ITEM CONTENT OF THE BASIC PERSONALITY INVENTORY

Daryl G. Kroner
E.R. Black
Millhaven Institution

John R. Reddon
Alberta Hospital Edmonton

ABSTRACT

Three raters evaluated the content of the 240 items of the Basic Personality Inventory (BPI) across the four domains (psychological functioning, situation, frequency, and time) proposed by Werner and Pervin (1986). Approximately 80% of the items referred to the present time, two-thirds of the items referred to a specific situation, and half the items referred to frequency of occurrence. Rated areas of psychological functioning corresponded with the putative nature of the content the scales (e.g., most of the items on the Anxiety scale referred to affective-feelings; most items on the Self-Depreciation scale referred to the cognitive-self). Results provide information for assisting in the utilization of the BPI in the assessment process.

INTRODUCTION

The basic building block of a self-report psychological assessment inventory is the test item. Scales are derived from items, inventories from scales, and higher order factors from inventories. Test item content influences what is being assessed, the validity of the inferences that are made from the assessment, and the utility of the assessment. Not only is it important to examine social desirability (Holden & Fekken, 1989; Paulhus, 1984), endorsement frequency (Reddon, Holden, & Jackson, 1983), stability (Parker, Hanson, & Hunsley, 1988), and readability (Reddon & Jackson, 1989), but examining other item characteristics such as reference to time, frequency, situations, affect, and behaviors could also prove beneficial. For example, if the MMPI is used to assess treatment outcome, change may not occur on the standard clinical scales between pre- and post-test because of the large proportion of items that are time bound and that refer to completed and irrevocable past events (Scapinello & Blanchard, 1987).

Using raters, Werner and Pervin (1986) examined the content of items in several structured personality inventories. Content was rated according to four domains: (a) area of psychological functioning, (b) reference to situations,

(c) reference to frequency, and (d) reference to time. They found that different personality inventories used time and reference to situations in varying degrees, suggesting that inventories may differ in their applications based on their item content. Over half the items rated referred to situations which indicated that personality is tapped by combining responses across different situations. Also posited was that the usage of behavioral items in an inventory may have implications for predicting specific behaviors. Their study examined samples of items from several inventories but did not incorporate a complete set of items from any particular inventory.

The purpose of the present study was to examine the item content of the Basic Personality Inventory (BPI; Jackson, 1989) with the same item rating system used by Werner and Pervin (1986). Previous BPI item studies have examined item endorsement and social desirability (Reddon et al., 1983), reliability (Holden, Helmes, Fekken, & Jackson, 1985, Kroner, Reddon, Beckett, 1991), item factor structure (Holden, Reddon, Jackson, & Helmes, 1983) and readability (Reddon & Jackson, 1989). By examining the item characteristics of the BPI according to the domains of psychological functioning, frequency, situations, and time within individual scales, a more complete picture of how personality/psychopathology is assessed with this inventory can be obtained.

METHOD

MATERIALS

The 240 items of the Basic Personality Inventory (BPI) served as the item pool. The BPI is comprised of 12 scales of 20 items each. Eleven of the scales are content scales and there is one critical item scale. Mood and personal/emotional adjustment are measured by the Hypochondriasis, Depression, and Anxiety scales. Scales measuring antisocial tendency include, Interpersonal Problems, Alienation, and Impulse Expression. The Denial scale is intended to be a measure of defensiveness and/or impression management. Persecutory Ideation and Thinking Disorder scales measure cognitive functioning. Self-perception and sociability are measured by the Self Depreciation and Social Introversion scales. Critical deviant behaviors are measured by the Deviation scale. In addition to the 12 standard scales, a measure of social desirability is scored which consists of the two most socially desirable items per scale (excluding Denial and Deviation; Reddon et al., 1983).

RATING SYSTEM AND RATERS

Following Werner and Pervin (1986), there were four descriptive domains against which each item was rated. Werner and Pervin's appendix (p. 628) served as the guide for the raters. Each domain was further divided into specific categories.

The *Area of Functioning Domain* comprised five categories: (a) cognitive-self (beliefs about the self, e.g., "I am a moral person"); (b) cognitive-beliefs, values, opinions (e.g., "People have it in for me"); (c) affective preferences (likes, dislikes, wishes, e.g., "I like to bowl"); (d) affective-feelings (emotions, e.g., "I get anxious a lot"); and (e) behavioral (reports of behavior and activities, e.g., "I often spend time gardening").

The *Situation Domain* was rated according to three categories: (a) setting or circumstance not referred to (e.g., "People are nice to me"); (b) general setting (e.g., "I get depressed at school"); and (c) specific situation (e.g., "I get anxious before an exam").

Frequency Domain had three categories: (a) no specific reference or uncertain reference (e.g., "I am an anxious person"), (b) general or specific reference (e.g., usually, often, rarely), (c) unclear if frequency, intensity, or both (e.g., "In arguments, I am easily downed").

There were five categories in the *Time Domain*: (a) no time reference (e.g., intelligence items); (b) past (e.g., "There have been times when..."); (c) present (includes current beliefs, e.g., "I am daring"); (d) future (e.g., "I hope to become a doctor"); (e) hypothetical (reference to notions like if, would, and might, e.g., "If I am attacked I am likely to fight back").

The BPI booklet (Jackson, 1989) was used to present the items which were independently rated by two psychologists and one psychometrist.

ANALYSIS

Interjudge reliability. Interrater reliability was measured by an uncorrected measure of agreement ($\underline{P}$, Fleiss, 1971). Agreement was computed for the items in each scale according to the four different domains. For example, $\underline{P}$ was measured across the no time reference, past, present, future, and hypothetical categories of the time domain for each scale.

Reliability was also computed using intraclass correlations to assess the amount of agreement across the 12 standard clinical scales. The number of endorsements for scales in each category (cognitive-self, cognitive-beliefs, affective-preferences, affective-feelings, and behavioral) were also tabulated. This reflected the amount that each judge viewed the items from each scale as being characteristic of specific content categories.

Preliminary Analysis. As with the Werner and Pervin (1986) study, distinctions between certain categories could not be reliably made. Two categories were, therefore, collapsed. Within the situation domain, the general settings/description and specific situation categories were combined. Items were then coded as to whether or not an item referred to a situation. The second collapse combined the no specific reference to frequency and the unclear frequency categories. Items were then coded as to whether or not an item referred to frequency.

RESULTS

The results are presented in Tables 1, 2, and 3 for agreement statistics, intraclass correlations, and percentage of content rated according to scale, respectively. Results in each table are briefly discussed but the detailed interpretation of each scale is in the "Summary of Scales" section.

Agreement statistics for the individual scales presented in Table 1 suggest that in general, the amount of agreement among the scales progressively increased from area of functioning to situation, to frequency, to time, with time items having the highest amount of agreement. The mean $\underline{P}$ statistic of agreement for the four domains of area of functioning, situation, frequency, and time were .57, .61, .83, and .90, respectively. For some scales the agreement statistics

TABLE 1
AGREEMENT STATISTICS FOR RATINGS
OF INDIVIDUAL SCALES

Scale	Area of Functioning	Situation	Frequency	Time	Mean
HYP	.53	.70	.90	.90	.76
DEP	.58	.50	.87	.83	.70
DEN	.53	.77	.80	.95	.76
INP	.60	.63	.77	.88	.72
ALN	.67	.73	.73	.93	.77
PEI	.68	.80	.83	.87	.80
ANX	.57	.74	.80	.87	.75
THD	.32	.40	.78	.92	.61
IME	.53	.34	.83	.92	.66
SOI	.62	.73	.90	.92	.79
SDP	.68	.40	.93	.90	.73
DEV	.53	.53	.83	.80	.67
SOD	.58	.63	.83	.95	.75
Mean	.57	.61	.83	.90	

Note. HYP=Hypochondriasis, DEP=Depression, DEN=Denial, INP=Interpersonal Problems, ALN=Alienation, PEI=Persecutory Ideation, ANX=Anxiety, THD=Thinking Disorder, IME=Impulse Expression, SOI=Social Introversion, SDP=Self Depreciation, DEV=Deviation, SOD=Social Desirability.

were quite low for areas of functioning and situations whereas frequency and time agreement statistics were very good for all scales. A P statistic of .95 would suggest that in classifying a randomly selected item, a second randomly selected judge would agree with the initial judgement 95% of the time (Fleiss, 1971).

The results in Table 2 record the intraclass correlations for each category. Due to the low intraclass correlation of the affective-feelings and future categories, they were not emphasized in the summary of the scales. These poor intraclass correlations and the collapsing of the frequency and situational categories demonstrate that some of Werner and Pervin's original categories may have limited utility.

The results in Table 3 are the mean percentage of items assigned to each category within each scale. For example, within the Cognitive-self category of the Denial scale, the percentage of the 20 Denial scale items that were rated as referring to the Cognitive-self was 10% by one rater, 20% by the other rater, and 40% by the last rater. The mean of 23.3% is recorded in Table 3. Within the time

TABLE 2
MEAN RATING AND INTRACLASS CORRELATIONS
FOR SUMMARY SCORES ACROSS ITEMS

Rating Area

Area of Functioning	$\underline{x}$	$\underline{r}$
Cognitive-self	24.3	.85
Cognitive	11.7	.99
Affective-preferences	11.4	.61
Affective-feelings	25.7	.23
Behavioral	25.0	.82
Situation		
Reference to situation	62.6	.93
Frequency		
Reference to frequency	44.9	.96
Time		
Past	9.2	.78
Present	78.8	.88
Future	2.6	.51
Hypothetical	9.5	.75

domain, no reference was made to the no time category, consequently, this category was deleted from the tables. The percentages referred to in the Summary of Scales section of this paper were derived from Table 3.

SUMMARY OF SCALES

Scales are presented according to the ordering given in the materials section.

Hypochondriasis (HYP). The behaviors category was the most frequently referred to within area of functioning. Reference to this behavior was the second highest of all the scales and the highest among the mood and personal adjustment scales. Surprisingly, 31.7% of the reference was made to descriptions or beliefs about the self, as one would expect the scale to deal more with physical complaints than descriptions or beliefs about the self. Minimal reference to the past would exclude using this scale solely as an indication of past dysfunctioning. No reference was made to preferences or beliefs about others. Reference to situation was made 51.7% of the time and frequency 60% of the time. The present was referred to 95% of the time. This scale is, therefore, reflective of a respondent's present status.

Depression (DEP). Reference to describing self was made 38.3% of the time and reference to feelings one-third of the time. Affective preferences and behaviors were represented about 10% of the time with cognitive beliefs being

TABLE 3
PERCENTAGE OF ITEMS RATED IN CONTENT CATEGORIES ACCORDING TO SCALE

	HYP	DEP	DEN	INP	ALN	PEI	ANX	THD	IME	SOI	SDP	DEV	SOD
Area of Functioning													
Cognitive-self	31.7	38.3	23.3	11.6	16.7	13.3	13.3	40.0	25.0	13.3	55.0	33.3	41.7
Cognitive-beliefs	0.0	6.7	1.7	21.7	36.7	41.7	8.3	3.3	0.0	3.3	16.7	0.0	10.0
Affective-prefer.	0.0	11.7	13.3	20.0	18.3	3.3	0.0	0.0	6.6	51.7	3.3	8.3	10.0
Affective-feelings	25.0	33.3	26.7	25.0	10.0	35.0	66.6	18.3	16.6	16.7	11.7	23.3	13.3
Behavioral	43.3	10.0	35.0	21.7	18.3	6.7	11.6	38.3	51.6	15.0	13.3	35.3	25.0
Reference to sit.	51.7	36.7	73.3	70.0	76.7	73.3	76.7	70.0	50.0	76.7	48.3	48.3	48.3
Reference to freq.	60.0	36.7	60.0	30.0	36.7	50.0	53.3	66.7	63.3	15.0	15.0	51.7	33.3
Time													
Past	3.3	5.0	21.7	6.7	13.3	11.7	3.3	3.3	1.6	1.7	6.7	31.7	10.0
Present	95.0	80.0	73.3	73.3	46.7	76.7	93.3	95.0	86.7	90.0	80.0	55.0	73.3
Future	0.0	13.3	0.0	1.7	5.0	1.6	1.7	0.0	6.7	0.0	1.6	0.0	0.0
Hypothetical	1.7	1.7	5.0	18.3	35.0	10.0	1.7	1.7	5.0	8.3	11.7	13.3	16.7

Note: See Table 1 for abbreviations.

represented the least. Reference to a situation and to frequency was made 36.7% of the time. This limited reference to situations would suggest that depression is measured through a global process. Eighty percent of the items referred to the present. DSM-III-R (American Psychiatric Association, 1987) characterizes a Major Depressive Episode as either a depressed mood or a loss of interest or pleasure on all or almost all activities. Approximately two thirds of the items in the Depression scale would encompass the above definition. Consequently, the Depression scale may include some additional correlates of depression that are not delineated in the DSM-III-R Major Depressive Episode.

Anxiety (ANX). Reference to affective feelings was made two-thirds of the time which is greater than for any other scale, indicating that this scale is tapping internal processes which are minimally effected by belief systems, behaviors, or affective preferences. With only 11.6% reference to behaviors, there was minimal emphasis on overt symptomology of anxiety. Similar to the Hypochondriasis scale, no reference was made to affective preferences. Reference to cognitive-self and cognitive-beliefs was 13.3% and 8.3%, respectively. Reference to situations was equivalent to Social Introversion and was greater than any of the other scales. Minimal reference was made to the past and the majority of items reflected affective states. The anxiety scale should be useful in the evaluation of anxiety disorders or treatments which predominantly involve emotions and feelings situationally based as compared to assessing general outward manifestations of anxiety. Having minimal reference to the past will not indicate if the anxiety is of a chronic nature.

Interpersonal Problems (INP). The Interpersonal Problems scale tapped all areas of psychological functioning as measured in this study more uniformly than the other scales. This scale gives a broad indication of a current antisocial orientation with regard to interactions with others. Reference to the present was made 73.3% of the time and hypothetical 18.3% of the time.

Alienation (ALN). Cognitive beliefs, values and opinions predominate with minimal reference to feelings which coincides with the emphasis of the scale on measuring attitudes. Compared to the other antisocial scales, cognitive-beliefs were endorsed more frequently. Reference to situation was the highest of the antisocial scales at 76.7%. With 36.7% of the items referring to frequency, this scale predominantly taps attitudes that are based on beliefs and preferences that make minimal reference to rate of occurrence. Compared to the other scales, the present was the least referred category and hypothetical was the most referred category. Consequently, elevations on this scale may be amenable to treatment by providing insight into the clients affect, beliefs, and behaviors.

Impulse Expression (IME). The manual considers this scale to indicate one who fails to consider the consequences of their actions and is prone to undertake risky and reckless actions. In concert with this definition, 51.6% of the items refer to behaviors, which was greater than any other scale. Beliefs about the self and affective feelings encompass over 40% of the items. Consequently, self-perception is an integral component of the Impulse Expression scale. This scale is comprised of more than just an outward expression, as the description and name of the scale would suggest. Elevated raw scores will include self-perceptions and feelings of impulsivity and will be indicative of impulsiveness expressed through a nonsituational, general personality characteristics. Reference to situations was

lower than the other two antisocial scales. Reference to frequency was made 63.3% of the time suggesting that a frequency-based criterion will have a major impact on scale scores. This reference to frequency occurred more than with the other antisocial scales. Minimal reference was made to the past and hypothetical categories.

Persecutory Ideation (PEI). More reference was made to cognitive beliefs than any other scale. If the cognitive beliefs and affective feelings categories were collapsed, then over three-quarters of the items referred to components that seem essential in assessing persecutory ideation. Consequently, elevated raw scores are going to be reflective of beliefs, values, opinions, or emotions. Having items refer to the beliefs about others category allows a framework for victimization by others to be tapped which is included in the manual's definition of this scale. Minimal reference was made to behaviors or affective preferences. The cognitive-self was referred to 13.3% of the time. Reference to situation and to frequency were made 73.3% and 50% the time, respectfully. In the time domain, approximately three quarters of the references were made to the present. The past was referred to 11.7% of the time. The Persecutory Ideation scale received the highest amount of agreement (P) across all four domains when compared to the other scales.

Thinking Disorder (THD). Due to the low amount of agreement among raters for this scale in the area of psychological functioning, only frequency, situation, and time dimensions will be discussed. Reference to situations was made 70% of the time with reference to frequency being made 66.7% of the time. Within the time domain 95% of the items refer to the present although the manual indicates there may be some carry over effects from past substance or alcohol abuse.

Social Introversion (SOI). By definition, this scale refers to a preference for solitary activities. Correspondingly, the majority of the items were rated as referring to preferences, such as likes, dislikes, and wishes. Cognitive-self, affective feelings, and behaviors received almost equal ratings. Minimal reference was made to cognitive beliefs. Elevated raw scores will refer to a specific situation in which introversion is currently being endorsed. This will be done through minimal reference to frequency. Ninety percent of the time reference was made to the present and 8.3% to hypothetical.

Self-Depreciation (SDP). According to the manual, this scale measures poor self-image and a pessimistic attitude about self and future. Cognitive beliefs about the self constituted the majority of items. Because only 11.7% of the items were reflective of feelings, elevated scores would not reflect how the respondent was feeling (e.g., feeling worthless), but one's cognitive beliefs (e.g., "I think I would make a very good leader"). Elevated scores would then be reflective of the respondents belief system about the self. This scale, therefore, would not be useful in measuring changes in self-perception that are affectively based, e.g., evaluating treatment which focuses on affect rather than cognitive beliefs. Reference to a situation was made 48.3% of the time and reference to frequency 15% of the time. With minimal reference to frequency, generalized traits are being tapped. Minimal reference was made to affective preferences. With regard to time, reference was made to the present 80% of the time. Reference to the past was made 6.7% of the time and hypothetical was referenced 11.7% of the time.

Although the manual suggests that this scale indicates attitudes about the future, minimal reference was made to the future.

Deviation (DEV). The Deviation scale is a critical item scale which has items that refer to areas such as suicide and substance abuse. Beliefs about the self and behaviors make up most of the items. Both reference to a situation and frequency was made about one-half of the time. In contrast to other scales, the Deviation scale made more reference to the past. Consequently, even though this scale contains critical items that may be the target of intervention, it may not be sensitive enough for measuring pre-post intervention effects. The present was referred to 55% of the time with the hypothetical being referred to 13.3% of the time.

Denial (DEN). The BPI manual suggests that the Denial scale measures defensiveness, denial of normal affective responses, and repression of unpleasant or emotionally charged cognitions. Although the present analysis indicated a cognitive and affective component to the scale, there was more reference made to behaviors. Also, 21.7% of the items referred to the past. It appears, therefore, that the Denial scale is not measuring the construct as defined in the manual. The high reference to situations implies that the denial process is measured through a range of different situations. Consequently, this scale may be of a more concrete and specific nature than a more global denial process. This reliance on behaviors could contribute to the scale having lower test-retest stability (Holden et al., 1985) and lower validity coefficients (Holden, Fekken, Reddon, Helmes, & Jackson, 1988). The amount of agreement ($\underline{P}$) for the time domain of the Denial scale was equivalent to that of Social Desirability which was higher than any other scale.

Social Desirability (SOD). Although intended to measure a response style, some have argued that facets of social desirability also tap content areas. This scale was, therefore, included in the analyses. Reference to cognitive feelings and affective preferences was made 10% of the time. Behaviors were referred to one quarter of the time. Cognitive self was referred to 41.7% of the time which was a high rating when compared to other scales. Reference to situation occurred 48.3% of the time and frequency was referred to one third of the time. In the time category, the present was referred to 73.3%, past 10%, and hypothetical 16.7%.

DISCUSSION

Similar to the inventories examined in the Werner and Pervin (1986) study, the BPI was derived from items that measured cognitions, preferences, emotions, and behaviors. With regard to time, the BPI assessed areas that are mostly concerned with the present. Very little indication of the past, future, or hypothetical was made. Consequently, the BPI's primary function is to reflect how the respondent is currently functioning. With about two thirds of the items being viewed through situations, the majority of reference to personality in the BPI was made through specific situations. Frequency of occurrence was referred to slightly less than one-half of the time.

Typically a single scale score from an inventory is used to represent a content area. There is, though, supplementary item content information that can be derived from items that may prove beneficial. By using these supplementary criteria to examine scale content, it can be suggested that different scales tap psycho-

pathology differently. Knowing that different scales utilize actions, thought processes, and experiences differently, appropriate use of the BPI and specific scales for evaluating different types or modes of intervention (i.e., cognitive versus behavioral) can be made. Also, knowing that the majority of items refer to the present may make the BPI a more appropriate instrument in measuring treatment outcome than other inventories, such as the MMPI, which have a large proportion of items that are time bound (Scapinello & Blanchard, 1987). Consequently, the different ways each scale assesses content should be taken into account when determining the purpose of the assessment.

A closer look at the way situations are categorized within scales could prove useful in assessing specific constructs. The role of situations is well known in that people with elevated personality trait scores will vary in their expression of that trait across different situations. In examining categories of situations, Wright and Mischel (1987) suggested that analysis of the situation-competency requirements will aid in predicting behavior from dispositional constructs. They defined situation-competency requirements as psychologically demanding situations in which individual dispositional constructs may be observed with particular clarity. Inquires into the way situations are incorporated in self-report inventories may prove fruitful in examining personality/behavioral relationships.

The way both behavioral and situational items are used in assessment would profit from closer examination. Werner and Pervin (1986) suggested that if the prediction of future behavior is to be measured, an inventory incorporating only behavioral items may be more appropriate in predicting behavior. This may not be the case. A recent report (Holden & Fekken, 1990) that examined the validity of behavioral items suggested that these items per se may not contribute to the validity of a construct. This, in conjunction with valid and targeted constructs increasing the predictability of behavior (Wright and Mischel, 1987), would indicate minimal utility of behavioral items if constructs are used for predicting behavior. To increase the predictive validity of a construct, Wright and Mischel suggested that incorporation of a situational component in items may be necessary. Together, these two studies (Holden & Fekken; Wright & Mischel) indicate that a distinction should be made between situational and behavioral items in future personality item analyses. Making this distinction also provides a basis to derive measures of scalability, a method of assessing consistency which estimates the extent that behavior parallels the situational demands. Scalability may help account for some of previously unaccounted variance between a self-report inventory and target behaviors (Lanning, 1988).

Having a large proportion of items not referring to situations may allow for more varied responses. For example, the item "I have a good deal of energy" will likely be responded to with more variance than an item referring to a specific situation such as "I think that I could commit a crime and get away with it". To illustrate this, the BPI Depression scale can serve as an example. According to the criterion used in the present study, the Depression scale had the least reference to situations. In a recent paper examining differences between computer and paper-and-pencil administration of the BPI among psychiatric patients, the Depression scale coefficient alpha was greater with the computer administration (Holden, Fekken, & Cotton, 1990). There were no other differences between the two forms of administration for scale means, variances, and alphas. Having less reference to

situations may increase the variability of responses via other variables such as mode of presentation and subject population. Consequently, the norming of non-situational items for different populations and different modes of assessment will likely be very important.

One of the intentions of incorporating frequency in the rating system in the Werner and Pervin (1986) study was to reflect the Act Frequency approach to examining personality. Buss and Craik (1983) proposed that examining frequency of acts is the basis for act trends and consequently, dispositional categories. Others (Block, 1989; Moser, 1989), though, have criticized this approach to personality assessment. Although the rating of frequency was an attempt to evaluate the Act Frequency approach, the way frequency was rated in the Werner and Pervin study does not provide the critical information to determine if an act frequency element is incorporated. In addition, as indicated by Werner and Pervin, the reference to frequency made little contribution because of the amount of frequency was not taken into account in the true-false answering format. For example, no differentiation was made between "always" or "very seldom". Furthermore, Holden and Fekken (1990) suggest that the usage of absolutes such as "always" or "never" can be endorsed or rejected because the item is uncharacteristic or because nothing can be categorically ruled out.

In summary, the nonscale item content for the different scales of the BPI corresponded with the area being assessed (i.e., more anxiety scale items referred to affect than to beliefs about others). Examining these item characteristics enables further inquiry into the different ways personality and psychopathology are assessed via self-report inventories. These nonscale item characteristics are not only important considerations for how personality is derived and measured (Holden & Fekken, 1990), but they may also be useful for assessing the clinical utility of inventories. Moreover, there may be implications for predicting future behavior from self-report. If a domain is adequately sampled (Jackson, 1971) and if personality dimensions, including nonscale item content, are reflective of target behaviors (Wright & Mischel, 1987; Jackson & Paunonen, 1985), then one can expect greater validity of self-report inventories and a higher congruence between self-report and behaviors.

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.) Washington, DC: Author.

Block, J. (1989). Critique of the act frequency approach to personality. *Journal of Personality and Social Psychology, 56,* 234-245.

Buss, D.M., & Craik, K.H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105-126.

Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76,* 378-382.

Holden, R.R., & Fekken, G.C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers. *Journal of Research in Personality*, *23*, 180-191.

Holden, R.R., & Fekken G.C. (1990). Structured psychological test item characteristics. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*, 35-40.

Holden, R.R., Fekken, G.C., & Cotton, D.H.G. (1990). Clinical reliabilities and validities of the microcomputerized Basic Personality Inventory. *Journal of Clinical Psychology*, *46*, 845-849.

Holden, R.R., Fekken, G.C., Reddon, J.R., Helmes, E., & Jackson, D.N. (1988). Clinical reliabilities and validities of the Basic Personality Inventory. *Journal of Consulting and Clinical Psychology*, *56*, 766-768.

Holden, R.R., Helmes, E., Fekken, G.C., & Jackson D.N. (1985). The multidimensionality of person reliability: Implications for interpreting individual test item responses. *Educational and Psychological Measurement*, *45*, 119-130.

Holden, R.R., Reddon, J.R., Jackson, D.N., & Helmes, E. (1983). The construct heuristic applied to measurement of psychopathology. *Multivariate Behavioral Research*, *18*, 37-46.

Jackson, D.N. (1971). The dynamics of structured personality tests: 1971. *Psychological Review*, *78*, 229-248.

Jackson, D.N. (1989). *The Basic Personality Inventory manual*. Port Huron, MI: Research Psychologists Press.

Jackson, D.N., & Paunonen, S.V. (1985). Construct validity and the predictability of behavior. *Journal of Personality and Social Psychology*, *49*, 554-570.

Kroner, D.G., Reddon, J.R., & Beckett, N. (1991). Basic Personality Inventory clinical and validity scales: Stability and internal consistency. *Journal of Psychopathology and Behavioral Assessment*, *13*, 147-154.

Lanning, K. (1988). Individual differences in scalability: An alternative conception of consistency for personality theory and measurement. *Journal of Personality and Social Psychology*, *55*, 142-148.

Moser, K. (1989). The act-frequency approach: A conceptual critique. *Personality and Social Psychology Bulletin*, *15*, 73-83.

Parker, K.C.H., Hanson, R.K., & Hunsley, J. (1988). MMPI, Rorscharch, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin*, *103*, 367-373.

Paulhus, D.L. (1984). Two-component models of social desirable responding. *Journal of Personality and Social Psychology*, *46*, 598-609.

Reddon, J.R., Holden, R.R., & Jackson, D.N. (1983). Desirability and frequency of endorsement scale values for items of the Basic Personality Inventory. *Psychological Reports*, *52*, 619-633.

Reddon, J.R., & Jackson, D.G. (1989). Readability of three adult personality tests: Basic Personality Inventory, Jackson Personality Inventory and Personality Research Form E. *Journal of Personality Assessment*, *53*, 180-183.

Scapinello, K.F., & Blanchard, R. (1987). Historical items in the MMPI: Note on evaluating treatment outcomes for a criminal population. *Psychological Reports*, *61*, 775-778.

Werner, P.D., & Pervin, L.A. (1986). The content of personality inventory items. *Journal of Personality and Social Psychology, 51*, 622-628.

Wright, J.C., & Mischel, W. (1987). A conditional approach to dispositional constructs: The local predictability of social behavior. *Journal of Personality and Social Psychology, 53*, 1159-1177.

Author Notes