

CANADIAN
PSYCHOLOGICAL
ASSOCIATION



SOCIÉTÉ
CANADIENNE
DE PSYCHOLOGIE

**Recommendations by the Canadian Psychological
Association for Improving the North American Safeguards
that Help Protect the Public
Against Test Misuse**

Marvin L. Simner, Ph.D.¹
Department of Psychology
University of Western Ontario
London, Ontario N6A 5C2
Canada

Approved by the Professional Affairs Committee
Working Group on Test Publishing Industry Safeguards,
September, 1994

Approved by the Board of Directors,
Canadian Psychological Association,
November, 1994

¹ Chair, Canadian Psychological Association Professional Affairs Committee Working Group
on Test Publishing Industry Safeguards

Background

In June 1992 the Professional Affairs Committee of the Canadian Psychological Association (CPA) formed a Working Group of representatives from several sections of the Association as well as from the Canadian test distributing industry. The purpose of this group was to examine the effectiveness of the two major safeguards that the North American test industry normally employs to help protect the public against test misuse.

The need for such an examination stemmed from a concern raised in a letter received by CPA over one of the safeguards, a three-level test classification system. This safeguard, which was developed more than 40 years ago by the American Psychological Association (APA, 1950), is a system that categorizes psychometric tests in terms of the professional qualifications or training needs that test users must meet in order to employ these tests in an appropriate manner (see Appendix A). Although initially intended as information to be included by a test publisher in a test manual, over the years many firms have employed the system to label the tests in their catalogues. These labels in turn are then used to restrict sales and thus help to ensure that tests which may be difficult to administer and/or interpret are sold only to qualified purchasers.

In the letter to CPA the author alleged that one firm had labelled inappropriately several of the tests in its catalogue. According to the author, this firm was making available to the public as Level A tests, tests which should have received a Level B classification. The author's concern was that by assigning a lower level to these tests the firm may have been selling the tests to individuals who lacked proper qualifications.

Because the allegation rested on the assumption that firms typically agree with one another when they make use of the three-level system, before responding to the allegation it was considered important to verify this assumption. To this end an investigation was undertaken of the catalogues issued in 1991/92 by a sample of 17 North American firms. The aim of the investigation was to determine if, in fact, there is consistency among firms when firms assign either a Level A, B, or C to a given test. During the course of the investigation information was also gathered on the second safeguard that the test industry normally employs to protect the public, namely, test user qualification statements. The findings from this investigation, which are contained in two preliminary reports (Simner, 1992, 1993), revealed a number of serious shortcomings in the implementation of both safeguards. Because it was these shortcomings which, for the most part, led to the recommendations in the present report, the outcome of this investigation coupled with the nature of the shortcomings that emerged from the investigation are summarized below.

SAFEGUARD #1. THREE-LEVEL TEST CLASSIFICATION SYSTEM

As mentioned, the investigation involved material in the catalogues issued by 17 firms, six of which

made use of the three-level system. Since the allegation dealt only with tests employed primarily in educational settings, the investigation was confined to tests of this nature. In total, 139 tests that matched this description were located in the catalogues and, of this number, 27 tests were listed in the catalogues of more than one firm.

A tally sheet was constructed that contained the A, B, and C Levels assigned by each of the six firms to each of the 27 tests. Prior to evaluating this data, however, it was important to consider that the APA system, as shown in Appendix A, provides only limited information on the professional qualifications associated with the three levels. Because of this limitation each firm had developed its own set of qualifications. Thus, rather than simply note whether the firms employed the same A, B, or C Level when they referred to the same test, it was necessary to make use of the operational definitions for these levels that the firms themselves had developed.

With this added information in mind it then became possible to compare, across firms, the actual training needed according to each firm in order to purchase each of the 27 tests in the sample. Contrary to the assumption that firms typically agree with one another when they employ the APA system, the results of these comparisons revealed that disagreements outnumbered agreements. Specifically, the comparisons showed that in 18 out of the 27 tests, one firm required less training than another firm to purchase the same test. Moreover, these disagreements among the firms were not confined to only a few firms but instead were scattered across all six firms. Therefore it would appear that the allegation of improper labeling, in reality, might not be a charge that applies only to the firm in question but instead may be a charge that reflects a fairly widespread practice in the test distributing industry.

What are the consequences of this practice? Consider, for example, the Strong Interest Inventory. This test, which is frequently used in educational settings for the purpose of making academic and career counselling decisions, was available for purchase from three of the firms. Moreover, this test was referred to as a B Level product by all three firms. Despite this common label, however, two of the firms disagreed with the third firm on the training needed to purchase a B Level product. As a result of this disagreement, whereas in order to purchase the Strong from one of the firms an individual needed the equivalent of a Master's degree, to purchase the Strong from the other two firms, an individual could have as little as one course in measurement from an accredited college. Hence, in terms of the first shortcoming, this example illustrates how a person who may not be qualified to purchase a test according to one firm that makes use of the three-level system, can still gain access to the test through another firm which not only makes use of the system but also assigns the same level in the system to that test.

The second shortcoming had to do with the category headings that the firms assigned to the tests in their catalogues. As shown in Appendix A, in the three-level system category headings such as achievement, proficiency, aptitude, projective, and mental are quite important because these headings are tied directly to the training needs in the system itself. Thus, for proper use of the system, all tests should appear in catalogues under headings that either match or closely resemble the headings in this system. Furthermore, in the case of an individual test listed in more than one catalogue, that test should appear

under the same heading in each catalogue. Unfortunately, however, consistency of this nature was rarely encountered in the catalogues under review. For instance, the Bender Visual Motor Gestalt Test was listed in four catalogues under as many as eight different headings. Aside from the fact that a number of these headings were not even mentioned in the three-level system, in the case of the few headings where a reasonable match was evident, the headings themselves were associated with different levels in the system. Thus, whereas the one firm that referred to the Bender as a test of mental retardation and therefore should have sold this test as a Level C product, the other firm that referred to the Bender as an achievement test could have sold this test, with equal legitimacy, as a Level A product.

The third shortcoming to emerge from this investigation had to do with the Level A designation. According to the three-level system, the guideline for assigning an A Level to a given test is whether that test "can be adequately administered, scored, and interpreted with the aid of the manual and a general orientation to the kind of organization in which one is working." No mention is made in this guideline of the need for a purchaser to be able to evaluate the psychometric properties of the test. Since three of the firms offered unrestricted access to A Level tests, the problem here is whether there are indeed tests to which purchasers without any background in testing should have such access. For instance, of the 27 tests in the sample, the following were designated by several of these firms as Level A products: AAMA Adaptive Behavior Scale-School Edition, the Gray Oral Reading Tests-Revised, the Test of Language Development-2, the Test of Written Language-2, and the Wide Range Achievement Test-Revised. To make proper use of these tests, however, an individual would need to be familiar with such measurement terms as standard scores, percentile ranks, age equivalents, and Rach Method. Because these are terms that very few individuals who lack measurement training would understand, it would certainly seem that access to tests of this nature should have been restricted to individuals who had completed at least one university course in measurement.

The final shortcoming associated with this first safeguard was the infrequent use that was made of the safeguard. Recall that of the 17 firms whose catalogues were reviewed, 11 firms failed to employ the three-level system. Nevertheless, many of these 11 firms had in their catalogues the same tests that were listed in the catalogues of the six firms that did use the system. Of particular concern in this regard were those tests that received C Level ratings by the firms that employed the system and, therefore, were restricted to individuals who had completed an advanced degree and had received training in measurement. Quite often these tests were treated by the firms that did not use the system as items that were readily accessible to all purchasers regardless of testing experience. Take, for example, the Quick Neurological Screening Test (QNST). This test was listed in one catalogue as a restricted C Level item, which is in line with comments made by Adams (1985) in his review of the QNST in the Ninth Mental Measurement Yearbook. Despite Adams' comments and the fact that for appropriate use the QNST should only be administered by individuals with graduate training in both neuropsychology and measurement, this test appeared in the catalogues of three other firms that did not employ the system. In the case of each firm the QNST was recommended for use by classroom teachers without the need for the teachers to have completed even one undergraduate course in either neuropsychology or measurement.

SAFEGUARD #2. TEST USER QUALIFICATION STATEMENTS

Unlike the first safeguard, which only six of the 17 firms employed, test user qualification statements, as the second major safeguard, appeared in the catalogues issued by 14 of the 17 firms. These statements, which prospective purchasers are required to complete, contain a series of questions that deal with the purchaser's prior training and experience in testing. The answers to these questions are then used to help firms decide whether a given individual is indeed qualified to buy a given test. Unfortunately, however, nine of the 14 firms that made use of test user qualification statements also had waiver clauses in their catalogues that exempted certain individuals from the need to complete these statements. Moreover, because the exemptions typically were based on the individual's occupation or professional affiliation and had little to do with the individual's training, these exemptions by themselves could easily give unqualified persons access to tests that otherwise would be restricted to appropriately trained personnel.

Teachers, for instance, frequently received an exemption under a waiver clause that allowed them to purchase both Level A and Level B tests. According to a recent survey by Rogers (1991), however, of 33 teacher training faculties across Canada, only five required teachers who majored in elementary education to complete a course in measurement. In fact, Rogers, in commenting on the findings from his survey, estimated that "approximately 60% of prospective elementary school teachers and up to three quarters of future secondary teachers will not complete an undergraduate measurement and evaluation course prior to beginning to teach" (p. 185-186). Thus there is good reason to believe that many teachers may not have the proper background to employ correctly either a Level A or a Level B test.

To illustrate the seriousness of this matter consider, for instance, the Bruininks-Oseretsky Test of Motor Proficiency and the Vineland Adaptive Behavior Scales (classroom edition), two tests that often are used in educational settings at the elementary level. Both tests were evaluated in Hammill, Brown, and Bryant (1989) by several independent reviewers. In the case of both tests, because the majority of the subscales were unacceptable from the standpoint of reliability and validity, the reviewers concluded that neither test could be recommended for use. Nevertheless, both tests were listed in one catalogue as Level B products and the waiver clause in that catalogue for a Level B product extended to classroom teachers. In light of the reviewers' comments it would certainly seem that the normal requirement, as stated in the same catalogue, of graduate training in measurement, guidance, etc. to purchase a Level B product should not have been waived in the case of these two tests and, in particular, that teachers with little or no training in measurement should not have been given ready access to either of these tests.

Consultation

As mentioned earlier, the findings summarized above are contained in two reports (Simner, 1992, 1993). Both reports were distributed to the members of the Working Group and to the Chairs of each CPA section followed by open meetings held during the 1992 and 1993 CPA Conventions to discuss the

findings along with a number of preliminary recommendations that also were contained in the two reports. Because many of the recommendations called for what several members of the Working Group referred to as a radical restructuring of the current set of test industry safeguards, it was suggested that the recommendations should be circulated to as large an audience as possible before they appeared in final form. The CPA Board of Directors, while approving the recommendations in principle, concurred with this suggestion. Hence, in 1993 both reports were sent for comment to the following organizations: all of the Canadian provincial psychological associations, the Canadian Association of School Psychologists, the APA Committee on Psychological Tests and Assessment, the European Test Publishers Group, the European Federation of Professional Psychological Associations' Task Force on Assessment, and the Association of Test Publishers. In addition, a number of individuals who were considered experts in the field of assessment were asked to share their views on these recommendations. All of the comments received were then consolidated into a third report (Simner, 1994) which was distributed to the Working Group and to the CPA Section Chairs. This third report served as the basis for further discussion at an open meeting held during the 1994 CPA Convention.

In essence, the final set of five recommendations that appear below and which were approved by the CPA Board of Directors in November, 1994, are the result of a consultation process that lasted more than two years and that involved input not only from the members of the Working Group but also from a number of individuals, professional associations, and trade organizations in Canada, the United States, and Europe. Although the members of the Working Group as well as the Board of Directors recognize that these recommendations do not provide perfect solutions to all of the problems of test misuse that are now being discussed in the literature (see for example Camara & Schneider, 1994; Dahlstrom, 1993; Matarazzo, 1990; Merenda, 1990; Sternberg, 1992), it is hoped that they will provide reasonable and timely solutions to at least some of these problems.

Recommendations

RECOMMENDATION 1. The three-level test classification system currently used by firms to categorize tests should either be replaced or supplemented by a purchaser classification system which recognizes that tests typically are employed for different purposes and that it is these different purposes which should determine whether an individual is qualified to purchase a given test.

Comment. While there was considerable agreement among the members of the Working Group as well as among those who took part in the consultation process with the need to revise the three-level system, at the same time questions were raised about the usefulness of this system. Because the system in its present form requires each test to be assigned either an A, B, or C rating, there was concern over whether such a limited rating scheme can adequately accommodate the range of skills that may be needed to administer properly the more than 2500 tests that are now in print (Sweetland & Keyser, 1991). Also, to make appropriate use of the present system, or even some expanded version of this system, it was felt that the ratings themselves should be awarded by an organization with no commercial ties to the tests being

rated. From a practical standpoint, however, it was considered unlikely that there would be any professional organizations or private firms with either the willingness or the resources to review and assign ratings, on a regular basis, to all the tests that are now entering the market. For example, the eleventh edition of the *Mental Measurements Yearbook* (Kramer & Conoley, 1992) lists 477 new or revised commercially available instruments that have appeared on the market in the three year period since the tenth edition of the yearbook was published. In fact, because of this recent proliferation of tests, for the first time in its history the Buros Institute was forced to restrict itself to reviewing only tests that are sold commercially in the United States.

The solution to this problem suggested by the Working Group is either to replace or supplement the current three-level system, which classifies tests, with a three-category system which classifies individuals who have legitimate reasons to purchase tests. The first category would consist of individuals who purchase and use tests for decision-making purposes (e.g., counselling clients, making hiring decisions, grade placement decisions, clinic referrals, etc.) as well as individuals who teach students who eventually will purchase and use tests for these purposes. The second category would refer to individuals who purchase and use tests for research purposes while the third category would include persons such as librarians who purchase and store tests for use by others as reference material.

Moreover, from the standpoint of safeguarding the public, because it was felt that the consequences of test misuse are likely to be far more serious when tests are employed for decision-making purposes, the training requirements for the first category of purchasers should be more demanding than the requirements for purchasers in the other two categories. At the same time, though, because it is widely recognized that not all tests require the same level of training to be used appropriately, even for decision-making purposes, different training requirements should still be specified in the catalogues for certain tests. For instance, tests such as the Bender Visual Motor Gestalt, the Luria-Nebraska Neuropsychological Battery, the Minnesota Multiphasic Personality Inventory, the Minnesota Test for Differential Diagnosis of Aphasia, and the Wechsler Intelligence Scale for Children, have traditionally required supervised experience in test administration and interpretation in the form of either a practicum or internship. Therefore, tests of this nature should remain restricted and should not be sold to individuals who fail to meet this additional requirement. Hence, in the case of firms that do not employ the three-level system, either all or at least a representative sample of these tests selected from the firm's own catalogue, should be named in a separate section of that catalogue.

Finally, to be fair to purchasers, the members of the Working Group also felt that allowances must be made for alternative training procedures. In other words, it is improper to assume that university level courses in measurement provide the only means for gaining an appropriate background in testing when, today, a number of professional organizations as well as firms offer continuing education workshops which may serve this purpose.

RECOMMENDATION 2. All first-time purchasers, regardless of background, should be required to

complete a test user qualification statement. Hence, firms that publish and/or distribute tests should remove from their catalogues all waiver clauses based on occupation, professional membership, level of graduate training, etc. that exempt certain individuals from the need to complete such a statement.

Comment. Much concern was expressed, not only by the members of the Working Group but also by the majority of those who were contacted by the Working Group, over the use of waiver clauses. In particular there was considerable agreement that neither membership in a professional organization nor an advanced degree should be used as a substitute for appropriate training in measurement. First, there are no stipulations in the membership requirements of professional organizations such as CPA or APA that call for competence in measurement. Second, many graduate programs no longer require students to complete courses in psychological measurement and only about one-quarter of the graduate departments now even rate their students as skilled in the use of psychometric methods (Aiken et al., 1990). Hence, there is simply no assurance that professional membership or an advanced degree, by itself, qualifies an individual to properly evaluate or administer a psychometric test.

RECOMMENDATION 3. The responsibilities assumed by test purchasers and by test distributors in order to safeguard the public against test misuse must be clearly defined.

Comment. The ethical standards adopted by CPA and APA that govern the conduct of their members in the use of psychological tests are quite clear. However, the distinction between these standards and the safeguards employed by the testing industry to protect the public against test misuse seems less clear, at least according to the industry representatives to the Working Group. Indeed there is some merit to this claim in that publishers and distributors of tests often are blamed for test misuse as illustrated in the comments received by Pope and Vetter (1992) in a national survey on ethical dilemmas encountered by members of the American Psychological Association. For example, one psychologist stated that "When the Binet IV came out, only one person was sent for training...We are often asked to add new tests without appropriate supervision. Test publishers aren't motivated to slow down sales by requiring training to purchase tests" (p. 405). At issue, of course, is who should bear the responsibility for ensuring that tests are employed appropriately. To clarify this matter the following divisions of responsibility were recommended.

Test Purchaser The ultimate responsibility for the appropriate use of a test should rest with the user of that test. This position is one that has long been accepted by the publishing industry and is expressed in the following statement which appears in the catalogues issued by most firms today: "A test user should know his (or her) own qualifications and how well they match the qualifications required for the uses of specific tests" (APA, AERA, NCME, 1974, p. 58, item G1.1).

At the same time, however, professional bodies as well as firms that adhere to a code of ethics or to standards of practice which contain provisions on testing, and that adjudicate complaints that stem from violations of these provisions, can help to ensure that a test purchaser will indeed abide

by the voluntary compliance component in this statement. Thus, a test purchaser should either be a member of such an association, registered by a provincial regulatory body, or be employed by such a firm. To be clear on this point the intent of the Working Group was to ensure that there is an external body that can act as a monitoring agent to which the public can turn in order to lodge a complaint when there is reason to believe that a test has been misused. For example, the most recent version of the ethics code issued by the American Psychological Association requires that psychologists "refrain from misuse of assessment techniques, interventions, results, and interpretations and take reasonable steps to prevent others from misusing the information these techniques provide" (APA, 1992, p. 1603, Standard 2.02b). The Ethics Committee of the American Psychological Association not only has the authority to respond to complaints that may arise from violations of this provision, it also has the right to admonish its members and has done so in the past when allegations of test misuse have been supported (for examples of sanctions that have been applied in cases of test misuse by the Ethics Committee see APA, 1987, p. 109-120).

Test Publishers/Distributors Publishers/distributors should be responsible for the appropriate use of tests only to the point of sales. This means that although it is essential to ensure that tests are properly advertised and sold only to individuals who meet approved training requirements, it should not be the responsibility of either a publisher or a distributor also to ensure that tests are used properly once they leave the firm. Nor should it be their responsibility to offer courses in testing. Instead, when new tests appear on the market that require additional skills, the responsibility for offering continuing education workshops or mini-courses to train individuals in the proper use of these new tests should be assumed by professional associations, graduate departments, etc.

RECOMMENDATION 4. Firms should be encouraged to insert in their catalogues the *Who May Purchase Tests* statement in Appendix D as well as the *Test User Qualification Statement* in Appendix E.

Comment. These two statements were developed by the Working Group in order to operationalize the provisions in the foregoing recommendations as well as to ensure that prospective purchasers are made aware of these provisions. Firms, of course, may add to the training requirements in both statements if they choose to do so. It is worth noting that, in addition to operationalizing the provisions mentioned above, clauses were added to both statements in order to address a number of further issues that were brought to the attention of the Working Group. For example, several organizations mentioned that it is not uncommon for firms to receive orders from schools, hospitals, and other agencies with no individual names on the order forms. When this happens it is impossible for a firm to know whether the person who initiated the order had previously completed a *Test User Qualification Statement* and, therefore, whether that person is indeed qualified to purchase the test is question. To deal with this matter the *Who May Purchase Tests* statement requires that institutional purchase orders be "countersigned by an individual who has on file a completed *Test User Qualification Statement*."

RECOMMENDATION 5. In order to encourage firms to make use of the *Who May Purchase Tests* statement as well as the *Test User Qualification Statement* a system should be established whereby firms that reproduce this material (or a close approximation thereof) in their catalogues receive recognition. Such a system, however, should not imply an endorsement by the Canadian Psychological Association of either the products or the business practices of a particular firm.

Implementation:

(A) Firms that request recognition should be granted permission to place the following statement in their catalogues.

The information in the *Who May Purchase Tests* statement as well as in the *Test User Qualification Statement* in this catalogue complies with the Canadian Psychological Association's recommendations on the proper sale and distribution of tests. Canadian Psychological Association (CPA logo)

To receive such recognition firms must submit, for review, copies of their most recent catalogues to the Head Office of the Canadian Psychological Association. These copies will be distributed by the Chair of the Professional Affairs Committee to the Chairs of the appropriate sections of the Association (e.g., Clinical, Educational/School, Industrial/Organizational). The Section Chairs (or their designates), in turn, will review the catalogues to ensure that the relevant material is in keeping with Recommendations 1, 2, and 3 and that appropriate examples of restricted tests are listed in paragraph four under Category 1 of the *Who May Purchase Tests* statement.

(B) Firms that satisfy the foregoing provision should receive further recognition by having their names appear in the CPA newspaper *Psynopsis* on an ongoing basis under the following heading.

The information in the *Who May Purchase Tests* statement as well as in the *Test User Qualification Statement* in the catalogues issued by the following firms complies with the Canadian Psychological Association's recommendations on the proper sale and distribution of tests.

Comment. Whether or not a firm acts upon a recommendation made by a professional organization, such as the Canadian Psychological Association, is strictly voluntary. In view of this shortcoming, there was considerable agreement with the need to create a method to encourage firms to comply with the recommendations contained in this report. The method described in this recommendation was suggested by the test industry representatives to the Working Group. It is worth noting that this method has the advantage of not only providing recognition to firms that seek recognition, but also of enabling the Canadian Psychological Association to monitor the catalogues of these firms. Since 11 of the 17 firms whose catalogues were reviewed in the investigation cited above, not only failed to employ the three-level system but also failed to designate any of the tests in their catalogues as restricted items, such monitoring by the Association of the *Who May Purchase Tests* statement could prove highly advantageous as a further safeguard to the public.

Acknowledgements

The Board of Directors of CPA acknowledges with appreciation the many valuable contributions made by the following members of the Working Group: James Breeze, Jean-Marc Chevrier, Rick D. Hackett, Denise Hughes, Nicholas Skinner, Steven Stein, and Beatrice Wickett. The Board also extends its appreciation to the following individuals who responded to the Working Group's requests for comments on earlier drafts of the recommendations: Dave Bartram, Mark G. Borg, Patrick J. Carney, Riet Dekker, Ian Florance, Hans Gerhardt, John Goyeche, Bertil Mardberg, Peter Merenda, Sam Mikail, Thomas Oakland, Sampo V. Paunonen, Gary J. Robertson, Sylvia Rosenfield, Phyllis Simner, Jennifer A. Veitch, and P. Anthony Vernon.

References

- Adams, R. L. (1985). Review of Quick Neurological Screening Test. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook*. Lincoln, NB: University of Nebraska Press.
- Aiken, L. S., West, S. G., Sechrest, L., Reno, R. R., Rodiger, H. L., III, Scarr, S., Kazdin, A. E., & Sherman, S. J. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of Ph.D. Programs in North America. *American Psychologist, 45*, 721-734.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Psychological Association. (1987). *Casebook on ethical principles of psychologists*. Washington, DC: American Psychological Association.
- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1597-1611.
- American Psychological Association, American Educational Research Association, and National Council on Measurements Used in Education. (1954). Technical recommendations for psychological tests and diagnostic techniques [Supplement]. *Psychological Bulletin, 51*, 1-38.
- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1974). *Standards for educational & psychological tests*. Washington, DC: American Psychological Association.
- APA Committee on Ethical Standards for Psychology. (1950). Ethical standards for the distribution of psychological tests and diagnostic aids: Code of standards for test distribution. *American Psychologist, 5*, 620-626.
- Camara, W. J., & Schneider, D. L. (1994). Integrity tests: Facts and unresolved issues. *American Psychologist, 49*, 112-119.
- Canadian Psychological Association. (1987). *Guidelines for educational and psychological testing*. Ottawa, ON: Canadian Psychological Association.
- Dahlstrom, W. G. (1993). Tests: Small samples, large consequences. *American Psychologist, 48*, 393-399.

References (cont.)

- Hammill, D. D., Brown, L., & Bryant, B. R. (1989). *A consumer's guide to tests in print*. Austin, TX: Pro-ed.
- Kramer, J. J., & Conoley, J. C. (Eds.). (1992). *The eleventh mental measurements yearbook*. Lincoln, NB: The Buros Institute of Mental Measurements.
- Matarazzo, J. D. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist*, 45, 999-1017.
- Merenda, P. F. (1990). Present and future issues in psychological testing in the United States. *Evaluacion Psicologica/ Psychological Assessment*, 6, 3-31.
- Pope, K. S., & Vetter, V. A. (1992). Ethical dilemmas encountered by members of the American Psychological Association. *American Psychologist*, 47, 397-411.
- Rogers, W. T. (1991). Educational assessment in Canada: Evolution or extinction? *The Alberta Journal of Educational Research*, 37, 179-192.
- Simner, M. L. (1992). *Preliminary report of the Professional Affairs Committee Working Group on Test Publishing Industry Safeguards*. Ottawa, ON: Canadian Psychological Association.
- Simner, M. L. (1993). *Interim report of the Professional Affairs Committee Working Group on Test Publishing Industry Safeguards*. Ottawa, ON: Canadian Psychological Association.
- Simner, M. L. (1994). *Final report of the Professional Affairs Committee Working Group on Test Publishing Industry Safeguards*. Ottawa, ON: Canadian Psychological Association.
- Sternberg, R. J. (1992). Ability tests, measurements, and markets. *Journal of Educational Psychology*, 84, 134-140.
- Sweetland, R. C., & Keyser, D. J. (Eds.). (1991). *Tests: A comprehensive reference for assessments in psychology, education, and business* (3rd ed.). Austin, TX: Pro-ed.

Appendix A

Training needs for each level in the three-level test classification system developed by the American Psychological Association (from APA, AERA, NCMUE, 1954, p. 11-12).

Training Needs

Level A - Tests or aids which can be adequately administered, scored, and interpreted with the aid of the manual and a general orientation to the kind of organization in which one is working. (E.g., achievement or proficiency tests).

Level B - Tests or aids which require some technical knowledge of test construction and use, and of supporting psychological and educational subjects such as statistics, individual differences, and psychology of adjustment, personnel psychology, and guidance. (E.g., aptitude tests, adjustment inventories with normal populations.)

Level C - Tests and aids which require substantial understanding of testing and supporting psychological subjects, together with supervised experience in the use of these devices. (E.g., projective tests, individual mental tests.)

Appendix B

The A, B, C Levels assigned to 27 tests by six firms that listed these tests in their 1991/92 catalogues.

Test	Firm					
	CPP	MHS	PCAN	PCOR	PMET	RPSY ¹
AAMD Adaptive Behavior Scale-School Edition		A		B		
Basic Personality Inventory					C	C
Behavioral Academic Self-esteem	B				B	
California Psychological Inventory	C			C	C	
Children's Apperceptive Story-Telling Test		B	B	C		
Children's State-Trait Anxiety Inventory	B				B	
Conners' Rating Scale		B		B		
Coopersmith Self-Esteem Inventory	B				B	
Detroit Tests of Learning Aptitude - Primary		B	B			
Gray-Oral Reading Tests - Revised		A	B			
Guilford-Zimmerman Aptitude Survey	B				B	

Appendix B (cont.)

Test	Firm					
	CPP	MHS	PCAN	PCOR	PMET	RPSY ¹
Jackson Personality Inventory					B	B
Jackson Vocational Interest Survey					B	B
Jesness Behavior Checklist	C	B				
Jesness Inventory of Adolescent Personality	C	B				
Murphy-Meisgeier Type Indicator for Children	B				B	
Myers-Briggs Type Indicator	B				B	
Psychological Screening Inventory					C	B
Readiness for Kindergarten	A				A	
School Readiness Survey	A				A	
School Situation Survey	B				B	
Strong Interest Inventory	B			B	B	
Test of Language Development - 2		A	A			

Appendix B (cont.)

Test	Firm					
	CPP	MHS	PCAN	PCOR	PMET	RPSY ¹
Test of Nonverbal Intelligence - 2		B	B			
Test of Written Language - 2			A	B		
Wide Range Achievement Test - Revised		A		B		
Wisconsin Card Sorting Test		A		C		

- ¹CPP = Consulting Psychologists Press
MHS = Multi-Health Systems, Inc.
PCAN = Psycan
PCOR = Psychological Corporation
PMET = Psychometrics Canada, Ltd.
RPSY = Research Psychologists Press, Inc.

Appendix C

Training needs for each level in the three-tier test classification system adopted by six test industry firms.

Firm		Training needs
1) Consulting Psychologists Press, Inc.	Level A -	Available to any purchaser
	Level B -	Satisfactory completion of a % course in the interpretation of psychological tests and measurements at an accredited college or university
2) Psychometrics Canada	Level C -	Fulfilled the Level B qualification plus one or more of the following: completion of an advanced degree in an appropriate profession, membership in an appropriate professional association, state licensure, and/or national or state certification.
	Level A -	Unrestricted
	Level B -	The user has completed courses in tests and measurement at a university or received equivalent documented training.
3) Multi-Health Systems, Inc.	Level C -	Requires training or experience in the use of tests and completion of an advanced degree in an appropriate profession (e.g., psychology, psychiatry), or membership in a relevant professional association (APA), or a state certificate in a relevant regulated profession (e.g., psychology).
	Level A -	User has completed at least one course in measurement, guidance, or an appropriate related discipline or has equivalent supervised experience in test administration and interpretation.
	Level B -	User has completed graduate training in measurement, guidance, individual psychological assessment, or special appraisal methods appropriate for a particular test.
4) Psycan	Level C -	User has completed a recognized graduate training program in psychology with appropriate course work and supervised practical experience in the administration and interpretation of clinical assessment instruments.

Appendix C (cont.)

Firm	Training needs
5) Research Psychologists Press	Level A - May be purchased for business and Psychologists educational settings.
	Level B - Available to those individuals who have completed an advanced level university course in psychological testing at the Master's level, as well as training under the supervision of a qualified psychologist.
	Level C - Available to those individuals who obtained a doctoral level degree in psychology or education, or who are members of qualified professional organizations, or who are under the direct supervision of a qualified psychologist.
6) Psychological Corporation	Level A - Verification of licensure or certification by an agency that is recognized by the Psychological Corporation to require training and experience in a relevant area of assessment that is consistent with the expectations outlined in the 1985 <i>Standards for Educational and Psychological Testing</i> .
	Level B - Verification of a master-level degree in Psychology or Education or the equivalent in a related field with relevant training in assessment or, verification of membership in a professional association (for example, APA) that is recognized by the Psychological Corporation to require training and experience in a relevant area of assessment that is consistent with the expectations outlined in the 1985 <i>Standards for Educational and Psychological Testing</i> .
	Level C - Verification of a Ph.D.-level degree in Psychology or Education or the equivalent in a related field with relevant training in assessment or, verification of licensure or certification by an agency that is recognized by the Psychological Corporation to require training and experience in a relevant area of assessment that is consistent with the expectations outlined in the 1985 <i>Standards for Educational and Psychological Testing</i> .

Appendix D

Who May Purchase Tests

Generally speaking, there are three major categories of legitimate test purchasers. The first category consists of individuals who purchase and use tests for decision-making purposes, whether in schools, hospitals, personnel departments, or other work settings. This category also applies to individuals who train students in the use of tests that will be employed for these purposes. The remaining categories consist of individuals who purchase tests for research purposes or for library use. *Regardless of category, however, all first-time purchasers must complete the Test User Qualification Statement in this catalogue and submit this statement together with their purchase order. Orders from individuals who do not have on file a completed Test User Qualification Statement will not be filled. Purchase orders from institutions (schools, hospitals, etc.) must be countersigned by an individual who has on file a completed Test User Qualification Statement.* The following criteria govern the sale of tests to individuals in each category.

Category 1: Tests purchased for teaching or decision- making purposes

An individual who employs a test for the purpose of teaching students, counselling clients, making hiring decisions, grade placement decisions, clinic referrals, etc. must have successfully completed a minimum of two university courses in tests and measurement. The first course, on basic principles of psychological measurement, should have been completed at the undergraduate or graduate level, and should have included information on such topics as scaling, transformations and norms, as well as information on factors that affect reliability and validity.

The second course, at the graduate level or its equivalent, should have included material on assessment which is appropriate to the test(s) being ordered. Because of the variety of tests listed in this catalogue, coupled with the overlapping content of many advanced level courses in testing, it is not possible to specify by name which courses are necessarily most suitable for which tests. A course entitled "Educational Assessment", for example, may contain material on achievement testing and vocational guidance coupled with a practicum, or it may only contain material on the administration and scoring of achievement tests. For this reason the (firm name) adheres to the position adopted by the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education that "a test user should know his (or her) own qualifications and how well they match the qualifications required for the uses of specific tests" (APA, AERA, NCME, 1974, p. 58).

To be clear on this matter, although it is essential for (firm name) to ensure that tests are properly advertised and sold only to individuals who meet approved training requirements, it is not the responsibility of (firm name) also to ensure that tests are used properly once they leave the firm. That responsibility remains with the test purchaser even in cases where the purchaser delegates the administration and/or interpretation of a test to someone else.

Appendix D (cont.)

In addition to having completed an appropriate advanced level course in testing, there are also certain tests which are widely recognized as requiring supervised experience in test administration and interpretation in the form of either a practicum or internship. In this catalogue these tests include, but are not necessarily limited to, the following: (firms should select appropriate examples from their own catalogues). Tests of this nature will not be sold to individuals who fail to meet this additional training requirement. The (firm name) retains the right to determine whether a person is indeed qualified to purchase a given test.

The (firm name) further recognizes that some individuals may have acquired an appropriate background in testing in ways other than through course work. Although it is not the responsibility of the (firm name) to offer courses in testing, to avoid imposing unfair restrictions on the sale of tests, if you feel that you are qualified to purchase a given test as the result of having successfully completed an appropriate advanced level workshop or mini-course sponsored by a professional association, college, university, etc. you should describe in a letter the nature of your background. The letter should then be submitted together with the purchase order form. Final approval to purchase a test will rest with the (firm name).

In addition to course requirements an individual who purchases a test for teaching or decision-making purposes must agree to abide by the principles that apply to appropriate test use as set forth in the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1985) and in the Guidelines for Educational and Psychological Testing (CPA, 1987). Such an individual also must either be a member of a professional association, registered by a professional association, or be employed by a firm that has endorsed these principles and has adopted a code of ethical conduct which is similar to the code of ethical conduct approved by the Canadian Psychological Association or the American Psychological Association. The reason for this last requirement is to ensure that there is an external body that can act as a monitoring agent to which the public can turn in order to lodge a complaint when there is reason to believe that a test has been misused.

Summary of the Major Provisions under Category 1

To purchase tests for teaching or decision-making purposes an individual must satisfy all five of the following criteria.

- 1) Successful completion of a university undergraduate or graduate level course on basic principles of psychological measurement that included information on such topics as scaling, transformations and norms, as well as information on factors that affect reliability and validity
- 2) Successful completion of a university graduate level course (or equivalent) that included

Appendix D (cont.)

material on assessment which is appropriate to the test(s) being ordered.

(An individual who feels qualified to purchase a given test as the result of having successfully completed an appropriate advanced level workshop or minicourse sponsored by a professional association, college, university, etc., should submit a description of the workshop together with the purchase order form.)

- 3) Willingness to abide by the principles that apply to appropriate test use as set forth in the *Guidelines for Educational and Psychological Testing* (CPA, 1987) and in the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1985).
- 4) Membership in a professional association, registered by a professional association, or employed by a firm that has adopted a code of ethical conduct similar to that approved by the Canadian Psychological Association or the American Psychological Association.
- 5) Completion of a Test User Qualification Statement. (All Test User Qualification Statements submitted to (firm name) remain on file for future reference. Therefore, once a Statement has been completed an individual is not required to complete a further Statement for future purchases. *All institutional purchase orders must be Countersigned by an individual who has on file a completed Test User Qualification Statement.*)

To purchase restricted tests such as (firms should list examples from their own catalogues), in addition to meeting all five of the criteria listed above, an individual must have successfully completed supervised experience in test administration and interpretation in the form of either an internship or practicum. The (firm name) retains the right to determine whether an individual is indeed qualified to purchase a given test.

Category 2: Tests purchased for research purposes

Tests and test materials that will be used for research purposes only may be purchased by college and university faculty members as well as by professional staff of hospitals and business organizations. Students enrolled in graduate programs may also purchase materials that will be used only for research purposes if the purchase order is countersigned by the student's supervisor or by the department head. Regardless of the purchaser, however, the order must be submitted on business or letterhead stationery and the purchaser must satisfy either the professional membership requirement, professional registration requirement, or the employment requirement specified under Category 1. In the case of students, enrollment in a graduate program that adheres to a code of ethical conduct similar to the code approved by the Canadian Psychological Association or the American Psychological Association will satisfy these requirements.

Appendix D (cont.)

Category 3: Tests purchased for library reference purposes

Tests and test materials for library use are sold to individuals upon the assurance that these products will be made available only to persons who meet the criteria specified under Category 1 and Category 2. This assurance must be submitted in writing on letterhead stationery before the order will be processed. The letter containing this assurance must be countersigned by an individual who satisfies either the Category 1 or Category 2 criteria.

Appendix E

Test User Qualification Statement

Name _____

Work Address _____

Work phone _____ Work fax _____

Highest Level of Education _____

A. Evidence of membership in an association, registration by an association, employment in a firm, or enrollment in a graduate program that abides by the principles of appropriate test use and has adopted a code of ethical conduct that is similar to the code approved by the Canadian Psychological Association or the American Psychological Association.

1) I am a member of the following organization(s)

AERA APA ASHA CPA CASP NASP Other _____

2) I am registered by _____

3) my registration number is _____

4) I am employed by _____

5) I am a enrolled in _____

B. Evidence of Appropriate Training in the Use of Tests

1) I have successfully completed the following course(s) in testing:

course name

Appendix E (cont.)

Institution

level (undergraduate, graduate or its equivalent)

2) I have successfully completed a practicum or internship in testing in the following area(s):

3) I have successfully completed an appropriate advanced level workshop or mini-course sponsored by a professional association, college, university, or firm that qualifies me to purchase the test(s) listed on the order form. I have included a letter that outlines the nature of this training.

C. Evidence of Acceptance of Responsibility for the Sound Use of Tests (see section entitled "Who May Purchase Tests").

1) I wish to purchase the test(s) listed on the attached order form under the following test purchaser category.

Category 1 _____

Category 2 _____

Category 3 _____

2) I plan to use the test for purposes other than the ones outlined under these categories. I have included a letter that fully describes these other purposes.

Your signature indicates that the information on this form is correct, that you agree to abide by the principles set forth under the category checked above, and that you also agree to abide by the regulations that apply to the copyrighted parts of the test(s) you wish to purchase.

Copyrighted parts of tests include test items, scoring algorithms, norms, test booklets, test protocols, etc. This material may not be reproduced in paper or electronic format without

Appendix E (cont.)

written permission from the publisher. Violation of copyright is a federal offense according to the Canadian Copyright Act, R.S.C. 1985, c.C-42.

Signature _____ Date _____