

Cautionary Comments Regarding the Myers-Briggs Type Indicator

David J. Pittenger *The University of Tennessee at Chattanooga*

The Myers-Briggs Type Indicator (MBTI; K. C. Briggs & I. B. Myers, 1998) is a popular measure of normal personality that its promoters claim has many applications. M. H. McCaulley (2000) offered an optimistic and enthusiastic account of how counselors can use this instrument in corporate settings. The present article evaluates several of the psychometric limitations and criticisms of the MBTI that warrant considerable caution when making inferences from its 4-letter type formula. The author concludes that the MBTI, while offering much intuitive appeal, may not yet be able to support the claims its promoters make.

McCaulley (2000) recently provided an optimistic and enthusiastic account of how consulting psychologists can and should integrate the Myers-Briggs Type Indicator (MBTI) into consulting work with business management. More specifically, she suggested that consulting psychologists can use the MBTI to help employees and managers enhance their interpersonal relations and thereby improve their ability to work effectively for the corporation. The fundamental assumption presented in McCaulley's article is that knowledge of coworkers' personality preferences, or MBTI type, will facilitate greater respect for individual differences, aid in assigning work responsibilities, and foster effective collaboration among employees. Indeed, McCaulley advocated using the MBTI as a component of the employee selection process. In addition, McCaulley described what she believed to be the striking difference in the worldview of counselors and managers and offered recommendations to help counselors convey the utility of the MBTI to corporate decision makers.

Although the MBTI is an extremely popular measure of personality, I believe that the available data warrant extreme caution in its application as a counseling tool, especially as consultants use it in various business settings. McCaulley (2000) offered decisive conclusions that attest to the validity of the MBTI four-letter type formula and its utility for consulting work. Unfortunately, she offered little empirical evidence to support the veracity of those claims. As such, these conclusions require additional examination.

Currently, the role of personality assessment in employment settings is receiving renewed interest. As Hogan, Hogan, and Roberts (1996) noted, although there remains considerable skepticism regarding the value of personality assessment in applied settings, they believe that well-constructed measures of normal personality can serve a useful role in personnel selection and development. Such an endorsement depends upon measures that evidence

David J. Pittenger is the Associate Provost for Academic Administration at the University of Tennessee at Chattanooga. Prior to this position he was the Department Head of the Department of Psychology and taught a variety of courses including Statistics, Research Methods, Learning Theory, and Introductory Psychology. His research interests include persistence of human behavior, coping behaviors, and psychometrics.

Correspondence concerning this article should be addressed to David J. Pittenger, Department of Psychology, The University of Tennessee at Chattanooga, 615 McCallie Avenue, Department 2508, Chattanooga, TN 37403. E-mail: david-pittenger@utc.edu

construct validity and afford useful predictions. Given the ramifications of applying personality measures in the work setting, it is necessary to maintain a critical analysis of how practitioners use tests in applied settings. In essence, I do not believe that the relevant data justify the conclusion that the MBTI is a direct measure of Jung's (1921/1971) theory of personality types. Furthermore, I do not believe that the current scoring procedures for the MBTI allow one to make important prophecies about individuals (see also, Pittenger, 1993).

In the remainder of this article, I review what I believe to be important psychometric data that raise significant caveats regarding the MBTI and its applications. The primary goal of this review is to provide an alternative perspective of the MBTI and its utility as a measure of personality. There are important consequences when psychologists apply personality measures in the work setting. The results of the test affect the individuals who are tested, the corporations who use those test results, and our profession to the extent that we endorse specific measurement practices and inferences from test scores. Consequently, it is necessary to maintain a critical analysis of the issues and ensure a lively dialectic regarding the use of normal personality assessments in corporate and other applied settings.

To organize my arguments, I have divided this article into four parts. First, I offer an abridged review of the theoretical structure of the MBTI, its scoring, and its interpretation. According to Myers, McCaulley, Quenk, and Hammer (1998), Briggs and Myers developed the MBTI as a measure of the personality types first articulated by Jung (1921/1971). The strong theoretical background of the MBTI affords many falsifiable hypotheses regarding the psychometric properties of the MBTI results. The second section reviews the relevant psychometric data that assess these hypotheses. The third section reviews the

relevant empirical literature that has examined the utility of the MBTI for industrial/organization applications. In the final section, I suggest that many of the uses of the MBTI, as endorsed by McCaulley (2000), lack empirical support, and that consulting psychologists should consider these facts before using the instrument. I also suggest programs of future research that might address several important matters I raise.

Theoretical Foundations of the MBTI

Before proceeding with a review of the empirical features of the MBTI, it is important to examine its theoretical foundation as an explication of Jung's (1921/1971) theory of personality. This foundation is essential for interpreting the psychometric properties of the instrument and the utility of its application for assessing an individual's personality. I will forego a detailed review of Jung's theory and the particular structure of the MBTI as other authors provide lucid and detailed reviews (Jung, 1921/1971; Macdaid, McCaulley, & Kaniz, 1986; McCaulley, 2000; Myers et al., 1998).

The important feature that warrants attention is the fact that Jung's theory and the MBTI are typologies (McCaulley, 2000). More specifically, the instrument treats personality types as distinctive groups. This perspective suggests that there are quantitatively and qualitatively different populations of people who express different personality characteristics. In other words, these populations will demonstrate relative homogeneity of variance within groups and heterogeneity of variance between groups (Block, 1971; Block & Ozer, 1982).

Most personality tests represent a trait perspective that characterizes personality as a construct best measured using a continuous variable that ranges between two extremes. For example, common measures of introversion and extroversion (e.g., NEO Personality Inventory [NEO-PI], Costa & McCrae, 1985) treat these constructs as op-

posing poles along a continuous scale. Scores derived from trait measures of personality reflect the degree or magnitude of the personality construct. Although casual descriptions of the trait may lapse into “type-as-label” discussions (Block & Ozer, 1982; e.g., Extrovert vs. Introvert), the underlying assumption is that the scale is continuous and that the “types” refer to the extremes of the population or scale. Moreover, trait theories presume that the majority of scores cluster near a common point along the scale, with fewer scores at the extremes, thus forming a unimodal and relatively symmetrical distribution.

By contrast, the MBTI is a type theory that emphasizes the 16 unique categories of personality created by the four type pairs (Extraversion¹-Introversion [EI], Sensing-Intuition [SN], Thinking-Feeling [TF], and Judging-Perceiving [JP]). The EI, SN, and TF dimensions derive from Jung’s (1921/1971) theory of personality. The JP scale is an embellishment that Myers and Briggs added to their interpretation of the theory (see Saunders, 1991, for a history of the MBTI and a biography of its creators).

The four-letter type formula (e.g., ENFP or ISTJ) is the most salient characteristic of the MBTI as it forms the central feature for interpreting the results of one’s MBTI results. When scoring the instrument, one converts the observed score for each of the four scales to a letter. Because the four scales each have two dimensions, there are 16 letter combinations, each representing a unique and distinct personality profile. In an earlier version of the MBTI, individuals who scored at the center of one of the scales received an *x* rather than a letter code to illustrate the intermediacy of the person’s score. The developers of the MBTI subsequently abandoned that practice and replaced it with a tie-breaking procedure to avoid assigning intermediate scores (Myers et al., 1998).

Because promoters of the MBTI view the type conditions as mutually exclusive

categories, the practice of test interpretation focuses on the four-letter type formula. As Carskadon (1979a) noted,

In general, continuous scores on the MBTI are probably overused. However much it may seem to make better intuitive or psychometric sense to use continuous scores in many instances, it should be remembered that the type scales are theoretically dichotomous in nature, and the indicator was designed with this in mind. In general it is probably better to use dichotomous classification except where there are specific considerations to the contrary. (p. 20)

McCaulley’s (2000) description of the MBTI clearly endorsed this perspective. In fact, McCaulley quoted Jung describing the fundamental nature of type development and then observed that “the most important unit of measurement for the MBTI is the four-letter type formula that indicates the choices for the four preferences plus their dynamic interaction” (McCaulley, 2000, p. 121). Both Jung’s theory and supporters of the MBTI treat personality as an invariant that is set at birth and tempered by experience. Thus, normally functioning adults have well-established, unambiguous, and stable personality preferences. McCaulley also affirmed that the MBTI follows type structure by asserting that the dimensions measured by the MBTI are dichotomous and not preferences that exist across a continuous scale that ranges from one pole to another. Finally, McCaulley asserted that the MBTI assesses the fundamental personality features that influence an individual’s cognitive processes. As I will illustrate subsequently, the type perspective and the dichotomous classification procedures are problematic. In addition, I will show that the link between the personality dimensions measured by the MBTI, cognition, and work behaviors is tenuous.

¹The spelling of extraversion, as opposed to extroversion, is unique to the MBTI and follows the practice established by Jung (1921/1971).

Psychometric Properties of the MBTI

Descriptive Statistics

Given the type theory orientation of the MBTI, one would predict between-group heterogeneity of variance and within-group homogeneity of variance for each of the four type preference dimensions. This prediction implies that the distribution of scores should be bimodal, with greater relative difference between the two preference types compared with the variance of scores within each type. In addition, the nature of the type preference predicts that few individuals should score at the midpoint of the scale. Although bimodality appears to be an essential characteristic of the distributions of scores, it is conspicuously absent. Evidence from several sources (Harvey & Murry, 1994; Hicks, 1984; McCrae & Costa, 1989; Stricker & Ross, 1962) indicates a continuous distribution of scores across each dimension. More recently, Bess and Harvey (2002) replicated the finding of unimodality of scores using the item response theory scoring scheme promoted in the current edition of the measure (Myers et al., 1998). As I have noted elsewhere (Pittenger, 1993), even the data presented in the MBTI manual provide only tentative evidence for a discontinuous scale and greater evidence for a continuous measure of personality traits.

The lack of bimodality and the high frequency of scores at the midpoint of the scales have profound implications for interpreting the MBTI's four-letter type formula. There is no evidence of separate populations of personality types using the standard scoring procedure. Thus, concluding that an *E* type is qualitatively different from an *I* type is indefensible unless there are corresponding data to suggest that the difference between the scale scores is sufficiently large to support such a distinction. Using a conventional 95% confidence interval, the standard error of measurement (SEM) for each scale is 20 or larger (Har-

vey & Murry, 1994; Pittenger, 1993). From a statistical perspective, the MBTI four-letter type formula may imply statistically significant personality differences where none exists.

Ipsative and dichotomous scaling procedures, such as those used by the MBTI, have several disadvantages and limit the ability to make reasonable prophecies of a person's behavior using the four-letter type formula. As a generality, using dichotomous scores reduces the predictive power of a continuous scale (Hunter & Schmidt, 1990) and can greatly increase the rate of Type I errors (Maxwell & Delaney, 1993). Specifically, converting continuous interval or ratio scaled scores to an ordinal or nominal scale removes important information, especially information related to variability.

Cohen (1983) demonstrated that dichotomizing a continuous scale reduces the shared variance with another variable by $.64(r^2)$; dichotomizing both scales reduces the shared variance by $.40(r^2)$. By implication, the four-letter type formula limits its own predictive power, especially if the criterion is a binary condition. Consequently, counselors who advocate the MBTI as a component of the job-selection or team-forming process, as McCaulley (2000) recommended, may well be promoting a decision-making tool whose own procedures reduce its predictive validity, especially in cases where the final decision is dichotomous (e.g., hire vs. do not hire).

Tenopyr (1988) offered a more emphatic caution regarding the use of forced-choice instruments (of which the MBTI is a member). Her analysis of the internal consistencies among scales using forced-choice items revealed the potential for considerable artifactual and inflated reliability estimates. Tenopyr thus concluded, "It appears that construct interpretation on the basis of forced-choice scales should be made with extreme caution. *In the case of some inventories, it would probably be better not to make these interpretations, or at*

least not rely on them heavily, with respect to important decisions about individuals” (p. 751; italics added). Applying item response theory procedures to examine the MBTI, Harvey and Murry (1994) found that the dichotomizing procedure produced between 26% and 32% loss of information for each of the scales.

These observations are troublesome as there are notably little data to support the type structure of the MBTI. Furthermore, the practice of converting scale scores to type categories reduces any predictive value that the test may afford. Consequently, those who use the conventional MBTI scoring practice to make inferences about others’ personality risk reaching conclusions that cannot be empirically justified.

Test-Retest Reliability

The theoretical framework supporting the MBTI provides a unique perspective for interpreting the test–retest reliability of the instrument. Recall that Jung and the promoters of the MBTI treat personality as an invariant that is set by adulthood. As such, one would predict that test-retest reliabilities will be high, especially for the four-letter preference categories.

There are several reports of the test-retest reliabilities of the four dimensions of the MBTI (Carskadon, 1977, 1979b; Howes & Carskadon, 1979; Stricker & Ross, 1962). These reports offer a consistent pattern that suggests that the reliability of the MBTI does not meet expectations derived from its theory. For example, Stricker and Ross found that across a 14-month period the reliabilities ranged from a low of $r(38) = .48$ for the TF scale to a high of $r(38) = .73$ for the EI scale. Schuerger, Zarrella, and Holtz (1989) also demonstrated that the attenuation of the reliability of the MBTI decreases at a rate over retest intervals that is comparable to other measures of personality. More recent reports (Capraro & Capraro, 2002; Salter, Evans, & Forney, 1997) confirm previous

reviews of the test-retest reliabilities of the MBTI scores. Although one might consider these reliabilities acceptable for conventional trait measures of personality, they are at odds with a type theory that predicts that nonpathological personality preferences should become and remain stable early in life.

Howes and Carskadon (1979) provided data that raise additional and equally important questions regarding the reliability of the four-letter type score. Their analysis indicated that a large portion of their participants received different type profiles when retested. Not surprisingly, the greatest proportion of changes occurred when the initial preference score was close to the middle of the scale (1 to 15 points on either side of the midpoint). When the initial score was within this intermediate range, 32% of the EI, 25% of the SN, 29% of the TF, and 30% of the JP labels shifted on the second testing. McCarley and Carskadon (1983) replicated these findings and demonstrated that across a 5-week test-retest interval, 50% of the participants received a different classification on one or more of the scales. Indeed, Myers et al. (1998) reported that 35% of individuals had a different four-letter type score after a 4-week interval.

These results are not surprising given the center-heavy distribution and heterogeneity of variance of the scale scores. Nor are these changes trivial. If we are to presume that “an extraverted sensing type will show extraversion differently from an extraverted thinking type” (McCaulley, 2000), then the alteration of one or more of the four-letter type formula represents a considerable change in personality. These data also raise profound questions regarding the advisability of using the four-letter typing system, while ignoring the magnitude of the scale scores, and raise questions regarding the veracity of any type interpretations for individuals with scale scores close to the midpoint of the scale.

Kummerow (1988) and Walck (1992) provided further evidence for the tenuous nature of the four-letter type formula. For example, Walck administered the MBTI to a large sample of individuals and then asked the participants to describe, in writing, their type preferences. Walck found that as many as a third of the participants believed that the MBTI four-letter type formula mislabeled them. Specifically, there were considerable inconsistencies regarding the TF and EI dimensions. Walck also reported that these disagreements were most likely to occur when the scale scores were close to the middle of the scale, thus raising further questions about the appropriateness of using type categories for individuals who do not evidence a clear preference for a specific MBTI personality dimension.

Factor Analytic Analysis

Factor analytic studies of the MBTI reveal consistent findings that also question the construct validity of the instrument. First, the factor structure is not consistent with the design or theory of the MBTI. For example, Sippes, Alexander, and Freidt (1985) found a factor structure that was inconsistent with the four factors specified by the MBTI and found only a marginal relation between these factors and the MBTI scales. Other researchers have reported findings that are inconsistent with MBTI theory (e.g., Johnson & Saunders, 1990; McCrae & Costa, 1989; Saggino, Cooper, & Kline, 2001; Saggino & Kline, 1996; Sippes & Alexander, 1987; Sippes & DiCaudo, 1988; Stricker & Ross, 1962; Thompson & Borrello, 1986a, 1986b). Similarly, Lorr (1991) found that cluster analytic techniques failed to produce results consistent with MBTI theory and, like others, faulted the instrument's scoring procedures used to categorize individuals.

A second problem is the correlation among the MBTI scales. To reiterate, promoters of the MBTI treat the four scales as

independent dimensions of one's personality preferences. If the scales are independent measures of different dimensions of personality, there should be little or no intercorrelation among the scales. The available data, however, suggest that there is a sizable correlation among the scales (Berr, Church, & Waclawski, 2000; Myers et al., 1998; Sippes & DiCaudo, 1988).

Other inconsistencies arise when one compares MBTI scores with the scores of other measures of personality. For example, the EI scale of the MBTI correlates with other measures of extroversion, including Eysenck's (Sippes & Alexander, 1987) or a version of the Big Five measure of personality (McCrae & Costa, 1989). This is an interesting finding because Jung's (1921/1971) account of extraversion is substantively different from other theories of extroversion (Pittenger, 1993).

Perhaps the most interesting finding is that the Big Five perspective of personality can readily subsume the variance explained by MBTI. For example, McCrae and Costa (1989) concluded that a five-factor model of personality more efficiently explains the MBTI factor structure. Furnham (1996) reached a similar conclusion after examining the correlation of scores between the subscale MBTI and the NEO-PI (Costa & McCrae, 1985).

Although McCrae and Costa (1989) argued that the MBTI is derivative of their version of the Big Five, it is possible that NEO-PI is a special case of the MBTI. By comparison, the MBTI has emerged from a theoretically enriched environment, whereas the Big Five perspective has emerged almost exclusively from empirical analysis (Goldberg, 1990, 1992, 1993). As such, the MBTI affords greater opportunity for systematic analysis of its construct validity. Nevertheless, those who wish to use the MBTI in corporate settings should recognize the depth of evidence supporting the predictive validity of other personality instruments, specifically those derived from

Big Five measures. As Furnham (1996) and Hogan et al. (1996) noted, the NEO-PI appears to have strong support for its utility to predict and explain work-related behaviors. Unfortunately, there are no published data to demonstrate the incremental validity of either instrument as a method for predicting job performance or other work-related behavior. Therefore, conclusions regarding the superiority of either the MBTI or other instruments are, at present, premature.

Summary of the Psychometric Properties of MBTI

A review of the surface features of the MBTI reveals several important inconsistencies between predictions derived from the nomothetic network that defines the MBTI and the empirical data. In total, these inconsistencies raise serious doubts about the inferences and conclusions one can make from the MBTI. Most notable among these problems is the tenuous nature of the dichotomous scoring procedure. I, like McCaulley (2000), belabor the dichotomous nature of the instrument because the entire structure of the MBTI depends upon this perspective of personality. Consequently, these inconsistencies are not trivial. The data related to the distribution of scale scores, the test-retest reliability, and the correlation with other measures of personality also raise doubts that one can make reasonable inferences from a four-letter type formula, especially when an individual's scale scores are close to the mean. Stated from a different perspective, the four-letter type formula may create the impression that there is meaningful difference between the personality profiles of two individuals when no such difference exists.

Analysis of the Utility of the MBTI in Corporate Settings

A central feature of McCaulley's (2000) presentation is the difference in type pref-

erence between managers and counselors. Although the pattern of differences between the two populations appears impressive, one must interpret the data with caution. The information presented by McCaulley represents the four-letter types and not the magnitude of the scale scores; consequently, the reader has no information to determine whether the intensity of these preferences is statistically significant. Therefore, a statistically significant χ^2 in the pattern of type preferences may be artifact of the forced-choice and dichotomous scoring procedure used by the MBTI. For example, Berr et al. (2000) reported the average scaled scores for their sample of senior managers. With the exception of the TF preference pair, all means were within 1 SEM of the center of the scale.

Other research examining the preference types of managers provides notably little information about the magnitude of the type preferences of managers. Gardner and Martinko (1996) conducted a literature review that examined the use of the MBTI in management research. They noted that of the 13 studies that examined the distribution of MBTI type preferences for managers, all were limited to reports of simple frequencies of type preference with no mention of scale scores. In addition, Gardner and Martinko observed that although "these studies indicate the proportions of types within managerial samples, they reveal little about managers' cognitions or behaviors" (p. 53).

Even if we accept the existence of a correlation between personality types and profession, the cause of that relation is, as yet, unclear. Gardner and Martinko (1996) offered three reasonable alternative explanations for the distribution of types among professions, especially those observed in corporate settings. The first is a conventional hypothesis that links personality to vocational preferences. This hypothesis predicts that people with specific personality types seek out specific professions or

are selected for a profession based on their personality. This hypothesis is the central feature of current research of personnel selection (Hogan et al., 1996). Although the potential link between personality and work performance is an interesting prospect, it is one that has yet to be broadly endorsed by the research community and should be presented, therefore, with caution.

A second account of the results is that factors related to social desirability, impression management, or stereotypes about one's profession bias self-reports on the MBTI. This concern is not unique to the MBTI, however, as it is a common criticism of all personality inventories (Hogan et al., 1996). Evidence provided by Walck (1992), however, suggests that many people may disagree with their assigned type category and underscores McCaulley's (2000) recommendation that individuals have the opportunity to determine the validity of the reported type.

Gardner and Martinko's (1996) third alternative hypothesis suggests that working in a specific managerial role may enhance or alter one's expression of personality. Because the majority of the research on type preference and job classification has been cross-sectional, it is not possible to resolve whether it is the personality preference that affects the job selection or if the selected job affects the personality profile. This problem is not unique to the MBTI, however, as few researchers have applied longitudinal methodologies to examine the relation between personality and job performance. Although this is the most speculative of the alternative hypotheses, it does suggest another caution regarding use of the MBTI.

McCaulley (2000) noted that the use of the MBTI can quickly become pervasive throughout an organization as individuals gladly share their MBTI type formula with others. Although a free exchange and discussion of individual differences among coworkers is admirable, it does raise the

specter of the fundamental attribution error. In other words, making personality a salient component of individual interactions ("Hello, I'm INTJ.") may exaggerate the tendency of individuals to make dispositional attributions while ignoring situational conditions that affect behavior. Such exaggerations can lead to unwarranted conclusions regarding the use of the MBTI. Elsewhere, I (Pittenger, 1994) have described my experiences with a wholesale application of the MBTI in an academic setting that resulted in the dissemination of dubious advice to students regarding their selection of courses and majors.

Reading McCaulley's (2000) description of the type preferences suggests that situational factors have little or no influence on individuals' behaviors or cognitions. Indeed, little in the MBTI theory appears to acknowledge the Person \times Situation interaction that is a common component of contemporary social-cognitive theory (Mischel & Shoda, 1995; Shoda & Mischel, 2000). For example, Mendoza-Denton, Ayduk, Mischel, Shoda, and Testa (2001) recently presented experimental evidence of Person \times Situation encoding of experiences in complex social situations. This encoding process attenuates acceptance of global stereotypes and enhances the analysis of self in different social contexts. Such results echo the findings of Gardner and Martinko (1996), who reviewed evidence that supports the conclusion that context may dominate one's decision style. Consequently, those who rely exclusively upon MBTI personality types to explain differences among persons may overlook an important contributor to differences among persons.

Conclusion

Gardner and Martinko (1996) provided a comprehensive review of the empirical literature examining the relation between the MBTI and various work-related behaviors. Their evaluation of the MBTI repre-

sents, at best, a guarded and conditional endorsement. Although they noted that there are studies that imply a link between the MBTI types and some important work performance measures, they were also careful to highlight the many methodological limitations of those studies. Furthermore, their review of the empirical link between personality types and work performance measures warrants greater qualification than McCaulley's (2000) comments imply. Specifically, they concluded that, "It is clear that efforts to detect simplistic linkages between type preferences and managerial effectiveness have been disappointing. *Indeed, given the mixed quality of research and the inconsistent findings, no definitive conclusion regarding these relationships can be drawn*" (Gardner & Martinko, 1996, p. 77, italics added). This conclusion echoes the conclusion of other authors who question the utility of the MBTI. Previously, Bjork and Druckman (1991) and Boyle (1995) reviewed the literature available to them and found no evidence supporting the utility of the MBTI. As Bjork and Druckman noted, "At this time, there is not sufficient, well-designed research to justify the use of the MBTI in career counseling programs. Much of the current evidence is based on inadequate methodologies" (p. 99).

Furthermore, there is a conspicuous lack of data demonstrating the incremental validity of the MBTI over other measures of personality. Such a finding is not surprising given the nascent revised interest in the role of personality measures in work settings. Nevertheless, given the attention that other personality perspectives receive from industrial-organizational psychologists (e.g., Big Five, McCrae & Costa, 1989; see also Hogan et al., 1996), it appears incumbent on counselors working in corporations to consider these alternatives as a part of their consulting work.

Equally important to recognize are the variables that a psychological test does not

measure. As Berr et al. (2000) noted, the correlations between conventional measures of personality and performance indicators are approximately $r = .20$. In other words, the shared variance between a personality dimension and work performance is 4%. Therefore, those who claim a link between personality and performance must acknowledge that the majority of work performance reflects many other influences.

Although the MBTI is an extremely popular test, the available evidence of its psychometric properties recommends caution regarding optimistic accounts of its ability to make empirically verifiable predictions. On many occasions, Messick (1965, 1980, 1982, 1989, 1995) called upon psychologists to recognize that test validity is an inference one derives from the scores and not an inherent property of the instrument. Specifically, a personality measure, such as the MBTI, is not inherently valid or invalid, only the interpretations made from the test results are. Moreover, Messick argued that inferences regarding the validity of a psychological measure must examine the potential consequences of using the instrument. Given the equivocal data related to the construction of the MBTI and the gravity of its application in work settings, I believe that counselors must consider with care claims made about the utility and value of the MBTI as a measure of personality and as a counseling tool.

The available evidence suggests that the MBTI does measure constructs related to personality; whether it measures the constructs identified by its underlying theory is not clear. Furthermore, it is not evident that the instrument can compartmentalize accurately, consistently, and unambiguously individuals' personality into the 16 type categories created by the instrument. Consequently, using the MBTI as a consulting tool in corporate settings may be, in some instances, the equivalent of making promises that one cannot keep. Such a conclu-

sion leaves us with the question, what is the utility of the MBTI as a consulting tool?

The popularity of the MBTI as a consulting tool most likely reflects the success of the publisher's marketing campaign and the intuitive and simple sounding nature of the instrument's scoring scheme. Compared to a conventional measure of the Big Five, it is probably comforting to learn that one tends to be intuitive and feeling, rather than learning that one has scored high on the neuroticism and low on the consciousness scales. Consequently, the MBTI can serve as a nonthreatening vehicle to introduce the concept of individual differences in personality and the relation between personally constructs and behavior to a general audience. The instrument might even serve as a catalyst for exercises that lead to improved esprit de corps among employees. However, presenting the data using the four-letter type formula rather than the scaled scores is a misrepresentation of the available evidence. The four-letter type formula may also be an overly simplistic account of complex personality dynamics and leave the recipient with a false impression that there is little left to doubt.

By contrast, other uses of the MBTI, especially those that affect an employee's job status, may have iatrogenic effects. Using the MBTI to select employees, to assign employees to work groups or assignments, or for other forms of employment evaluation are not justified for the simple reason that there are no available data to recommend such decisions. As such, counselors promoting applications of the MBTI should clearly represent the known limitations for specific uses of the instrument.

It is impossible to endorse applications of a personality measure without sufficient empirical evidence. Although McCaulley (2000) offered many claims regarding the conformation and utility of the MBTI in corporate settings, the available evidence does not support these predictions. Therefore, consulting psychologists require ded-

icated research that tests the specific predictions regarding the MBTI and other measures of personality.

An advantage of the MBTI is that its strong theoretical structure affords specific predictions regarding the link between personality and behavior. At the same time, there are other measures of personality whose structure portends useful applications in corporate settings. Therefore, it will be useful for future research to examine the predictive validity of these instruments relative to each other. Furthermore, it will be useful, as Hogan et al. (1996) noted, that analysis of the predictive power of personality tests incorporate other measures of performance, including intelligence, knowledge and experience, and skill.

The fact that the MBTI evolved within a strong theoretical structure can also be a liability so long as the structure of the theory, the scoring procedures, and the canons of score interpretation do not keep pace with the available evidence. That the MBTI shares variance with other measures of personality suggests that it may not be a unique measure of personality or that it measures unique personality dimensions. Consequently, those interested in using the MBTI should examine the advantages of replacing the four-letter type formula with more traditional magnitude assessments of personality.

References

- Berr, S. A., Church, A. H., & Waclawski, J. (2000). The right relationship is everything: Linking personality preferences to managerial behaviors. *Human Resource Development Quarterly, 11*, 133–157.
- Bess, T. L., & Harvey, R. J. (2002). Bimodal score distributions and the Myers-Briggs Type Indicator: Fact or artifact? *Journal of Personality Assessment, 78*, 176–186.
- Bjork, R. A., & Druckman, D. (1991). *In the mind's eye: Enhancing human performance*. Washington, DC: National Academy Press.
- Block, J. (1971). *Lives through time*. Berkley, CA: Bancroft Books.

- Block, J., & Ozer, D. J. (1982). Two types of psychologists: Remarks on the Mendelsohn, Weiss, and Feimer contribution. *Journal of Personality and Social Psychology*, *42*, 1171–1181.
- Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some psychometric limitations. *Australian Psychologist*, *30*, 71–74.
- Briggs, K. C., & Myers, I. B. (1998). *Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Capraro, R. M., & Capraro, M. M. (2002). Myers-Briggs Type Indicator score reliability across studies: A meta-analytic reliability generalization study. *Educational & Psychological Measurement*, *62*, 590–602.
- Carskadon, T. G. (1977). Test-retest reliabilities of continuous scores on the Myers-Briggs Type Indicator. *Psychological Reports*, *41*, 1011–1012.
- Carskadon, T. G. (1979a). Clinical and counseling aspects of the Myers-Briggs Type Indicator: A research review. *Research in Psychological Type*, *2*, 2–31.
- Carskadon, T. G. (1979b). Test-retest reliabilities of continuous scores on Form G of the Myers-Briggs Type Indicator. *Research in Psychological Type*, *2*, 83–84.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*, 249–253.
- Costa, P., & McCrae, R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Furnham, A. (1996). The Big Five versus the big four: The relationship between the Myers-Briggs Type Indicator (MBTI) and the NEO-PI face factor model of personality. *Personality and Individual Differences*, *21*, 303–307.
- Gardner, W. L., & Martinko, M. J. (1996). Using the Myers-Briggs Type Indicator to study managers: A literature review and research agenda. *Journal of Management*, *22*, 45–83.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.
- Goldberg, L. R. (1992). The development of markers for the Big Five factor structure. *Psychological Assessment*, *4*, 26–42.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*, 26–34.
- Harvey, R. J., & Murry, W. D. (1994). Scoring the Myers-Briggs Type Indicator: Empirical comparison of preference score versus latent-trait methods. *Journal of Personality Assessment*, *62*, 116–129.
- Hicks, L. E. (1984). Conceptual and empirical analysis of some assumptions of an explicit typological theory. *Journal of Personality and Social Psychology*, *46*, 1118–1131.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, *51*, 469–477.
- Howes, R. J., & Carskadon, T. G. (1979). Test-retest reliabilities of the Myers-Briggs Type Indicator as a function of mood changes. *Research in Psychological Type*, *2*, 67–72.
- Hunter, J. E., & Schmidt, F. L. (1990). Dichotomization of continuous variables: The implications for meta-analysis. *Journal of Applied Psychology*, *75*, 334–349.
- Johnson, D. A., & Saunders, D. R. (1990). Confirmatory factor analysis of the Myers-Briggs Type Indicator—Expanded analysis report. *Educational & Psychological Measurement*, *50*, 561–571.
- Jung, C. G. (1971). *Collected works of C. G. Jung: Vol. 6. Psychological types* (H. G. Baynes, Trans., revised by R. F. C. Hull). Princeton, NJ: Princeton University Press. (Original work published 1921)
- Kummerow, J. M. (1988). A methodology for verifying type: Research results. *Journal of Psychological Type*, *15*, 20–25.
- Lorr, M. (1991). An empirical evaluation of the MBTI typology. *Personality & Individual Differences*, *12*, 1141–1145.
- Macdaid, G. P., McCaulley, M. H., & Kaniz, R. I. (1986). *Myers-Briggs Type Indicator: Atlas of type tables*. Gainesville, FL: Center for Applications of Psychological Type.
- Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, *113*, 181–190.
- McCarley, N. G., & Carskadon, T. G. (1983). Test-retest reliabilities of scales and subscales of the Myers-Briggs Type Indicator and of criteria for clinical interpretive hypotheses involving them. *Research in Psychological Type*, *6*, 24–36.
- McCaulley, M. H. (2000). Myers-Briggs Type Indicator: A bridge between counseling and consulting. *Consulting Psychology Journal: Practice and Research*, *52*, 117–132.
- McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. *Journal of Personality*, *57*, 17–40.

- Mendoza-Denton, R., Ayduk, O., Mischel, W., Shoda, Y., & Testa, A. (2001). Person \times Situation interactionism in self-encoding (I Am . . .when . . .): Implications for affect regulation and social information processing. *Journal of Personality and Social Psychology, 80*, 533–544.
- Messick, S. (1965). Personality measurement and the ethics of assessment. *American Psychologist, 20*, 136–142.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35*, 1012–1027.
- Messick, S. (1982). Test validity and the ethics of assessment. *Diagnostica, 28*, 1–25.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.; pp. 13–103). New York: Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246–268.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologist Press.
- Pittenger, D. J. (1993). The utility of the Myers-Briggs Type Indicator. *Review of Educational Research, 63*, 467–488.
- Pittenger, D. J. (1994). The cross-disciplinary ethical responsibilities of psychology faculty. *Ethics & Behavior, 4*, 199–208.
- Saggino, A., Cooper, C., & Kline, P. (2001). A confirmatory factor analysis of the Myers-Briggs Type Indicator. *Personality & Individual Differences, 30*, 3–9.
- Saggino, A., & Kline, P. (1996). The location of the Myers-Briggs Type Indicator in personality factor space. *Personality & Individual Differences, 21*, 591–597.
- Salter, D. W., Evans, N. J., & Forney, D. S. (1997). Test-retest of the Myers-Briggs Type Indicator: An examination of dominant functioning. *Educational & Psychological Measurement, 57*, 590–597.
- Saunders, F. W. (1991). *Katherine and Isabel: Mother's light, daughter's journey*. Palo Alto, CA: Consulting Psychologists Press.
- Schuerger, J. M., Zarrella, K. L., & Holtz, A. S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56*, 777–783.
- Shoda, Y., & Mischel, W. (2000). Reconciling contextualism with the core assumptions of personality psychology. *European Journal of Personality, 14*, 407–428.
- Sipps, G. J., & Alexander, R. A. (1987). The multifactorial nature of extraversion/introversion in the Myers-Briggs Type Indicator and Eysenck Personality Inventory. *Educational & Psychological Measurement, 47*, 543–552.
- Sipps, G. J., Alexander, R. A., & Friedt, L. (1985). Item analysis of the Myers-Briggs Type Indicator. *Educational & Psychological Measurement, 45*, 789–796.
- Sipps, G. J., & DiCaudo, J. (1988). Convergent and discriminant validity of the Myers-Briggs Type Indicator as a measure of sociability and impulsivity. *Educational & Psychological Measurement, 48*, 445–451.
- Stricker, L. J., & Ross, J. (1962). *A description and evaluation of the Myers-Briggs Type Indicator* (Research Bulletin #RB-62-6). Princeton, NJ: Educational Testing Service.
- Tenopyr, M. L. (1988). Artifactual reliability of forced-choice scales. *Journal of Applied Psychology, 73*, 749–751.
- Thompson, B., & Borrello, G. M. (1986a). Construct validity of the Myers-Briggs Type Indicator. *Educational & Psychological Measurement, 46*, 745–752.
- Thompson, B., & Borrello, G. M. (1986b). Secondary-order factor structure of the MBTI: A construct validity assessment. *Measurement and Evaluation in Counseling Development, 18*, 148–153.
- Walck, C. L. (1992). The relationship between indicator and type and "true type": Slight preferences and the verification process. *Journal of Psychological Type, 23*, 17–21.