

Ethical Standards for Language Testing Professionals: An Introduction to Five Major Codes

Yujie Jia
(Hong Kong Polytechnic University)

Abstract

This article investigates five major ethical codes in the language testing field. Whereas the two ILTA codes and the JLTA codes were found to be quite similar, the ALTE and EALTA codes were seen to differ in many regards. This article raises several questions about codes of ethics in general and prospects for their future development.

Keywords: professional codes of ethics, ethical work codes, testing codes of practice, professionalism in language testing

摘要

此文重点强调社会责任感对于语言测试行业的重要性。研究比较了语言测试领域 5 个主要的《考试道德规范》。研究表明国际语言测试组织和日本测试组织的非常相似，而 2 个欧洲语言测试组织的则大不相同。最后针对道德规范整体上以及将来规范的制定提出了几个问题。

关键词：职业道德规范；道德规范书；测试规范；语言测试职业化

Since the 1990's one of the important trends in the language assessment field has been to give more attention to the ethical practices behind language testing. Buchholz considers professional codes as 'the enlivening spirit of a profession' (as cited in Williamson, 1993, p. 20). When confronted by ethical dilemmas, professionals can generally find guidance in their official codes. Brill also states, 'Code of ethics are windows into a profession' (2001, p. 223). In other words, they provide insights into what the members of that profession consider important – and indirectly as to what they regard as non-essential.

As Hook and White (2001, p. 2) state, "a code of ethics functions as a tool for professional self-definition." In other words, in the process of specifying what persons in a given profession should and should not do, those individuals are also socially defined. In this respect, ethical codification can be viewed as a means of formulating social self-identity, and subsequently projecting that image to the public at large. Hence the process of codification involves more than the mere exposition of duties and rights within a profession - it is also linked to social identity formation and role projection.

Though language testing is not on the same foothold as medical science or law in terms of social status, it can be considered a profession in many regards. One sign of its growing professional status is the emergence of codes of ethics.

Thrasher (2001, p.14) considers a code of practice to be essential for language testers in providing professional guidance about test quality. This article mentions others who have also stressed the importance of ethical codes for our profession.

"the process of codification involves more than the mere exposition of duties and rights within a profession - it is also linked to social identity formation and role projection."

Codes of Ethics: Towards a Definition

What constitutes an ethical code (or code of ethics)? Davies (2005, p. 46) states that an ethical code is "a public written statement by an organization of its principled approach to its professional and/or business activities". In other words, ethical codes refer to "a set of social practices that has a purpose, namely, the promotion of the common welfare" (Davies, 2004, p. 98).

According to Fortin (1996, p. 1720) successful codes of ethics are not idealized descriptions of behavior – they are evolving codifications of baseline standards in a given field. Fortin (1996, p. 1715) adds:

The first trap to be avoided is that of angelically high standards. If the standards are set too high, the goal of any code of ethics, which is to promote acceptable moral conduct . . . will not be achieved.

Fortin also points out that codes of ethics, by themselves, do not create virtue: what they can do is to increase the likelihood of *responsible behavior* mainly by “appealing to the moral sense of those whose activities they govern” (1996, p. 1720). For this to occur, there needs to be an ongoing process of engagement with a given code and a basic acceptance of the ideas behind it. Longstaff (1994, par. 51) echoes this point by stating codes of ethics must be internalized by their practitioners to be effective. For that reason,

“successful codes also need to be simple enough so that they can actually be remembered and universal enough so that different persons from varied cultures in a given profession can agree with them.”

successful codes also need to be simple enough so that they can actually be remembered and universal enough so that different persons from varied cultures in a given profession can agree with them.

Although consensual agreement among practitioners is the ideal which codes should strive for, Darnay (2001, par. 13) suggests that we should not be naïve enough to assume that all persons in a profession will be motivated to follow a given code. For this reason, many professional codes contain sanctions for violations, ranging from formal reprimands to dismissal. However, even if practitioners in a given profession wish to adhere to a code, some face working conditions that make it difficult (or impossible) to fulfill all obligations specified in a code. For example, some nurses face huge workloads that hinder the fulfillment of even basic points in their code (Hook & White, 2001, p. 11). Some EFL teachers are also asked to create some in-house entrance exam without being given the time or training to do so properly. In such cases, fulfillment of the code should not be solely regarded as a matter of individual responsibility – part of the responsibility also rests with the broader systematic conditions in which practitioners are placed.

For reasons such as this, many codes typically seek to embody a dynamic balance competing discourses, some of which are listed in Table 1.

Table 1. A partial list of the competing discourses expressed in many professional ethical codes.

idealized norms	vs.	actualized behaviors
internal appeals	vs.	external sanctions
practitioner rights	vs.	practitioner responsibilities
voluntary “oughts”	vs.	mandatory “musts”
individual conscience	vs.	universal norms
client autonomy	vs.	agency control
abstract values (“whys”)	vs.	concrete specifications (“whats”)

Codes have a valuable function in shaping behaviors, yet they should also be flexible enough to recognize that many test developers operate under diverse (and often less than optimal) conditions.

An Analysis of Five Codes

The current study compares five professional language testing codes: (1) *The Code of Ethics* and (2) *The Guidelines for Practice* of the International Language Testing Association, (3) *The Principles of Good Practice* of the Association of Language Testers in Europe, (4) *The Guidelines for Good Practice in Language Testing and Assessment* of the European Association for Language Testing and Assessment, and (5) *The Code of Good Testing Practice* of the Japanese Language Testing Association.

In this paper, the terms “guidelines”, “principles” and “codes of practice” are used interchangeably. Table 2 presents the key points of the five codes, after which their similarities and differences are discussed individually.

Table 2. A comparison of five different codes of ethics for testing as of March 2009.

Code	ILTA Code of Ethics	ILTA Guidelines for Practice	ALTE Principles of Good Practice for ALTE Examinations	EALTA Guidelines for Good Practice in Language Testing & Assessment	JLTA Code of Good Testing Practice
Salient points	9 principles & annotated gloss expressing altruistic aims with ample loopholes	10 duties of test designers/writers & 6 points about test administration & other concerns -10 rights and 10 responsibilities of test takers are also included	Detailed guidelines on how to ensure validity, reliability, positive impact, practicality, and quality of service for all stakeholders	9 queries about testing & assessment education plus 31 considerations in developing tests	Very similar to the ILTA Guidelines with an extra proviso that 'Items written by non-native speakers of the language being tested must be checked by competent native speakers'
Most recent version & current status	Adopted in March 2000	Drafted in July, 2005 and adopted in June 2007	Drafted in 1991-93 and adopted in 1994, then revised in October 2001	Adopted in May 2006	Proposed in September 2002 and adopted in 2006

(1) *The ILTA Code of Ethics*

The ILTA Code of Ethics comprises a prelude and nine principles elaborated through a series of annotations. These principles draw upon moral philosophy and serve to promote good professional conduct. They also offer a benchmark of satisfactory ethical behavior by language testers (ILTA, 2000). This code is closely associated with the ILTA Guidelines for Practice (ILTA, 2007).

A product of compromise, different parts of the code can be interpreted in various ways to justify practices that might seem contradictory. For example, concerning confidentiality, the code states, 'a careful balance must be maintained between preserving confidentiality as a fundamental aspect of the language tester's professional duty and the wider responsibility the tester has to society' (ILTA, 2000, p.2). Based on this wording, examiners might feel justified in compromising client confidentiality if it seemed to benefit society as a whole. Since many ethical questions involve weighing competing claims, and far-reaching decisions often come down to difficult choices in limited time frames, it should be no surprise if decisions made by language testers sometimes seem gray. In this vein Mackay, as cited by Ross (1994, par. 20), remarked:

The fundamental point about ethical behaviour is that it willingly takes the rights, the needs and welfare of others into account. It is not behavior which simply responds to the pressure of law, regulation or code: it is behaviour which synthesizes a number of conflicting pressures and competing claims, and never quite feels neat and tidy. Black and white judgments rarely emerge from sensitive ethical debate.

Peters (2008, par. 6) has also made a significant point when stating that even if ethical benchmarks (which codes generally attempt to embody) are not realized perfectly, such codes are still valuable because they offer guidance based on historical precedent. Since language tests have been used for questionable practices such as ethnic cleansing at various points in history (McNamara & Roever, 2007, pp. 150-164), ethical test purpose needs to be at the forefront of test designers and administrators' minds.

Another part of the ILTA Code which might cause confusion is the conflict between Principles 8 and 9. Principle 8 requires language testers to recognize their obligation to the testing organization for which they work, even when they may not themselves agree with them, whereas Principle 9 states that language testers have the right to withhold their professional services on the grounds of conscience (ILTA, 2000, pp. 5-6).

These criticisms should not undermine the many positive points in the ILTA Code. For instance, the Code of Ethics stresses the importance of fostering scholarly interactions in the field. Principle 5 also requires language testers to 'develop their professional knowledge, sharing this knowledge with colleagues and other language professionals' (ILTA, 2000, p. 4).

(2) *The ILTA Guidelines for Practice*

Whereas the ILTA Code of Ethics outlines general ethical ideals, its Guidelines of Practice are more precise in specifying minimum standards of conduct and detailing what behaviors are regarded as unprofessional (ITLA, 2000, p. 1). The ILTA Guidelines for Practice spell out some basic considerations for good testing in all situations as well as on some special occasions. They mention the responsibilities and rights for the following stakeholder groups: test designers, test writers, institutions preparing or administering high stakes examinations, those preparing and administering publicly available tests, test users and test takers. The fundamental rationale for the ITLA Guidelines for Practice is to promote test validity, reliability, and test fairness.

Part 1, Section C of the ILTA Guidelines contends that items written by non-native speakers of the language being tested must be checked by someone with a high level of competence in the language (2007, p. 3). They do not specify that items written by native speakers need to be similarly checked. This dual standard is an issue which should be raised and tactfully resolved in a later revision of the ILTA code.

"Rather than envisioning ethical codes as tablets carved in stone, it might be more useful to think of them as invitations to become more critically engaged with key issues."

Rather than envisioning ethical codes as tablets carved in stone, it might be more useful to think of them as invitations to become more critically engaged with key issues. As Ross (2000, par. 12) comments, codes of ethics must be more than carefully preserved archival records – they should be "living documents" with which members of a profession are intimately acquainted. They also periodically reviewed to make sure that the "goodness of fit" between what they espouse accords well with what most members of the profession currently believe in.

Basically speaking, the ILTA Guidelines for Practice express many altruistic points. For instance, they state that test takers have the right to present concerns about the testing process or their results and to receive information about procedures that will be used to address such concerns. The ILTA Guidelines take into account various stakeholder groups, suggesting that test design is a complex social process of balancing opposing interests.

(3) *The ALTE Principles of Good Practice for ALTE Examinations*

Unlike the ILTA's Code of Practice, which mainly presents a set of professional standards for stakeholders, the ALTE Principles of Good Practice for ALTE Examinations focus on technical standards for test construction and evaluation. This document sets out in considerable detail the principles that ALTE members should adopt in order to achieve high professional standards.

One good point about the ALTE Principles of Good Practice is that the first stage of a rational approach to examination development, administration and validation must involve planning to identify the considerations and constraints which will be relevant to the examination development project and which determine how examination usefulness will be achieved. However, they can also be regarded as little more than a cursory review of some of the important concepts or principles in the language testing field.

(4) *The EALTA Guidelines for Good Practice*

The EALTA Guidelines for Good Practice in Language Testing and Assessment recommend assessment literacy training for teachers as well as maintaining sound standards for test development in national or institutional testing agencies. EALTA claims that many general principles also apply for different audiences: respect for the

students/examinees, responsibility, fairness, reliability, validity and collaboration among the parties involved. Unlike the other codes discussed in this article, the EALTA Guidelines for Good Practice are set down in the form of questions. For example, Article (A), Section 2 offers this rhetorical inquiry:

How aware are trainees made of the range of assessment procedures appropriate to their present or future needs?

Such queries have the advantage of raising language testers' awareness without coercive innuendos. However, it could be argued that there are in fact no explicit guidelines - merely a range of questions to consider.

(5) The JLTA Code of Good Testing Practice

The JLTA Code of Good Testing Practice covers six themes: 1) basic considerations for good testing practice in all situations, 2) responsibilities of test designers and test writers, 3) obligations of institutions preparing or administering high-stakes exams, 4) obligations of those preparing and administering commercially available exams, 5) responsibilities of users of test results, and 6) special considerations.

The JLTA Code of Good Testing Practice and ILTA Guidelines for Practice are one and the same except that the JLTA Code doesn't incorporate the rights and responsibilities of test takers. One other notable point is that the JLTA document indicates items written by non-native speakers of the language being tested must be checked by 'competent native speakers', whereas ILTA refers to 'someone with a high level of competence in the language'. I believe there is enough evidence to suggest that competency is not an exclusive domain of native speakers, and not all native speakers have the skills needed for test editing.

Conclusion

Several points stand out from this preliminary analysis. Firstly, the comparison of the codes of practice for the four associations mentioned here reveals that (1) All the codes except the EALTA's consist of a set of explicit expectations by which language testers can potentially be judged. The EALTA Guidelines are perhaps best described as a list of inquiries to reflect on when producing tests. (2) Each code attempts to improve the quality of language testing practices and assessment procedures and improve test fairness. (3) They mirror widely accepted beliefs about validity, reliability, washback and test fairness in the language testing community. Table 3 illustrates some of the salient features of the major codes discussed in this paper.

Table 3. *A comparison of some of the features of different language testing codes as of 2009.*

Test Characteristics	ILTA Code of Ethics	ILTA Guidelines for Practice	ALTE Principles of Good Practice for ALTE Examinations	EALTA Guidelines for Good Practice in Language Testing & Assessment	JLTA Code of Good Testing Practice
Concern for client confidentiality	O	O	Δ	X	O
Native speaker checking of tests	X	X	X	X	O
Concern for ongoing education & professional development	O	O	O	X	O
Concern for responsibilities & rights of test takers	O	O	X	X	O
Concern about examinee cheating	X	X	X	X	X
Mention of sanctions	O	X	X	X	X

SYMBOLS: O = a feature of that code Δ = a partial feature of that code X = not a feature of that code

Secondly, these findings also raise questions about the appropriacy of standardizing codes of ethics in language assessment. Such standardization could yield consistency, yet practically speaking, it would be a daunting task to keep track of the way language tests are designed across the globe. We should remember that the vast majority of foreign language test developers are ordinary teachers with little, if any, specialized test design training. Although it might be interesting to consider the possibility of having specified test brands labeled as "ILTA Code compliant", such proposals are beyond the scope of this paper. Nonetheless, when reflecting on code standardization we must consider what (if any) economic incentives code standardization might have.

A third point is that all these codes devote little attention to sanctions and penalties. Only the ILTA Code of Ethics suggests any penalties. It states that anyone who fails to uphold the Code may lose his or her ILTA membership on the advice of the ILTA Ethics committee. Since the ILTA is a voluntary organization with only 243 members as of April 2009, whether or not this is a "serious penalty" is open to question. The absence of sanctions may tempt some people to believe that codes of ethics are toothless (Wiley, 2000, p. 112). Thus if professional ethical codes are to be truly effective, my belief is that sanctions are necessary.

Focusing on professional codes in the language testing field, this article has introduced the professional codes of conduct of four associations. Future studies should offer a more systematic analysis of each code from economic, structuralist, and sociological perspectives. This article has also made these specific suggestions regarding professional language testing codes:

- (1) That any dichotomy between native and non-native speakers be avoided in the future wording of any codes.
- (2) That language testing codes specify at least some form of sanctions.
- (3) That language testing codes become more standardized.
- (4) That when people join a language testing association, they receive a copy of its code. As it stands, people can join some language testing organizations without being informed of the codes existence or being aware that there is any responsibility to adhere to the specific code.

It is hoped that this study will promote more interest in professional codes of ethics and encourage more research in this area.

Acknowledgement

The original version of this paper is from a guided study I did with Prof. Alan Davies in 2008. I am very appreciative of his support and encouragement.

References

- Association of Language Testers in Europe. (2001). *Principles of good practice for ALTE examinations*. Retrieved March 30, 2009 from <http://www.alte.org/cop/principles.php>
- Brill, C.K. (2001). Looking at the social work profession through the eye of the NASW Code of Ethics. *Research on Social Work Practice, 11* (2), 223-234.
- Buchholz, W. J. (1989). Deciphering professional codes of ethics. *IEEE Transactions on Professional Communication, 32* (2), 62-68.
- Davies, A. (1997a). Demands of being professional in language testing. *Language Testing, 14* (3), 328-339.
- Davies, A. (1997b). Introduction: The limits of ethics in language testing. *Language Testing, 14* (3), 235-241.
- Davies, A. (2004). Introduction: Language testing and the golden rule. *Language Assessment Quarterly, 1* (2&3), 97-107.
- Davies, A. (2005). *A glossary of applied linguistics*. Mahwah, N.J.: Lawrence Erlbaum Associates.

- European Association for Language Testing and Assessment. (2006). *EALTA Guidelines for good practice in language testing and assessment*. Retrieved on March 30, 2009 from <http://www.ealta.eu.org/documents/archive/guidelines/English.pdf>
- Fortin, P. (1996, May 8). Address to a Special Joint Committee on a Code of Conduct of the Canadian Parliament. Retrieved on March 30, 2009 from http://www.parl.gc.ca/35/Archives/committees352/sjcc/evidence/o3_96-05-08/sjcc03_blk101.html
- Hook, K. G. & White, G. B. (2001). Code of ethics for nurses with interpretive statements: An independent study module. Retrieved on May 1, 2009 from nursingworld.org/mods/mod580/code.pdf
- International Language Testing Association. (2000). *Code of ethics for ILTA*. Retrieved on March 30, 2009 from <http://www.iltaonline.com/code.pdf>
- International Language Testing Association. (2007). *ILTA guidelines for practice*. Retrieved on March 30, 2009 from http://iltaonline.com/index.php?option=com_content&task=view&id=122&Itemid=133
- Japan Language Testing Association. (2007). *The JLTA Code of Good Testing Practice*. Retrieved on March 23, 2009 from <http://www.avis.ne.jp/~youichi/COP.html>
- Longstaff, S. (1994). Developing Codes of Ethics: Why codes of ethics fail. Retrieved on April 30, 2009 from <http://www.ethics.org.au/about-ethics/ethics-centre-articles/ethics-subjects/codes-of-ethics/article0124a.html>
- McNamara, T. & Roever, C. (2007). *Language Testing: The Social Dimension*. (Language Learning Monograph Series) Malden, MA & Oxford, UK: Blackwell Publishing.
- Newfields, T. (2001). Voices in the field: An interview with Randy Thrasher. *SHIKEN: The JALT Testing & Evaluation SIG Newsletter* 5 (1), 12-14. Retrieved on March 23, 2009 from http://jalt.org/test/thr_new.htm
- Peter, E. (2008). Seeing our way through the responsibility-vs-endeavour conundrum: The code of ethics as both product and process. *Nursing Leadership (CJNL)*, 21 (2) 28-33. Retrieved on April 29, 2009 from <http://www.longwoods.com/product.php?productid=19872>
- Wiley, C. (2000). Ethical standards for human resource management professionals: A comparative analysis of five major codes. *Journal of Business Codes*, 25, 93-114.
- Williamson, W.J. (1993). *Ethics of Technical Communication: An Annotated Survey of Scholarship, 1980-1993*. Retrieved on April 10, 2009 from http://www.uni.edu/~drbill/professional/teaching/dr/MS_Project_1994.pdf