The Psychological Report: A Review of Current Controversies

Gary Groth-Marnat
Private Practice and Pacifica Graduate Institute

Leah Suzanne Horvath
New York University Counseling Services

A summary of the most frequent controversies in report writing is covered. These include length, readability, acknowledging use of poorly validated measures, use of computer-based narratives, inclusion of test scores, degree of integration, inclusion of client strengths, and development of a feedback report. Available research is summarized along with suggestions for future research. © 2005 Wiley Periodicals, Inc. J Clin Psychol 62: 73–81, 2006.

Keywords: psychological report; controversies

The psychological report is the end product of the assessment process. As such, it combines a wide number of factors. These include interpreting the referral question, test selection, case conceptualization, theoretical orientation, processes involved in the interview, interpretation of test scores, integrating various sources of information, prioritizing what is most important to include, and the selection of the best means of presenting the information. It is thus not surprising that psychologists have often come to different conclusions about how best to construct a report. The following discussion will summarize and elaborate on issues and controversies in the psychological report.

Length

Although the average psychological report is between five and seven single-spaced pages (Donders, 2001) the optimal recommended length varies. Some clinicians tout the benefits of a brief report—cutting quickly to the core of the assessment findings and allowing for
efficient access to their meaning and implications for treatment (Stout & Cook, 1999). Other practitioners advocate for a more comprehensive report, with descriptions of client history, behavioral observations, detailed descriptions of the results and a thorough summary and recommendations section. In practice, there appears to be much variability in how long reports are—with one study of custody evaluations noting that psychological reports ranged from 1 to 54 pages (Horvath, Logan, Walker, & Juhasz, 2000).

Universal agreement on length may not be possible as the appropriate length of the report may depend on the setting for the assessment, as well as the referral source, referral question, and intended consumer. Within a forensic context Ackerman (this issue, 2006, pp. 59–72) distinguishes between brief (1–3 page), standard (2–10 page), and comprehensive (10–50 page) reports. The shortest reports generally come from medical settings where rapid feedback, ease of team members accessing information, and the model set by medical reports all encourage quite short formats. For example, Stout and Cook (1999) report that assessments conducted for medical professionals are preferably one page, specific, use bullets to highlight major points, and focus on symptoms, diagnosis, and treatment applications. On the other hand, psychological reports in forensic settings need to be considerably longer because they need to include much more detailed client histories, descriptions of strengths and weaknesses, review additional records, and include a thorough summary and recommendations section. A one-page bulleted report would be inappropriate for this type of setting.

The above considerations indicate practitioners should be aware of the rationale for various report lengths. They should also develop the skills and flexibility necessary for evaluating when various lengths are required. This may represent the greatest challenge for brief reports where extracting only the most essential information is warranted. Finally, negotiating with the referral source on the relative length of the report should be a routine practice.

Readability

It has also been noted that psychologists need to improve the readability of their reports (Brenner, 2003; Harvey, 1997, this issue, 2006, pp. 5–18). Psychological reports are consumed not only by other mental health professionals, but also by clients, parents, court systems, and school personnel. Reports need to be written at a level commensurate with the audience. Overuse of psychological jargon was the most frequent complaint found in past and present research on consumer satisfaction (Brenner, 2003; Harvey, 1997, this issue, 2006, pp. 5–18; Tallent & Reiss, 1959), and this complaint was true for consumers who were mental health professionals as well as parents, clients, and teachers. Although practitioners may believe that their reports are written at an appropriate reading level, it is likely that they are in fact too advanced for most consumers.

Harvey (this issue, 2006, pp. 5–18) determined that the reasons for this are most likely due to training that emphasized test scores rather than the client’s everyday experience and overall context. Additional reasons included being given model reports that were too technical, incorrectly assuming that a wide range of people understood jargon, lack of consensus regarding technical terminology, greater time efficiency writing reports that were more difficult to understand, and confusion regarding the intended audience. Harvey offers suggestions for ways practitioners can make their reports more readable, including shortening sentences, minimizing difficult words, reducing jargon and acronyms, omitting passive verbs, increasing the use of subheadings, using everyday descriptions of client behavior, using time-efficient strategies, and collaborating with the client. In addition, treatment recommendations should be as specific and concrete as possible
and be phrased in such a way as to convince recipients that the situation is serious, that treatments will be effective, that the benefits will outweigh the inconvenience, and that they have some control over treatment decisions (Meichenbaum & Turk, 1987).

Acknowledging Presence of Poorly Validated Measures

Psychological tests used in clinical assessment have varying degrees of validity, and research has shown that the validity of psychological tests is similar to many medical tests (Meyer et al., 2001). The question that has been debated in the psychological assessment literature is whether or not to include measures in an assessment battery that have not been validated or that have some questionable validity (Michaels, this issue, 2006, pp. 47–58). There are many reasons why a practitioner chooses particular assessment instruments over others, and these reasons have been debated elsewhere (Garb, 1998; Hunsley, 2002; Silver, 2001). If poorly validated measures have been used, it has been argued that practitioners should note that the measures do have questionable validity. In particular, interpretations that are based on data from poorly validated measures should be described in the report with some caution or with a disclaimer that includes the information on how test validity may have affected the results. When psychologists describe their measures in the report, an explanation for the use of these measures can help support the use of measures with limited validity.

One difficulty with the above argument is that it assumes there is a one-to-one correspondence between a test score and an interpretation. When practitioners write an integrated report that utilizes a hypothesis testing approach, they base their interpretations on the convergence of multiple sources of information (see Beutler & Groth-Marnat, 2003). Simply listing the test and resulting interpretation would fail to take this integrative process into account. However, the degree to which practitioners integrate and hypothesis test varies greatly. There are some reports that have a minimal amount of integration and read more like “test results” (usually organized according to individual tests rather than relevant domains of the client’s functioning). In these cases, it may be appropriate for practitioners to comment on the relative validity or the degree of certainty that can be placed on the interpretations. Such statements become less relevant for practitioners writing integrated reports.

Use of Computer-Based Narratives

Computer-based test interpretation (CBTI) is a frequently debated area (Butcher, Perry, & Atlis, 2000; Butcher, Perry, & Hahn, 2004; Groth-Marnat & Schumaker, 1989). Computer-based test interpretations provide information on the client’s symptoms, interpersonal and intrapersonal styles, and likely responses to treatment, along with information on critical items (i.e., suicide and homicide risk) that the individual endorsed. However, a central issue is the accuracy of these interpretations. A computer does not have access to the client’s relevant history and does not observe how the client presented for testing, or how the client responded to the test environment (Butcher et al., 2000). The use of rigorous clinical judgment is something that cannot be programmed into a computer (Butcher et al., 2004). Some researchers have expressed concerns about the validity of these interpretations. Perhaps even more concerning is the seeming sophistication of the lengthy CBTI narratives, which make these interpretations appear valid to untrained consumers and even to some professionals (Butcher et al., 2000; Groth-Marnat & Schumaker, 1989; Snyder, 2000). Clinicians must not assume these data are accurate and appropriate for their client (Butcher et al., 2000, 2004). Indeed, it is estimated that half of
the interpretations made using CBTIs are false (Butcher et al., 2004). This means that if a clinician unquestioningly incorporates these interpretations into a report, half of the interpretations will be erroneous. Given these concerns and perhaps in anticipation of some of them, the American Psychological Association (APA) adopted interim standards on computerized test scoring and interpretation in 1966 (Fowler, 1985) and developed ethical guidelines for their use in 1986 (APA, 1986).

The greatest controversy in report writing comes when clinicians unquestionably accept the computer-based interpretations. The extreme example of this is when report “writers” simply cut and paste large segments from CBTI narratives and include this material in their reports. Certainly ethical issues are raised by this procedure in that it means that inaccurate interpretations will be presented to the referral source as if they were accurate (see Michaels, this issue, 2006, pp. 47–58). Survey data by McMinn, Ellens, and Soref (1999) indicated that 42% of psychologists rated this as generally unethical whereas 19% stated they did not know and 38% rated it as generally ethical (also McMinn, Buchanan, Ellens, & Ryan, 1999).

The above information indicates that psychologists need to evaluate each statement in CBTI’s within the context of the client’s other data and decide which statements are valid and applicable and which are not (Cates, 1999). This is a difficult task, as poor integration of CBTI and clinician data can result in reduced validity of the report (Snyder, 2000). Lichtenberger (this issue, 2006, pp. 19–32) provides strategies for enhancing clinician–computer interpretations including integrating other sources of data, exploring contradictory hypotheses, use of decision trees, use of actuarial formulas (when present), and being aware of and challenging clinician-based sources of error. Ultimately, the psychologist writing the report has responsibility for the content of the report regardless of how the data and interpretations were generated.

### Inclusion of Test Scores

There is considerable controversy on whether or not to include test scores in psychological reports (see Ackerman, this issue, 2006, pp. 59–72; Michaels, this issue, 2006, pp. 47–58). Those in favor of including test data note that it is helpful for other psychologists who are reading the report to have access to these scores (Freides, 1993; Matarazzo, 1995). Psychologists may have different ways of interpreting test scores and allowing other professionals access to test scores allows each the opportunity to better understand the client in terms of their own interpretation of the results. The actual scores will indicate both supporting as well as nonsupporting information related to the inferences made in the report. This means that the psychologist is more likely to be held accountable for his or her conclusions. A further reason is, within forensic settings, it reduces the need for discovery because the most relevant information has already been provided. Actual data also gives the report a more precise, scientific feel to it. Another important reason for including scores is that it enables follow-up ipsative comparisons of changes in a client’s performance. This means that improvement or deterioration in a client’s functioning can be noted upon retesting. Finally, including scores in an appendix means that psychologists may have more freedom to focus on an integration of findings and qualitative aspects of the client in the body of the report itself.

Those opposed to the inclusion of test scores do not discount that, under some circumstances, the data may be useful to other professionals. However, they argue for keeping careful control over who gets access to this data. Keeping data between qualified professionals is also consistent with previous APA ethical guidelines (see Naugle & McSweeney, 1996). A counter argument to this is that the scores are meaningless to those not
trained in assessment and thus would not be harmful to the client and would not violate test integrity. However, though some test data may be meaningless to nonpsychologists (such as scores on some neuropsychological tests or Rorschach test summaries), other scores of more common assessments (such as IQ tests) may, in fact, be meaningful to the general public. Not only might the scores be meaningful but also they are likely to be subject to frequent misinterpretation. In those cases, it is argued that it is inappropriate to include data and report specific scores to the client or others such as schools or family members.

A further reason not to include scores routinely is that even qualified persons may misinterpret their meanings. This is because the simple score does not take into account such things as behavioral observations, cultural factors, level of motivation, malingering, testing the limits, and the presence of fatigue. Often the underlying reasons for a score are not apparent. For example, a low score on Digit Symbol-Coding can occur for a variety of reasons. It is the clinician’s job to sort through a variety of possibilities and conclude, for example, that the low Digit Symbol-Coding score was due to slow processing speed (vs. memory impairment, poor sequencing, malingering, or a perfectionistic problem solving style).

Ethical considerations relate to finding a balance between a client’s right to and control over information and the need to avoid harm (see Michaels, this issue, 2006, pp. 47–58). The most recent APA ethical guidelines (APA, 2002) have given more control to the client in deciding where and to whom their records (including test data) should be released. This now means that they could give authorization for release of test data to unqualified persons if they chose to do so. This presents the possibility of an ethical dilemma if there was the potential for harm to the client. A further issue is that “test data” may extend to actual test items that may appear on clinician notes, questionnaires themselves, or computerized reports where critical items might be listed. Sometimes such items are included on a psychological report as a means of providing qualitative information. This may result in a dilemma between client rights to information versus the psychologist’s ethical and legal responsibilities to test security and copyright laws.

Research within a neuropsychology context ($N = 137$) found that the majority of clinicians (63%) do not routinely append test scores (Pieniadz & Kelland, 2001). The most frequent reason given was to protect the data from misunderstanding or misuse by unqualified persons (83%). Other reasons were that they were more concerned with patient processes (qualitative information) rather than the end score (46%) to protect even qualified professionals from misinterpreting the data because they did not have access to potentially mediating factors (46%), and to minimize possible intrusions into the patient’s privacy (25%). The two most important reasons for the subgroup who routinely append scores (35% of sample) were to be thorough (100%) and to facilitate comparisons with past or future testing (95%). They also indicated that appending scores would help within legal contexts (43%), allowed them to focus on qualitative information thereby enhancing efficiency (62%), and on rare occasions, scores were required by the referral source (2%).

The above indicates that the decision to append test scores can be based on the following factors. First, in some cases, the referral source may expect actual scores. If these scores are released to a properly qualified person, then it would be appropriate to append them to the report. A good example would be if a referring psychologist needs the scores to compare the client’s performance with a past or future assessment using the same tests. In contrast, if they might be misused or misinterpreted by the referral source, the scores should not be released. Second, a clinician’s theoretical orientation may impact their decisions to append scores. More qualitative, process-oriented clinicians are prob-
ably less likely to append scores than persons oriented to a quantitative, fixed battery approach. Third, whether or not to append scores may vary across instruments. IQ scores may be subject to misinterpretation by nonprofessionals and thus should probably be guarded more carefully. In contrast, summary scores for a Rorschach test are likely to be sufficiently opaque that they would only be meaningful to a properly qualified professional. Finally, test scores need to be released when specifically requested by the client. However, ethical and legal considerations need to be made related to client rights, potential harm, test security, and copyright laws.

Integration of Reports

The extent that reports integrate different sources of information varies greatly. Reports low on integration generally present the conclusions test by test and rely heavily on test scores. They often neglect contradictions between various scores and ignore the overall context of the client. They usually spend minimal time discussing the meaning of the scores for the client. Typical phrases that suggest a poorly integrated report include “. . . people with these profiles are often described as . . .,” “scores indicate that . . .,” or “Mr. Smith had an elevated score on scale 3 which indicates . . .”. In contrast, integrated reports involve the combination of data from multiple assessment methods, including records, interviews, observations, and multiple psychological tests. An integrated report is not merely the presentation of findings from each of these unique assessment tools, but the blending of these findings into a meaningful understanding of the client in the context of the client’s life (Beutler & Groth-Marnat, 2003; Lewak & Hogan, 2003). Multiple measures are used in the integrated report to obtain new information or to better explain findings which may be only weakly supported or vague without the inclusion of additional techniques (Cates, 1999). Integrated reports rely on good clinical judgment skills to make sense of contradictory findings and best understand results in terms of the assessment context, the client’s background, and the referral question.

Inclusion of Client Strengths

Psychological reports have sometimes been criticized for focusing too much on what is wrong with a client and ignoring a client’s strengths. This probably stems from the medical model, which focuses only on identification of the problem and recommendations for how to solve the problem. Snyder, Ritschel, Rand, and Berg (this issue, 2006, pp. 33–46) argue that the psychological report should not only focus on client difficulties, but also elaborate on the strength and hope that exist in each individual. Focusing only on negatives will limit the psychologist’s view of the client with the risk of overpathologizing the individual. Additionally, a deficits only psychological report fails to recognize relevant client strengths that could be important ingredients in enabling the client to overcome deficits. Finally, clients have progressively more access to their reports such that the impact of a client reading it needs to be taken into consideration. Clients reading reports that focus almost exclusively on their deficits are likely to be demoralized and more likely to feel alienated from the practitioners they are working with and possibly the mental health care system in general. This also raises ethical issues related to avoiding harm to the client.

In contrast, some clinicians argue that a report should not include information other than that requested by the referral source. This is consistent with the fact that, in actual practice, few referral sources actually request information regarding a client’s strengths. This means that if a report writer decides to include client strengths, they typically do so
without having been requested to do so by the referral source. The result may be a report that is too long, does not focus sufficiently on the “referral question,” and, as a result, may not be as favorably received by the referral source. An intermediate position to this controversy is, at a minimum, to routinely ask referral sources if they would like information on client strengths, ask clients about their strengths during interview, and include a subheading on relevant client strengths.

Feedback Report

Feedback in the assessment process can also be a source of some disagreement among clinicians—perhaps not so much in theory as in practice. In the fast-paced environment of an inpatient psychiatric unit for example, psychologists may find that there is little time available for feedback sessions with patients, though undoubtedly most practitioners would agree that feedback is an important component of the assessment process. Research has found that clients who received test feedback were considerably more satisfied with the assessment than those who did not receive feedback (Finn & Tonsager, 1992, 1997), and clients who received feedback also showed a significant decline in self-reported distress compared to clients who did not receive feedback (Finn & Tonsager, 1992).

Finn (2003) described having a discussion session about the assessment results with the client and referring clinician and sending written feedback to the client and the clinician. Providing a feedback report is an excellent addition to the psychological report—offering a hard copy summary of the test results in language that is clear and with examples from the client’s life that are applicable to the data (Pope, 1992). Such a feedback report can serve as a reminder of the key components of the assessment findings and can address the client’s specific assessment questions if applicable (Lewak & Hogan, 2003). Finn and Tonsager (1997) report that clients derive the greatest therapeutic benefit when feedback is ordered according to how well information matches with the client’s own view of him or herself. The most congruent information should be presented first and then feedback should proceed with gradually less-congruent information. Similarly, a feedback report that summarizes the test feedback in this manner would further ensure that the client has a solid understanding of the test results. Offering a feedback report in addition to, or in lieu of, a formal report may provide client consumers with a more user friendly and useful summary of their assessment experience.

Summary and Conclusions

The above controversies were selected because they highlight variations in report writing style, have proponents on different sides of the issues, and some of the topics were partially covered in separate articles in this special series. However, research is generally lacking regarding many of these areas. The following is a listing of what seems to be the most salient research topics organized in the order the controversial topics were presented:

- How do consumers rate report length?
- What can clients themselves tell us about the psychological report?
- How do psychologists actually use computer-based narratives?
- What are the patterns of inclusion of test scores by clinical psychologists?
- How do clinical psychologists decide to include or not to include test scores?
- To what extent are reports integrated?
- How do consumers rate integrated versus nonintegrated reports?
• How do referral sources rate the inclusion of client strengths?
• What is the impact on clients and client treatment of including client strengths?
• What is the impact on clients of receiving a feedback report?

To summarize this review, the most useful and ethical psychological reports take into consideration the needs of the consumer (Cates, 1999). This means that the format of the report, the choice of language, the inclusion of test data, and the use of feedback reports should all be decided after careful consideration of the consumer’s needs. Reports should also take into consideration the client’s cultural context, ensuring that the report is structured around the client’s culture. They should also provide a balance between positive and pathological aspects in the client’s psychological functioning. Finally, computer-based interpretations and nonvalidated assessment measures should be included only when used by those experienced with these devices and only with careful consideration of their accuracy. The psychological report will likely continue to engender debate in the field, but with proper use it can also continue to be a unique and invaluable contribution to the field of mental health.

References