

ISSUES SHAPING THE FUTURE AND ETHICS OF PSYCHOLOGICAL TESTING

REUBEN, NDUBUISI

**Department of Educational Foundations
Rivers State University, Nkpolu – Oroworukwo
Port Harcourt, Rivers State**

And

MAXWELL D. EREMIE, PhD.

**Department of Educational Foundations
Rivers State University, Nkpolu – Oroworukwo
Port Harcourt, Rivers State**

Abstract

This paper discussed issues shaping the future and ethics of psychological testing. It considered some of the professional concerns that continually play essential role in the present state and future of psychological testing such as theoretical concerns, test adequacy, human rights, labelling, privacy invasion, tests constructors/test users' responsibility, dehumanization, usefulness of tests, and current fashion in psychological testing such as the development of new tests, increase in public awareness and influence on psychological testing and the hope of new and improved test. With the increase in the level of awareness of test users and psychological needs of humans, especially with the changes in human body Chemistry which sometimes may have some psychological implications in human ability and personality. Psychologist should be more creative in building new tests that will meet future testing needs of the fast growing population and be persistence in modifying the existing tests to match ability and personality that are always changing and can be measured only within the context in which they are present.

Introduction

Students who major in physical sciences have chosen to study phenomena that are relatively easy to carry out measurement. For example, if such a student want to test for Carbon and Hydrogen. The procedure is simple and straight, all that the student needs to do is: put a little amount of pure Copper (II) Oxide in a dry test-tube and heat gently for a while. Add a little amount of the unknown substance and then heat strongly. Test the colourless gas evolved with lime-water on a glass rod; it turns milky due to the presence of CO₂ gas. Also test the gas with anhydrous CuSO₄; it turns blue due to the presence of water vapour. In psychology, measurement is more difficult. Psychologists are rarely interested in testing simple qualities such as colour change. Instead, they usually pursue latent variables or complex traits such as intelligence or aggressiveness, which one can neither see nor touch with no yardsticks available to measure such characteristics (Mislevy, 2002; Bootz & Mcknight, 2006).

Psychological testing refers to all the possible uses, applications, and underlying concepts of psychological and educational tests (Kaplan & Saccuzzo, 2009). They are tools used by psychologist in gathering data about people. It is an objective procedure for sampling and quantifying human behaviour to make decision about a particular construct or latent variable using standardized stimuli, and techniques of administration, scoring and interpretation. These tests are mainly employed in evaluating individuals. For example, individuals who achieve high on IQ test are seen as having a high degree of intelligence than those who achieve low scores. Hence, the most important purpose of testing is to recognize the ability difference among those participating in the test. In the educational sector today, individuals are first exposed to tests and assessment. This plays a significant role in the evaluation of students such as admission, graduation, retention, placement, and scholarship awards. Tests are also widely employed in screening of individuals for career and vocational, for certification and licensing of Individuals for a number of occupations, selection and placement of workers within government and private sector organizations (Camara, 1997).

In a clinical setting, a test may provide a sample of the behaviour the patient finds disturbing. For example, a patient may suffer an irrational fear of an object that is not actually dangerous, such as goat. Because of this fear, the individual cannot step out of his house unless accompanied by an adult. To measure the magnitude of the irrational fear, the psychologist may ask the individual to approach a goat held in a cage. The distance from the goat that indicates a report of anxiety is taken as an indication of the severity of the client's avoidance behaviour. This can be used to evaluate the effectiveness of any subsequent planned intervention to reduce the problem. Given the diverse and important purpose of tests, evaluators have become increasingly concerned with the issues shaping the field of testing such as professional, moral, ethical issues and questions of validity, fairness, intended uses. Some of these professional concerns that continually play essential role in the present state and future of psychological testing are theoretical concerns, and test adequacy.

Theoretical Concerns

The dependability (reliability) of tests results is one of the most significant considerations underlying tests (Thomas & Selthon, 2003); (Tryon & Bernstein, 2003). Reliability defined as the degree to which knowledge, and skills learned is correct and accurate. That is the extent to which tests are relatively free of measurement error (Abe, 2012). Reliability places an upper limit on validity. A test that is totally unreliable is meaningless. There may be exception to this statement, but general application demands that tests possess some form of reliability or stability. As a direct consequence, whatever is being measured must itself have reliability. To say that a test has reliability implies that tests results are attributable to a systematic source of variance which is reliable itself (American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (AERA, APA, & NCME, 1999; APA, 2002).

Most psychological tests today measure a presumably emotionally balanced entity -either the person as he or she currently functions or some temporally emotional balanced characteristics of the person. In providing current in-depth functioning, psychologists suggest that the individual functions this way in a fairly stable, though perhaps short term manner that is free from outside control or influence of the situation or environment. In other words; psychologists assume that they can give a detailed account of the individual in absolute terms; as if in a vacuum. Psychologists may opine that the individual is emotionally unstable or that the individual is out of contact with the state of things as they actually exist, or else provide a diagnostic label such as Schizophrenic or neurotic. In like manner, and even more noticeably, psychologists support enduring qualities that become noticeable themselves overtime without concern of immediate or long term external (situational, environmental, and so forth) factors. Hence, they assume that what they are measuring exist in absolute terms (APA, 2002). Whether measuring what is presently going on or situation that is temporally stable, examiners always accept as true without proof that the systematic source of variance measured by the test results entirely from the individual rather than some other factors.

When psychologists try to measure a stable quality or characteristic of a person and discovers less than perfect temporal reliability, they assume that the imperfections proceed from test-related inadequacies, like measurement error, or from minor fluctuating subject variables like fatigue. Thus, they assume then that the characteristic or variable being measured is stable; it is present, only the test instrument limits one's ability to measure it. This implies that the more accurate a test, the more the results would be (Kaplan & Saccuzzo, 2009). Evaluators assume that individuals possess stable characteristic such as intelligence and stable response tendencies like traits that keep at a certain level and functions freely or independent of the environment. However, studies by Cacioppo, Bernston, and Anderson (1991) revealed that even the best tests are yet to achieve such temporary stability. This implies that examiners cannot readily attribute differences overtime solely to assessment error or fluctuating subject variables. Therefore, the primary assumption beneath tests is not completely correct; hence, the social environment influences behaviour (Corrigan, Bogner, Mysiw, Clinchot, & Fugate, 2001). Spear (2007) stated that the question of trait does not only apply to personality psychology but extend to the entire body of psychology. In the beginning, human personality theorist tended to view personality as made up of stable and lasting behavioural dispositions (traits). For example, Sigmund Freud and some of his associates, believed that the individual early or childhood experiences, memories, traumas, and anxieties often resulted in behavioural traits, or dispositions that

persisted all through existence. Freud's assumption was however challenge by psychologists who believed that human personality is not fixed and stationary but changes with time and circumstance, or situations and external factors that influence human behaviour.

Human behaviour may be the outcome of long-term stable traits; internal or external conditions or environments that happened before, coexist with, and follow behaviour; or some other fact or such as the interaction between behavioural tendencies (traits) and environments. Traits and environment tests have a long way to go before they will be able to measure the particular determinants of human behaviour or current conceptualizations and the underlying assumptions of tests are not precise enough to make accurate predictions (Bandura, 1986; Mischel, 1968; Ziskin, 1995). This implies that psychological tests can be no better than science of psychology that underlies them. As the science of psychology clarifies basic theoretical issues, testing conforms to the available knowledge. Thus, the single most important theoretical assumption of test-which human behaviours or characteristics are stable and can be freely measured without considering the environment of the individual, is questionable. Bandura (1986); Mischel (1968); and Ziskin (1995) pointed out that although human being exhibits core stability, they continually change. Personality tests usually do not produce the same or similar result after a long period; this is because as the individual adjusts to the environment, he or she changes. This account for why most definitions of intelligence include the ability to adapt or change based on the circumstances or situations. The truth is human being change all the time. Garlick (2002); Sternberg (2001); and Zandra (2003) highlighted that all normal people possess the ability to adapt to changing circumstances or situations. This ability involves a combination of factors that change. These factors are referred as the individual's index of competency which is usually correlated with score of most of the ability test in use today. A person with a high index of competency can adapt more successfully and perhaps discover more effective solutions to environmental challenges than people with a low index. Though, reacting to the environment may not only change behavioural tendencies or traits but also the index of competency. However, an environment that is extremely demanding such as one that pressured a person to call on latent reserves may increase the index of competency. In psychological testing, ability and personality are always changing and can be measured only within the context in which they are present.

Moral Issues in Psychological Testing

The field of psychological testing is being shaped by moral issues such as human rights, labelling, and privacy intrusion

Human Rights

Several different kinds of human rights are recognized in psychological testing. Test takers are usually treated with courtesy, irrespective of gender, ethnicity, state or nation of origin, religious affiliation, and age, etc. Test takers are usually tested with measures that meet professional benchmark or standards that are appropriate, relieve explanation prior to test examination on the kind(s) of tests to be conducted. Individuals who do not want to subject themselves to testing should not and ethically cannot be forced to do so, hence, the individuals freedom to decline, and freedom to withdraw is highly respected unless situation(s) where the testing is mandated by law or government (APA, 2002). The second right due tests takers is there right to know their test results(s) and interpretations as well as the bases of any judgement or decision that affect their lives. Before now, protecting the security of test was of utmost importance. Nowadays, one must still take all the necessary precautions to guard test security, but not at the expense of an individual's rights to know the basis of detrimental or adverse decisions. Test interpreters are under ethical obligation to protect human rights. Potential test takers have the responsibility for knowing and demanding their rights.

Labelling

There is nothing absolutely wrong in diagnosing people with kidney problem or disease, but labelling people with certain Medical disease such as Acquired Immunodeficiency Syndrome (AIDs) and psychiatric disorders can be damaging. For example, a reasonable percentage of the generality of the public are has little understanding of Schizophrenia, when diagnosing this kind of disease, it is advisable to use least stigmatizing label consistent with accurate representation. Labels have the capacity to affect one's access to help. For instance, chronic Schizophrenia is not curable, as such labelling

someone a chronic schizophrenia may be so harmful (McReynolds, Ward, & Singer, 2002). Since the disorder is not curable, nothing can be done. Since nothing can be done, why should anyone bother to help? Since no help is provided, the individual is a chronic case (Corrigan, 2007).

Theoretical is another problem which people unfortunately often justify with psychological tests. Szase (1961) pointed out that a medical label such as schizophrenia means that the individual is ill or diseased. Since nobody can be blamed for becoming ill, a psychiatric or medical label indicates that the individual is not responsible for the health challenge. Though, it may well be that individuals who are label as medically disturbed must take responsibility for their lives if they are to recover or get better. Taking responsibility for one's life implies exercising some degree of control rather than simply being the victim of uncontrollable external forces. People who take control or responsibility for their actions should be able to tolerate more stress, frustration and pain than those who feel like passive victims. Obviously, an individual who assume responsibility or control has more capacity to alter negative conditions than those who does not. However, labelling process may not only stigmatized the individual but also lower tolerance for stress and make treatment more difficult. Thus, judging from the potential negative effects of labels, an individual should have the right not to be labelled (Hammel, Shaffer & Erdberg, 2000; Shaffer, Erdberg, & Haroian, 1999).

Privacy Invasion

When people react or respond to psychological tests, they have little idea what is being revealed. But in many cases they feel that their privacy has been invaded in a way that is not justified by the tests benefits (Brayfield, 1965). Dahlstron (1969) stated that the issue of privacy invasion is based on serious misunderstanding. He maintained that because tests have been oversold, the public doesn't their limitations. Ambiguity of the motion of invasion of privacy is an important issue in psychological tests. There is nothing absolutely wrong or detrimental in trying to find out about a person. It is only the wrong application or use of the information gathered from the person that amount to invasion of the person's privacy. Psychologists are ethically, morally and often legally bound to maintain confidentiality and do not have the audacity to reveal any information concerning test taker(s) than is necessary to achieve the purpose for which testing was initiated. More so, psychologists must notify or inform test takers of the limit of confidentiality. Individuals must comply in order to be tested. They also have the right to refuse or declined to be tested. Confidentiality is guaranteed by law in most states that have laws governing the practice of psychology: like the United States of America (APA, 2002). This means that as a general rule, personal information obtained by the psychologist from any source is communicated only when the person permits it, unless circumstances in which withholding such information may be injurious to the individual or the larger society, as well as situations that that demand subpoenaed records. Individuals or test takers have the right to know the limits of confidentiality and to know that test information can subpoenaed as administered as evidence before a jury or court of law (Benjamin & Gollan, 2003).

Test Constructors and Test Users' Responsibility

The testing profession has become increasingly stringent and precise in defining the ethics and responsibility test designers and test users. This is because even the best test can be misused. In the right circumstance, almost any test can be useful, but when inappropriately used, even the best test can be dangerous to the individual (APA, 2002). A major concern is the utilization of tests with different populations. A test that is reliable and valid for group A may not be valid and reliable for group B. In light of this issue, psychologist who administered tests are instructed to employ instruments whose validity and reliability have been established for use with members of the population being tested and to utilize assessment techniques that are most appropriate to a person's best preferred language. However, in the cause of interpreting test results, it is necessary that psychologists take into consideration the characteristics of the individual being assessed, such as personal, situational, linguistic, and cultural differences that might affect their judgement or reduce the accuracy of their interpretations.

To reduce the potential damage that may arise from test, test user should be made to know the reason for using a particular test, the consequences of using the test, and the necessary procedures to maximize fairness. Thus, test users must possess enough knowledge to understand the principles

governing test development (Dana, 2000; Merrel, 2003; Naar-King, Ellis, & Frey, 2003). Franklin (2003) pointed out that it is the responsibility of the constructor to provide necessary information judging by the current standard for test use, the test designers must provide test manual with enough information to allow appropriate use of test, including adequate reliability and validity information, well defined scoring and administration standards, and a clear description of the group sample (AERA, APA, & NCME, 1999; APA, 2002). In addition, test developers are to ensure that interpretations of data are based on the test results and are justified and that the test is properly utilized. Test users must take the responsibility of finding out all the pertinent information before using any test. The test user must have sufficient knowledge to differentiate between a test that meets present test standard and one that does not. Hence, there is no excuse for employing inadequately documented instrument that has damaging consequences (Embreston & Hershberg, 1999).

Issues of Social Concern

In psychological testing, social issues such as dehumanization, usefulness of tests and access to psychological testing services are of essential importance. This aspect will be limited to dehumanization and usefulness of tests only.

Dehumanization

Some forms of testing lurk any human from judgement making process. Seen as becoming wider spread with the increase in computer based testing. For instance, some corporations provide computerized analysis of Minnesota Multiphasic Personality Inventory (MMPI-2) and other test results (Kaplan & Saccuzzo, 2009). Such technology tends to reduce test takers freedom and uniqueness. With high speed information communication technology (computers) and centralized data banks the probability that computers will someday provide important evaluation judgements about human lives is on the increase. Therefore, society must weigh the negative and positive benefits of growing application of modern information communication technology to the testing field. It is important people make the right decision early before an unpleasant but unalterable situation start to exist. With the emergence of computers in the field of psychological testing and its acceptance by psychologist in analyzing and storing of tests results, it may be very difficult to alter this new innovation or reverse the trend especially with the evolution of industry 4.0, where machines and objects as well as information communication technology are employed for dynamic management of complex processes because of its flexibility. Thus, anything that will threaten the use of computers in psychological testing and interpretation of tests results must be evaluated and the benefits sound to greatly outweigh the risks can the decision be accepted socially.

Usefulness of Tests

The important issue in testing is not whether the tests are perfect but whether they are useful to the individual, or the society. Tests need not be perfect in every area. Society often finds users for initial rough or simple instruments that are become precise with research and development (McKnow, 2007; Meyer et al, 2003; Sawyer, 2007). For instance, scientist believed that the sun revolved around the earth, the available methods and the principles were useful in that they led to some precise predictions, even though the theories beneath were incorrect. In like manner, the assumptions beneath today's test may be fundamentally incorrect and the resulting test instruments far from perfect. The test however, may still be useful as long as they provide data that leads to better predictions and understanding that can otherwise be obtained. However, as new knowledge emerges, society must continually weigh the disadvantages of tests against the advantages. The disadvantages of tests include; possible misinterpretations which may adversely affect the life of an individual or may systematically discriminate against a specific cultural group or setting (Fish, 2002; Henry, Bryson, & Henry, 1990). The benefits of tests include; the potential for increased precision and fairness in decision making process. Predictably, the resolution of this recurring issue will profoundly affect the field of psychological testing.

Current Fashions in Psychological Testing

Among the current fashions or issues in psychological testing is the development of new tests (higher standards, improved technology, and objectivity), increased in public awareness and influence, and computer and internet application.

The Development of new Tests

Studies have shown that hundreds of new tests are being published each year. The impetus for developing these new tests comes from professional disagreement over the best strategies for measuring human behaviour, the nature of these behaviours, and theories of these human characteristics. An example is the 2004 modern version of the Kaufman Assessment Battery for Children (KABC-11); this is an individual ability test for children between 3 and 18 years of age. The test consist of 18 subsets combined into five global scales called sequential processing, simultaneous processing, learning, planning, and knowledge (Kaufman & Kaufman, 2004a). The impetus also stem from public and professional pressure to fair, accurate and unbiased instrument. If tests are used, developers o tests stand to benefit financially, then new tests will be developed and marketed in other to make more financial profit.

Studies have also indicated that majority of new tests developed are based on the sample principles and theories underlying the already established tests. Though most newly developed tests are justified on the grounds that they are either psychometrically superior to the existing tests or are more specific and indeed more appropriate for particular problems. However, some of the newer tests such as; the Fear Survey Schedule (FSS), the Differential Attitude Scale, Irrational Belief Test etc, are based on models, theories, and concepts that fundamentally differ from those that underlie traditional tests. These traditional tests stem from modern concepts and theories from learning, social, physiological, and experimental psychology (Wood, Nezworski, Lilienfield, & Garb, 2003).

Increase Public Awareness and Influence on Testing

Increase public awareness of the nature and usefulness of tests has led to increasing external influence on testing. Before this time, the public had little or knowledge about psychological tests. Today, there is wide spread awareness among the general public on the need and importance of psychological tests and other forms of test. This is as a result of increase in information communication technology, and the importance attached to life and living that has led to increase demand for psychological services including testing services. The greatest benefit of increase public awareness of tests has been the extra focus on safeguarding human rights. As more individuals share the responsibility of encouraging the proper use of tests by becoming better informed of their rights and insisting on receiving their rights, the probability of misappropriation and abuse of tests will be reduced. However, the pressure of public interest in psychological tests has led test experts to an even greater awareness about safeguarding the rights and dignity of the testee (Shakow, Hilgard, Kelly, Stanford, & Shaffer, 1947).

Computer-Based Testing

One of the major trends in testing is the use of computers. Computers are being used in many different ways. For example, in adaptive computerized testing, different sets of test questions are administered through computer to different test takers, depending on each the trait being measured (Mills, Potenza, Fremer, & Ward, 2003; Weiss, 1983, 1985). Likewise in ability testing, the computer adjusts the level of item difficulty according to the test takers response. If the test takers answer is incorrect, then an easier item is given; if correct, then a more difficult item appears next. Computers are also being employed to administer, score, and interpret psychological tests. Even Tradio-Medical doctors nowadays use computers in testing their clients before prescribing drugs. Computers are also being used to generate tasks that cannot be presented by traditional methods such as the Strong-Campbell Interest Inventory, the Multidimensional Aptitude Battery, and other tests that measure anxiety, depression, and psychological reactance (Costa, De Carvalho, Drummond, Wauke, & De Guimaraes, 2002).

The Hope of New and Improved Tests

Psychologists believe that the dominant role of some of the popular test such as Stanford-Binet and Wechsler tests is far from secure. These two intelligence scales are probably technically adequate as they will ever remain. They can be improved through minor versions to update test stimuli and provide larger and even more representative normative samples with special norms for particular groups via additional research to extend and support validity evidence. However, despite the variations in the modern Binet and WAIS-III, the fundamental characteristics and the underlying concepts still take the format of the original scales. The probability that these two intelligence scales will be change in the future at least once or twice if not more by similar tests with superior standardization and normative information or less bias against certain minority is high. More so, the emergence of the Kaufman Assessment Battery for Children, Structured Personality Testing, and the MMPI-2 is already pioneering the 21st century.

Conclusion

All psychological tests are based on theories of human functioning. The validity of these theories and the underlying assumption is far from proven. More so, there seem to be no consensus or generally agreed assumption of the essence of human personality, normal or abnormal. With the increase in the awareness of tests users created by them for testing the need for improving the existing psychological test is necessary as some of the tests today may not be able to meet the psychological needs of individuals consisting the changes that take place in our body chemistry which sometimes may have some psychological implications on human personality or trait. As Kaufman Assessment Battery for Children, Structured Personality Testing, and the MMPI-2 is already pioneering the 21st century, Psychologist should be more creative in building new tests that will meet future testing needs of the fast growing population and be persistence in modifying the existing tests while accomplishing the goals of psychological testing.

References

- Abe, C.V. (2012). *Monitoring and evaluation: A panacea of quality assurance in higher education*. Ibadan: Stirling Holden Publishers Ltd.
- Kaplan, M.R., & Saccuzzo, D.P. (2009). *Psychological testing: Principles applications and issues*. Seventh edition. Wadsworth: engage learning.
- Camara, W.J. (1997). Use and consequences of assessment in the USA: Professional, ethical and legal issues. *European Journal of Psychological Association*, 13(2), 140-152.
- Thomas, J.C. (2013). Planning data collection and performing analysis. In J.C. Thomas & M. Hersen (Eds.), *Understanding research in Clinical and Counselling Psychology*. Mahwah, NJ: Erlbaum, 27-68.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.
- American Psychological Association (2002). *The ethical principles of psychologists and code of conduct*. Retrieved 19/10/2018 from <http://www.apa.org/ethics/code2002.html>.
- Tryon, W.W., & Benstein, D. (2003). Understanding measurement. In J.C. Thomas & M. Herson (Eds.). *Understanding research in clinical and counselling psychology*. Mahwah, MJ: Erlbaum, 67-68
- Cacioppo, J.T., Bernston, G.G., & Anderson, B.L. (1991). Physiological approaches to the evaluation of psychotherapeutic process and outcome, 1991: Contribution from social psychophysiology. *Psychological Assessment: Journal of Counselling and Clinical Psychology*, 3, 321-336
- Corigan, J.D., Boguner, J.A., Mysiw, W.J., Clinchot, D., & Fugate, L. (2001). Lie satisfaction after traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 16(6), 543-555.

- Spear, J.H. (2007). Prominent schools or other active specialties: A fresh look at some trends in psychology. *Review of General Psychology*, 11, 363-380.
- Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, 36, 117-128.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Ziskin, J. (1995). *Coping with psychiatric and psychological testimony* (5th ed.). Los Angeles. Law and Psychology Press.
- Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review*, 109(1), 116-136.
- Stembery, R.J. (2001). Successful intelligence: Understanding what Spearman had rather than what he studied. In J.M. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement*. Mahwah, NJ: Erlbaum.
- Zautra, A.J. (2003). *Emotions, stress, and health*. London: Oxford University Press.
- Shibre, T., Kebede, D., Alen, A., Negash, A., Deyasa, N., Fekada, D., Jacobson, L., & Kullgreen, C. (2003). Schizophrenia: Illness impact on family members in a traditional Society – Rural Ethiopia. *Social Psychiatric Epidemiology*, 38(1), 27-34.
- Szase, T.S. (1961). *The myth of mental illness*. New York: Harper & Row.
- Shaffer, T.W., Erdberg, P., & Horiaian, J. (1999). Current non-patient data for the Rorschach, WAIS-R, and MMPI-2. *Journal of personality Assessment*, 73, 305-316.
- McReynolds, C.J., Ward, D.M., & Singer, O. (2002). Stigma, discrimination, and invisibility: Factors affecting successful integration of individuals diagnosed with Schizophrenia. *Journal of Applied Rehabilitation Counseling*, 34(4), 32-39.
- Hamel, M., Shaffer, T.W., & Erdberg, P. (2000). A study of non-patient preadolescent Rorschach protocols. *Journal of Personality Assessments*, 75, 280-294.
- Dahlstrom, W.G. (1969). Recurrent issues in the development of MMP1. In J.N. Butcher (Ed.), *MMP1: Research developments and clinical applications*. New York: McGraw-Hill.
- Brayfield, A.H. (Ed.). (1965). Testing and public policy. *American Psychologist*, 20, 857-1005.
- Benjamin, G.A., & Gollan, J.K. (2003). Evidentiary standards and rules of evidence. In G.A. Benjamin & J.K. Gollan (Eds.), *Family evaluation in custody litigation: Reducing risks of ethical infractions and malpractice guide book (pp.17-28)*. Washington, D.C: American Psychological Association.
- Dana, R.H. (2000). *Handbook of cross-cultural and multicultural personality assessment*. Mahwah, NJ: Erlbaum.
- Franklin, R.D. (2003). *Prediction in forensic and neuropsychology: Sound statistical practices*. Mahwah, NJ: Erlbaum.
- Embretson, S.E., & Hershberg, S.L. (1998). *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: Erlbaum.
- Merrell, K.W. (2003). *Behavioural, social and emotional assessment of children and adolescents (2nd ed.)*, Mahwah, NJ: Erlbaum.
- Naar-King, S., Ellis, D.A., & Frey, M.A. (2003). *Assessing children's wellbeing: A handbook of measures*. Mahwah, NJ: Erlbaum.
- McKnow, C. (2007). Concurrent validity and clinical usefulness of several individual administered tests of children's social-emotional cognitive. *Journal of Clinical Child and Adolescent Psychology*, 36(1), 29-41.

- Meyer, G.I., Eyde, L.D., Kay, G.G., Moreland, K.L., Dies, R.R., Eisman, E.J., Kubiszyn, T.W., & Reed, G.M. (2003). Psychological testing and psychological assessment: *A review of evidence and issues*. In A.E. (Ed.), *Methodological issues and strategies in clinical research (3rd ed.)*. Washington, D.C: American Psychological Association, 265-435.
- Sawyer, R. (2007). Indicators of usefulness of test scores. *Applied Measurement in Education*, 20(3), 255-271.
- Fish, J.M. (2002). *Race and intelligence: separating science from myth*. Mahwah, NJ: Erlbaum.
- Henry, P., Bryson, S., & Henry, C.A. (1990). Black student attitudes towards standardized tests. *College Students Journal*, 23, 346-354.
- Mislevy, R.J. (2002). *Psychometric principles in students' assessment*. Loss Angeles, CA: Centre for the study of Evaluation National centre for research on Evaluation Standards and Students Testing Graduate School of Education & Information Studies University of California Los Angeles.
- Bootzin, R.R., & McKnight, P.E. (2006). *Strengthening research methodology: Psychological measurement and evaluation (1st ed.)*. Washington, D.C: American Psychological Association.
- Kaufman, A., & Kaufman, N. (2004a). *KABC-II manual*. Circle pines, MN: American Guidance service.
- Wood, J.M., Nezworski, M.T., Liliefield, S.O., & Garb, H.N. (2003). *What's wrong with the Rorschach? Science confronts the controversial inkblot test*. San Francisco: Jossey-Bass.
- Mills, C.N., Potenza, M.T., Framer, J.J., & Ward, C. (2002). *Computer based testing: Building the foundation for future assessments*. Mahwah, NJ: Erlbaum.
- Shakow, D., Hilgard, E.R., Kelly, E.L., Stanford, R.N., & Shaffer, L.F. (1947). Recommended graduate training in clinical psychology. *American Psychologist*, 2, 539-558.
- Weiss, D.J. (1983). *New horizons in testing*. New York: Academic Press.
- Weiss, D.J. (1985). Adaptive testing by computer. *Journal of Consulting and Clinical Psychology*, 53, 774-789.
- Costa, R.M.E., De Carvalho, L.A., Drummond, R., Wauke, A.P.T., & De Sa Guimaraes, M. (2002). The UFRJ-UERJ group: *International Virtual Reality Experiments in Neuropsychiatry, Cyber Psychology and Behaviour*, 5(5), 423-431.
- Wainer, I, B. (2003). *Principles of Rorschach interpretation (2nd ed.)*, Mahwah, N.J: Erlbaum.