1999

Ethical Perspectives and Practice Behaviors Involving Computer-Based Test Interpretation

Mark R. McMinn
George Fox University, mmcminn@georgefox.edu

Brent M. Ellens

Erez Soref

Follow this and additional works at: http://digitalcommons.georgefox.edu/gscp_fac

Part of the Psychology Commons

Recommended Citation
McMinn, Mark R.; Ellens, Brent M.; and Soref, Erez, "Ethical Perspectives and Practice Behaviors Involving Computer-Based Test Interpretation" (1999). Faculty Publications - Grad School of Clinical Psychology. Paper 185.
http://digitalcommons.georgefox.edu/gscp_fac/185

This Article is brought to you for free and open access by the Graduate School of Clinical Psychology at Digital Commons @ George Fox University. It has been accepted for inclusion in Faculty Publications - Grad School of Clinical Psychology by an authorized administrator of Digital Commons @ George Fox University. For more information, please contact arolfc@georgefox.edu.
ETHICAL PERSPECTIVES AND PRACTICE BEHAVIORS INVOLVING COMPUTER-BASED TEST INTERPRETATION

Mark R. McMinn
Brent M. Ellens
Erez Soref
Wheaton College

The debates of the 1980s regarding responsible use of computer-based test interpretation (CBTI) software have mostly disappeared, as CBTI use has become common practice. We surveyed 364 members of the Society for Personality Assessment to determine how they use CBTI software in their work and their perspectives on the ethics of using CBTI in various ways. Psychologists commonly use CBTI software for test scoring and to provide a complementary source of input for case formulations. Most do not use CBTI software as the primary way to formulate a case, nor as an alternative to a written report. Controversy and uncertainty were expressed about importing sections of CBTI narratives into psychological reports. We distinguish between support and replacement functions of CBTI use, arguing that adequate research evidence should be present before using CBTI as a replacement for established assessment procedures.

Keywords: Computer-based test interpretation, personality assessment, professional ethics, psychological testing, report writing

Over a decade ago Matarazzo (1986) voiced concern about the proliferation of computer-based test interpretation (CBTI) software, noting that many such computer programs are “unvalidated plus all mean and no sigma” (p. 14). Matarazzo suggested that many CBTI products failed to consider the nuances that might result in a particular scale elevation or profile pattern, noting that in professional practice “variation is often the rule rather than the exception” (p. 20). He suggested that most CBTI software does not emulate the sophisticated clinical decision-making process that psychologists develop over years of training.

Matarazzo’s observations were part of a larger discussion within the professional psychology literature of the 1980s and early 1990s (Fowler, 1985; Garb, 1992; Matarazzo, 1983, 1986; Moreland, 1985; Rubenzer, 1991; Spielberger & Piotrowski, 1990). The American Psychological Association (APA) temporarily published cautionary Guidelines for Computer-Based Tests and Interpretations (APA, 1986)—a document that is no longer in print. Even optimistic proponents issued warnings about the importance of systematic research before uncritically accepting the merits of CBTI software (Butcher, Keller, & Bacon, 1985; Moreland, 1985).
Unfortunately, if this research was done at all, it was not reported in mainstream psychology journals. The *Journal of Personality Assessment* proposed a series of critical reviews on CBTI products (see Moreland, 1990), but few were forthcoming. Moreover, the critical discussions of the late 1980s and early 1990s seem to have given way to widespread use and approval of computerized assessment tools, including CBTI software.

Among survey respondents belonging to the Society for Personality Assessment and the Clinical Psychology Division of the APA, almost two-thirds indicate that they use computers to assist with psychological testing, and over one-third reported using CBTI software (Ball, Archer, & Imhof, 1994). In a recent survey regarding use of technology in psychological practice, McMinn, Buchanan, Ellens, and Ryan (in press) found surprisingly little controversy surrounding the use of computerized test administration, scoring, and interpretation software, with most respondents viewing these computer applications as generally ethical. In the midst of a decade bringing enormous economic pressures threatening the survival of psychological assessment, could it be that the scientific standards fueling productive discussions about CBTI in the 1980s have been overshadowed by pragmatic considerations such as efficiency and cost effectiveness?

Software has become dramatically more sophisticated in the past decade, and the number of CBTI vendors has declined, leaving mostly products marketed by reputable corporations with clear commitments to product research and a good understanding of psychological assessment standards. However, it seems important to establish and maintain ethical and scientific standards apart from the vendors of CBTI products. In this regard, we have made little progress (and perhaps have slipped backward) in the past decade.

To this end, we were interested in assessing psychologists’ current behaviors and ethical perspectives regarding CBTI. Collecting survey data is generally an effective way to determine current CBTI use among psychologists, but a difficulty with existing survey data is the lack of precision in determining how psychologists use CBTI software (McMinn, 1998). Previous survey questions about CBTI use have typically been embedded in general surveys about time requirements (Ball, Archer, & Imhof, 1994) or technology (McMinn et al., 1998), or have requested opinions regarding general statements about the value of CBTI (Spielberger & Piotrowski, 1990). The survey questionnaire reported here was intended to provide a more specific, focused evaluation of CBTI use among assessment psychologists.

**Method**

A three-part questionnaire was sent to 600 randomly selected members of the Society for Personality Assessment. Three waves of mailing were used: the questionnaire and a cover letter were sent in the initial mailing, a reminder postcard was sent 2 weeks later, and another questionnaire and cover letter were sent 2 weeks after the postcard.

Part I of the questionnaire requested basic demographic and practice information including age, gender, highest academic degree, years in practice, and the number of psychological assessments done per year. Part II posed the following scenario:

Assume that a clinical psychologist has been asked by an attending professional to evaluate a psychiatric inpatient. The psychologist has access to *Computer-Based Test Interpretation* (CBTI) software, and has also been trained in using *standard clinical methods* for test interpretation (e.g., actuarial methods, consultation, published case studies, and clinical judgment based on prior clinical experience). Please rate each of the following actions on the two scales: Ethics and Personal Use.

A list of six potential uses for CBTI followed, ranging from using CBTI for scoring purposes only, to using CBTI as an adjunct to standard clinical interpretation methods, to using CBTI as the exclusive basis for a psychological report (items are listed in Table 1). Respondents rated each scenario on a 5-point scale on the extent to which the psychologist’s behavior was ethical, and on another 5-point scale as to the frequency with which they engage in similar uses of CBTI. This scale format is similar to that used by Pope,
<table>
<thead>
<tr>
<th>Item</th>
<th>Rating</th>
<th>Occurrence in your practice?</th>
<th>Ethical?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The psychologist uses CBTI software rather than hand scoring</td>
<td></td>
<td>16  4 25 25 31</td>
<td>3 1 5 11 80</td>
</tr>
<tr>
<td>2. The psychologist first uses standard clinical methods for the</td>
<td></td>
<td>17  8 23 28 25</td>
<td>2 2 3 17 75</td>
</tr>
<tr>
<td>case formulation, and then uses the interpretative printouts from</td>
<td></td>
<td>54 17 18 8 3</td>
<td>17 25 19 25 13</td>
</tr>
<tr>
<td>the CBTI software as a source of additional information (i.e., the CBTI narrative provides a second opinion).</td>
<td></td>
<td>87 10 2 1 0</td>
<td>53 33 9 4 1</td>
</tr>
<tr>
<td>3. The psychologist writes a report using “cut and paste” technology,</td>
<td></td>
<td>94  4 2 0 0</td>
<td>79 12 6 1 2</td>
</tr>
<tr>
<td>thereby including portions of the CBTI narrative in the report.</td>
<td></td>
<td>17  8 23 26 26</td>
<td>1  3 4 28 63</td>
</tr>
<tr>
<td>4. The psychologist uses the CBTI interpretive report(s) as the</td>
<td></td>
<td>16  4 25 25 31</td>
<td>3 1 5 11 80</td>
</tr>
<tr>
<td>primary resource for the case formulation (i.e., standard clinical</td>
<td></td>
<td>17  8 23 28 25</td>
<td>2 2 3 17 75</td>
</tr>
<tr>
<td>methods have little or no effect on the formulation).</td>
<td></td>
<td>54 17 18 8 3</td>
<td>17 25 19 25 13</td>
</tr>
<tr>
<td>5. The psychologist provides the referring professional with the</td>
<td></td>
<td>87 10 2 1 0</td>
<td>53 33 9 4 1</td>
</tr>
<tr>
<td>CBTI interpretive report(s) in lieu of writing a report.</td>
<td></td>
<td>94  4 2 0 0</td>
<td>79 12 6 1 2</td>
</tr>
<tr>
<td>6. The psychologist considers the CBTI interpretive report(s) in</td>
<td></td>
<td>17  8 23 26 26</td>
<td>1  3 4 28 63</td>
</tr>
<tr>
<td>tandem with standard clinical methods and arrives at a case</td>
<td></td>
<td>16  4 25 25 31</td>
<td>3 1 5 11 80</td>
</tr>
<tr>
<td>formulation (i.e., the CBTI narrative and standard clinical methods</td>
<td></td>
<td>17  8 23 28 25</td>
<td>2 2 3 17 75</td>
</tr>
<tr>
<td>are both important in the initial formulation).</td>
<td></td>
<td>54 17 18 8 3</td>
<td>17 25 19 25 13</td>
</tr>
</tbody>
</table>

Notes. Rating codes: Occurrence in your practice? 1 = never, 2 = rarely, 3 = sometimes, 4 = fairly often, 5 = very often; Ethical? 1 = unquestionably not, 2 = under rare circumstances, 3 = don't know/not sure, 4 = under many circumstances, 5 = unquestionably yes. CBTI = Computer-based test interpretation. The number of responses for Items 1 through 6 do not always total 100% because of rounding.


Results

Of the 600 questionnaires sent, 364 respondents returned the questionnaire, resulting in a response rate of 63.4% (after excluding the undeliverable surveys). Of the 364 respondents, 244 were male (67%), 115 were female (31.6%), and 5 (1.4%) did not report their gender. Approximately 86% of the respondents were between the ages of 30 and 60, while approximately 14% were over 60 years. Nearly 93% of respondents reported their ethnicity as European-American, while 0.6% reported as African-American, 1.5% as Asian-American, 1.2% as Latino, and 0.6% as Other. The mean number of assessments per year was 81.6, and the mean years in practice was 17.5. Most (87.6%) respondents reported holding a PhD, and others (12.1%) reported holding a PsyD. The vast majority (96.1%) were licensed as psychologists, and 15.4% of respondents reported being ABPP Diplomates.
The numbers of psychological tests administered and CBTI uses per month are summarized in Table 2. Distributions on each of the tests were positively skewed, with a few psychologists administering many more tests than average.

Response patterns on the six CBTI scenarios are reported in Table 1. Three of the six behaviors were generally deemed ethical (more than 80% rated the behavior as a 4 or 5 on the 5-point ethics scale): using CBTI for scoring purposes, using CBTI printouts as a supplement to standard clinical methods, and using CBTI printouts in tandem with standard clinical methods. Two behaviors were generally deemed unethical (more than 80% rated the behavior as a 1 or 2 on the 5-point ethics scale): using CBTI printouts as the primary resource for case formulation, and providing the CBTI printout to the referring professional in lieu of writing a report. One behavior received equivocal ethics ratings (more than 20% rated the behavior as a 1 or 2 and more than 20% rated the behavior as a 4 or 5 on the 5-point ethics scale): using “cut and paste” technology to include portions of a CBTI printout in a psychological report.

We also looked for group differences on each of the ethics and practice ratings, using a conservative alpha of .005 to control for the inflated risk of Type I error with multiple hypothesis tests. Men were more likely than women to “cut and paste” from CBTI reports to their own clinical reports, $t(343) = 3.9, p < .001$, and to view this as ethical, $t(338) = 4.1, p < .001$. Men were also more likely to score psychological tests with CBTI software than women, $t(341) = 2.9, p = .005$, and younger psychologists were more likely than older psychologists to use CBTI software for scoring, $t(341) = 3.0, p < .005$. No differences were observed based on the number of psychological assessments done per year.

**Discussion**

**Test Use**

Among the tests listed on the survey, the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) is administered most often and is the most common subject of CBTI use. This is not surprising considering that the MMPI-2 is among the most commonly used personality test used by clinical psychologists (Watkins, Campbell, Nieberding, & Hallmark, 1995), and that the Microtest Q CBTI software is widely available through National Computer Systems.

**Behaviors Deemed Appropriate**

Psychologists appear to be comfortable using CBTI for scoring purposes, with 85% having done so, at least rarely, and 91% reporting it to be ethical (score of 4 or 5 on the 5-point scale). CBTI scoring can save time (Alexander & Davidoff, 1990) and reduce errors. When Allard, Butler, Faust, and Shea (1995) had 8 trained technicians and professionals hand score 43 protocols from the Personality Diagnostic Questionnaire-Revised (PDQ-R; Hyler, Skodol, 1995),

<table>
<thead>
<tr>
<th>Test</th>
<th>Tests administered in an average month</th>
<th>CBTI uses in an average month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>MMPI-2</td>
<td>4.8</td>
<td>16.0</td>
</tr>
<tr>
<td>MCMI-III</td>
<td>1.1</td>
<td>3.5</td>
</tr>
<tr>
<td>SCL-90-R</td>
<td>0.7</td>
<td>3.4</td>
</tr>
<tr>
<td>Rorschach</td>
<td>3.3</td>
<td>5.6</td>
</tr>
<tr>
<td>Wechsler intelligence tests</td>
<td>3.8</td>
<td>8.1</td>
</tr>
<tr>
<td>Other</td>
<td>9.4</td>
<td>32.3</td>
</tr>
</tbody>
</table>

Note. CBTI = Computer-based test interpretation; MMPI-2 = Minnesota Multiphasic Personality Inventory-2; MCMI-III = Millon Clinical Multiaxial Inventory-III (Millon, 1994); SCL-90-R = Symptom Checklist-90-R (Derogatis, 1994).
Oldham, Kellman, & Doidge, 1992), they found 53% of the resulting profiles contained errors and 19% contained errors significant enough to affect clinical diagnosis. Assuming correct data entry and proper software development, the error rate of computerized scoring is negligible.

Psychologists also reported comfort with using CBTI interpretive printouts in tandem with standard clinical methods in formulating an assessment case. This was true whether the CBTI printout was considered a secondary source (Item #2) or an equal source of information (Item #6). Test interpretation software provides psychologists access to the interpretive expertise of test developers. Such software has been labeled “expert system” because it allows the computer to serve as an expert consultant to the psychologist. For example, Dr. John Exner has been highly involved in developing both the Comprehensive System for Rorschach scoring (Exner, 1993) and in the Rorschach Interpretative Assistance Program (RIAP; Exner & Ona, 1995) software that helps clinicians interpret Rorschach results. Similarly, Dr. Theodore Millon has been involved in the development of the various Millon tests such as the Millon Clinical Multiaxial Inventory (Millon, 1983) and the Millon Adolescent Personality Inventory (Millon, Green, & Meagher, 1982) and in the development of the Microtest Q interpretive software for these tests. Automated interpretive systems such as these magnify the availability of experts for each test.

However, even when CBTI software is developed by experts the question must be raised as to whether an expert on a particular test can, by using algorithms and interpretive narratives, provide expert opinion on a specific patient whom he or she has never known. Accordingly, psychologists are not comfortable with indiscriminate use of CBTI products.

**Behaviors Deemed Inappropriate**

Most psychologists reported never using CBTI as the primary resource for case formulation, and 86% believe it would be unethical to do so (score of 1 or 2 on the 5-point scale). Likewise, psychologists do not provide CBTI printouts to referring professionals in lieu of writing their own report, nor do they see this as ethical. Thus, it appears that psychologists are comfortable using CBTI information in tandem with standard clinical methods of interpretation, but do not see CBTI as a replacement for these standard methods. Perhaps they recognize the dangers that opponents of CBTI voiced over a decade ago—even an expert system cannot consider all the human variation and clinical nuances that are evident to an experienced psychologist who has had personal interaction with the patient.

Expert systems are based on interpretive approaches that have limited accuracy and utility (Ehrenworth & Archer, 1985). Given the associations of objectivity and accuracy that many individuals have when considering computer software, output from these expert systems might be misperceived as reports based on pure actuarial data. As Butcher et al. (1985) note, “there are no purely actuarial automated interpretive reports, partly because available actuarial systems leave a large percentage of examinees unclassified” (p. 807). Expert systems provide interpretative output based on algorithms—mathematical procedures that simplify human variation into a finite set of decision rules. Here we return to Matarazzo’s (1986) warning that “variation is often the rule rather than the exception” (p. 20). Because CBTI products do not adequately anticipate all possible human variation, they are inaccurate or misleading at times—what Matarazzo decried as “all mean and no sigma” (p. 14). When expert systems with their inherent limitations diffuse the sense of responsibility experienced by the local expert (i.e., the psychologist doing the evaluation) they do potential damage to the veracity of the assessment itself.

**Equivocal Behavior**

One item emerged as equivocal: the psychologist using “cut and paste” technology, thereby including actual statements from the CBTI narrative in his or her written report. For this item, 42% rated the behavior as generally unethical (rating of 1 or 2) whereas 38% rated it as generally ethical (rating of 4 or 5). An additional 19% reported not knowing if this is ethical.

With the technological advances of recent years, it is often possible to obtain CBTI reports electronically. Sentences, paragraphs, or entire sections can
be electronically copied into a word processing file and used as part of a psychologist’s clinical report. Though most psychologists do not do this routinely, there is disagreement about the ethics involved. If certain paragraphs are copied, should they be clearly identified as coming from a CBTI report, or can they be used as if they were the psychologist’s own words? If these paragraphs are identified as coming from a CBTI report, can the psychologist edit them? Are plagiarism laws and standards applicable in these circumstances? Does the software developer share legal liability for the report? Psychologists apparently have varying opinions and a high rate of uncertainty about matters such as these.

Group Differences
The finding that men are more likely than women to “cut and paste” from CBTI reports to their own clinical reports may be partly related to the observation that men use computers at a higher rate than women. This is suggested by the observation that men are more likely to score psychological tests with CBTI software than women are despite the fact that most respondents, regardless of their gender, viewed CBTI scoring as ethical.

Younger psychologists are more likely than older psychologists to use CBTI software for scoring, perhaps because younger psychologists have generally received more training in computer technologies and have been exposed to CBTI as part of their graduate education.

Conclusion
It is helpful to distinguish between the support functions of CBTI and the replacement functions of CBTI. By support functions, we refer to using CBTI as an adjunctive tool along with standard methods of psychological assessment. Assuming adequate training of the psychologist, using CBTI for support functions seems a relatively safe practice—an opinion shared by the vast majority of our respondents. For example, if a psychologist who is well grounded in research regarding MMPI-2 profile interpretation first examines a profile and then seeks a “second opinion” through CBTI, this may help the psychologist critically evaluate his or her interpretation. Our respondents were generally comfortable using CBTI as a secondary source of information or in tandem with standard clinical methods.

By replacement functions, we refer to using CBTI to eliminate some analogous aspect of the assessment process. It would seem that replacement functions lead to more dangerous applications of CBTI, calling for an adequate research base before considering such functions ethical. One example of a replacement function that is deemed ethical by our psychologist respondents—and one with an emerging research base to support its use (Allard et al., 1995)—is using CBTI to score psychological tests. This is not only more efficient for many objective personality tests, it is also more accurate. Indeed, if research continues to show improved scoring accuracy with CBTI software, it might eventually be deemed unethical not to use computerized scoring for personality tests. Of greater concern is using CBTI to replace the interpretive or reporting work of assessment psychologists. Consistent with our views and the relative dearth of research evidence to support its use, survey respondents reported ethical danger in using CBTI as the primary resource for case formulation or using a CBTI report in lieu of a psychologist’s report.

It is not entirely clear whether using “cut and paste” technology to transfer narrative from a CBTI report directly to a psychological assessment report is primarily a support function or a replacement function. Survey respondents disagreed about the ethics of this, yet almost 30% reported engaging in this “cut and paste” behavior sometimes, fairly often, or very often. One could argue that this is a support function, especially if small portions of the narrative are used to support and add credibility to a formulation articulated in the psychologist’s own words. One could also argue that this is a replacement function, intended to make report writing more efficient by using computer-generated narratives as an alternative to the painstaking process of carefully writing a report. In this latter case especially, we argue that there is not adequate research support to justify this practice.

This distinction between support and replacement functions of CBTI leads to broader questions
pertaining to professional training and practice. As computerized methods become more common in psychological training, might it be tempting to train psychologists in CBTI replacement methods while minimizing training in the traditional research and interpretation methods that have formed the basis of assessment psychology over the course of this century? Is it appropriate for assessment psychologists to obtain informed consent from patients regarding their interpretive assistance programs? These questions illustrate the need for pursuing clear guidelines, standards, and review procedures for CBTI products and practices. Much as the American Psychological Association (APA) has led the way in establishing guidelines for test development, it seems fitting to look to national organizations of psychologists to provide guidance for CBTI use.

References