and-pencil and computer versions of the self-directed search. Journal of Counseling and Development, 67, 249-252.

Rogers, T.B. (1974). An analysis of the stages underlying the process of responding to personality test items. Acta Psychologica, 38, 205-213.

Sanitioso, R., & Reynolds, J.H. (1992). Comparability of standard and computerized administration of two personality questionnaires. Personality and Individual Differences, 13, 899-907.

Sternberg, R.J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillside, NJ: Erlbaum.

Sternberg, R.J. (1989). Inside intelligence. American Scientist, 74, 137-143.

Tellegen, A. (1988). The analysis of consistency in personality assessment. Journal of Personality, 56, 621-661.

Tyron, W.W., & Mulloy, J.M. (1993). Further validation of computer-assessed response time to emotionally evocative stimuli. Journal of Personality Assessment, 61, 231-236.

Van Merrienboer, J.G., Jelsma, O., Timmermans, J., & Sikken, J. (1989). Computer versus experimenter controlled administration of the matching familiar figures test: Mean test scores and reliabilities. Educational and Psychological Measurement, 49, 883-891.

Watson, C.G., Thomas, D., & Anderson, P.E. (1992). Do computer-administered Minnesota Multiphasic Personality Inventories underestimate booklet-based scores? Journal of Clinical Psychology, 48, 744-748.

Whitney, P., Kellas, G. & Ferraro F.R. (1990). Reading ability and semantic processing: Isolating speed of access. Journal of Educational Psychology, 82, 479-485.

APPENDIXES

6.1 Appendix A: Participant Consent Form

Statement of Informed Consent

1. I understand that my participation in this study is voluntary and that I may withdraw from the study at any time.

2. I understand that my name will not be used in connection with my participation in this study.

3. I understand that my responses to this study will be recorded and timed.

4. I understand that there will be no "trick questions" used in this study.

5. I am volunteering for this study on my own free will.

6. I understand that I will be compensated $5.00 for my participation in this study.

---

signature

---

print name

10_____

assigned number

Appendix B: Clarity Ratings Instructions

Dear Volunteer,

On the following 10 pages are listed 44 questions from the 16 Personality Factor test (16 Personality Factor Questionnaire) (Cattell, Eber, & Tatsuoka, 1970), Fifth Edition. My study, entitled "Effect of Item Clarity and Probability of Item Endorsement on Response Latencies on Personality Test Items", will utilize these questions to gather evidence on response latency, or the amount of time required to respond to such items. Item clarity of 16 Personality Factor test items will be defined by item ratings of a group of graduate students in the School Psychology program.

Please rate all 44 questions along a seven point scale ranging from 1 ("very vague") to 7 ("very clear"). Clear questions being those more easily understood and vague question those having an ambiguous or confusing meaning.

Please complete the attached questionnaire sheets and return them in the post-paid envelope.

If you have any questions please do not hesitate to contact me. My telephone number is (812) 232-4917.

Thank-you in advance for participating in this part of the study.

Sincerely,
Jim Kinney

Appendix C: 16 Personality Factor Test independence-dependence scale items

Note: items are numbered according to their delivery in the computer program administration.

1. When people do something that bothers me I usually;
a. let it go;
b. ?
c. mention it to them.

3. In joining a new group, I usually seem to fit in right away.
a. true
b. ?
c. false

5. There's usually a big difference between what people say they'll do and what they actually do.
a. true
b. ?
c. false

7. A lot of people will "stab you in the back" in order to get ahead themselves.
a. true
b. ?
c. false

9. I'd prefer to deal with people who are:
a. conventional and polite in what they say;
b. ?
c. direct and speak up about problems they see.

11. I like to think up better ways of doing things rather than to follow well-tried ways.
a. true
b. ?
c. false

13. If I had to cook or build something, I'd follow the directions exactly.
a. true, why take chances
b. ?
c. false, I'd probably try to make it more interesting

16 Personality Factor Test independence-dependence scale items, continued

15. In a situation where I'm in charge, I feel comfortable giving people directions.
a. true
b. ?
c. false

17. People think of me as more:
a. cooperative;
b. ?
c. assertive.

19. I am shy and cautious about making friends with new people.
a. true
b. ?
c. false

21. It's important to pay attention to other people's motives.
a. true
b. ?
c. false

23. People form opinions about me too quickly.
a. hardly ever
b. ?
c. often

25. I don't like people who are "different" or unusual.
a. true, I usually don't
b. ?
c. false, I usually find them interesting.

27. I'm more interested in:
a. seeking personal meaning in life;
b. ?
c. a secure job that pays well.

29. What this world needs is:
a. more steady, solid citizens;
b. ?
c. more reformers with opinions about how to improve the world.

16 Personality Factor Test independence-dependence scale items, continued

31. If people are doing something wrong, I usually tell them what I think.
a. true
b. ?
c. false

33. I tend to get embarrassed if I suddenly become the center of attention in a social group.
a. true
b. ?
c. false

35. Starting conversations with strangers:
a. never gives me any trouble;
b. ?
c. is hard for me.

37. It's wise to be on guard against smooth talkers because they might take advantage of you.
a. true
b. ?
c. false

39. People are lazy on the job if they can get away with it.
a. hardly ever
b. ?
c. often

41. I find people more interesting if their views are different from most people's.
a. true
b. ?
c. false

43. I like people who:
a. are stable and conventional in their interests;
b. ?
c. seriously think through their views about life.

45. Work that is familiar and routine makes me feel;
a. bored and sleepy;
b. ?
c. secure and confident.

16 Personality Factor Test independence-dependence scale items, continued

47. If being polite and pleasant doesn't work, I can be tough and sharp if I need to.
a. true
b. ?
c. false

49. If we were lost in a city and my friends didn't agree with me on the best way to go, I'd:
a. make no fuss and follow them;
b. ?
c. let them know that I thought my way was best.

51. I have always had to fight against being too shy.
a. true
b. ?
c. false

53. When I'm in a group, I usually sit and listen and let others do most of the talking.
a. true
b. ?
c. false

55. If people are frank and open, others try to get the best of them.
a. hardly ever
b. ?
c. often

57. It seems that more than half the people I meet can't really be trusted.
a. true, they can't be trusted.
b. ?
c. false, they can be trusted.

59. I like to think out ways in which our world could be changed to improve it.
a. true
b. ?
c. false

61. In my newspaper, I'd rather read:
a. articles on current social problems.
b. ?
c. all the local news.

16 Personality Factor Test independence-dependence scale items, continued

63. I believe in complaining if I receive bad service or poor food in a restaurant.
a. true
b. ?
c. false

65. When others don't see things my way, I can usually get them to come around.
a. true
b. ?
c. false

67. I consider myself a very socially bold, outgoing person.
a. true
b. ?
c. false

69. I'm usually the one who takes the first step in making new friends.
a. true
b. ?
c. false

71. I suspect that people who seem friendly to me could be disloyal behind my back.
a. hardly ever
b. ?
c. often

73. Many people are too fussy and sensitive and should toughen up for their own good.
a. true
b. ?
c. false

75. More trouble arises from people:
a. questioning and changing methods that are already satisfactory;
b. ?
c. turning down promising, new approaches.

77. When I find I differ with someone on social views, I prefer to:
a. discuss what our differences mean;

16 Personality Factor Test independence-dependence scale items, continued

b. ?
c. discuss something else.

79. I most enjoy a meal if it consists of familiar, everyday foods rather than new, unusual foods.
a. true
b. ?
c. false

81. If I notice that another person's line of reasoning is wrong, I usually:
a. point it out;
b. ?
c. let it pass.

83. I enjoy having some competition in the things I do.
a. true
b. ?
c. false

85. I find it hard to speak in front of a large group.
a. true, I usually find it very hard
b. ?
c. false, it doesn't bother me

87. In social groups I tend to feel shy and unsure of myself.
a. true
b. ?
c. false

Appendix D: Psycholinguistically Neutral Items

Note: items are numbered according to their delivery in the computer program administration.

2. To fix a flat tire you would probably use:
a. a car jack;
b. ?
c. a hammer and nails.

4. On a hot and sunny day, it is usually cooler in the shade.
a. true
b. ?
c. false

6. If there's smoke and flames coming from your neighbor's house, you should call the fire department.
a. true
b. ?
c. false

8. When someone is "as drunk as a skunk" they probably should not try to drive.
a. true
b. ?
c. false

10. If you were shipwrecked you'd rather have:
a. some way to contact help;
b. ?
c. a very expensive work of art.

12. If you wanted to pay someone to fix your car you might talk to an auto-mechanic.
a. true
b. ?
c. false

14. Dropped into the ocean, a rock would float on the surface of the water.
a. true, rocks are heavy
b. ?
c. false, it'd float because rocks are lighter than water

Psycholinguistically Neutral Items, continued

16. When people are given a choice, they usually prefer to remain alive.
a. true
b. ?
c. false

18. A good employee is most often:
a. responsible;
b. ?
c. dishonest.

20. A baby might scream and cry when it is hungry.
a. true
b. ?
c. false

22. Some people like to bicycle and others like to golf.
a. true
b. ?
c. false

24. It costs more to buy a car than a bike.
a. true
b. ?
c. not true

26. A "dopey" person is someone who is not acting smart.
a. true, they don't seem smart
b. ?
c. false, dopey means smart.

28. Bananas and oranges are:
a. two kinds of fruits;
b. ?
c. two kind of vegetables, not fruits.

30. Being brave means you are:
a. not afraid, very courageous;
b. ?
c. simple minded, or unable to use much of your brain.

32. When making an international call, it usually costs more than calling locally.
a. true
b. ?
c. false

Psycholinguistically Neutral Items, continued

34. Some people get nervous when they have to speak in front of a large group of people.
a. true
b. ?
c. false

36. The Atlantic and Pacific:
a. are two animal names;
b. ?
c. are two ocean names

38. If you wanted to find out what direction you were facing you might use a compass.
a. true
b. ?
c. false

40. It is a good idea to look both ways before crossing the street.
a. true
b. ?
c. not true

42. Many people use cars to drive to places like work and grocery stores.
a. true
b. ?
c. false

44. Doctors are people who:
a. work in fields to grow food and crops;
b. ?
c. try to heal people who are sick.

46. A person covered with dirt could become cleaner by;
a. taking a shower;
b. ?
c. rolling in mud.

48. Not all books are about fictional characters, some books are about real people in the world.
a. true
b. ?
c. false

Psycholinguistically Neutral Items, continued

50. If someone wanted to build a house in an area known for heavy flooding, they'd probably want to build their house:
a. at the bottom of a valley;
b. ?
c. high on a mountain

52. "Small" and "little" are opposites of "big and "large."
a. true
b. ?
c. false

54. It is much faster to travel by plane than it is to travel by boat or by bicycle.
a. true
b. ?
c. false

56. If the weather is freezing cold, many people wear clothing to help them stay warm.
a. true
b. ?
c. not true

58. Gasoline is something people put in their cars so that they can drive them.
a. true, gas makes cars runs.
b. ?
c. false, gasoline is a food.

60. One way to ask someone to come to a party is to send them an invitation.
a. true
b. ?
c. false

62. To make movies, you'd probably use:
a. a camera and some film.
b. ?
c. old bits of paper.

64. If someone is allergic to eggs or bread they probably should not eat French toast.
a. true
b. ?
c. false

Psycholinguistically Neutral Items, continued

66. To travel from Rome to New York, you would probably drive a bus or car.
a. true
b. ?
c. false

68. Books and magazines can be found in the library.
a. true
b. ?
c. false

70. Most people go to the beach when it is cold and rainy.
a. true
b. ?
c. false

72. One place where you would probably find sand and hot air is in the desert.
a. hardly ever
b. ?
c. often

74. A log book might be used by the captain of a ship to record events.
a. true
b. ?
c. false

76. A watch is something that:
a. you wear on your head;
b. ?
c. tells you what time it is.

78. When a person is talking about a baseball diamond, they are talking about:
a. three bases and one home plate;
b. ?
c. a gem stone.

80. The largest ocean in the world is much bigger than a pond or a large wading pool.
a. true
b. ?
c. false

Psycholinguistically Neutral Items, continued

82. If a person went to the zoo, they would probably see lots of:
a. animals and birds;
b. ?
c. planets and stars.

84. When you mix blue and yellow together you get green.
a. true
b. ?
c. false

86. It is easier to sleep lying down than standing up.
a. true, it is easier to sleep lying down
b. ?
c. false, it's easier standing up.

88. Most monkeys learn how to ski before they learn to climb.
a. true
b. ??
c. false