American Translators Association Scholarly Monograph Series (ATA)

As of 1993 John Benjamins has been the official publisher of the ATA Scholarly Monograph Series. Edited by Françoise Massardier-Kenney, under the auspices of the American Translators Association, this series has an international scope and addresses research and professional issues in the translation community worldwide. These accessible collections of scholarly articles range from issues of training, business environments, to case studies or aspects of specialized translation relevant to translators, translator trainers, and translation researchers.

Volume XIV

Testing and Assessment in Translation and Interpreting Studies. A call for dialogue between research and practice
Edited by Claudia V. Angelelli and Holly E. Jacobson

Testing and Assessment in Translation and Interpreting Studies

A call for dialogue between research and practice

Edited by

Claudia V. Angelelli
San Diego State University

Holly E. Jacobson
University of Texas at El Paso

John Benjamins Publishing Company

Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48–1984.

Library of Congress Cataloging-in-Publication Data

Testing and assessment in translation and interpreting studies : a call for dialogue between research and practice / edited by Claudia V. Angelelli, Holly E. Jacobson.
p. cm. (American Translators Association Scholarly Monograph Series, ISSN 0890-4111 ; v. XIV)
Includes bibliographical references and index.
1. Translating and interpreting--Evaluation. I. Angelelli, Claudia (Claudia V.) II. Jacobson, Holly E.
P306.2.T495 2009
418'.02--dc22 2009025926
ISBN 978 90 272 3190 1 (HB; alk. paper)
ISBN 978 90 272 8902 5 (EB)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Introduction: Testing and assessment in translation and interpreting studies: A call for dialogue between research and practice 1
Claudia V. Angelelli and Holly E. Jacobson

Part I. The development of assessment instruments: Theoretical applications

Using a rubric to assess translation ability: Defining the construct 13
Claudia V. Angelelli

Moving beyond words in assessing mediated interaction: Measuring interactional competence in healthcare settings 49
Holly E. Jacobson

Part II. The development of assessment instruments: Empirical approaches

The perks of norm-referenced translation evaluation 73
June Eyckmans, Philippe Anckaert and Winibert Segers

Revisiting Carroll's scales 95
Elisabet Tiselius

Meaning-oriented assessment of translations: SFL and its application to formative assessment 123
Mira Kim

Assessing cohesion: Developing assessment tools on the basis of comparable corpora 159
Brian James Baer and Tatyana Bystrova-McIntyre

Assessing software localization: For a valid approach 185
Keiran Dunne

Part III. Professional certification: Lessons from case studies

The predictive validity of admission tests for conference interpreting courses in Europe: A case study 225
Šárka Timarová and Harry Ungoed-Thomas

Getting it right from the start:
Program admission testing of signed language interpreters 247
Karen Bontempo and Jemina Napier

Standards as critical success factors in assessment:
Certifying social interpreters in Flanders, Belgium 297
Hildegard Vermeiren, Jan Van Gucht and Leentje De Bontridder

Assessing ASL-English interpreters:
The Canadian model of national certification 331
Debra Russell and Karen Malcolm

Author index 377

Subject index 381

INTRODUCTION

Testing and assessment in translation and interpreting studies

A call for dialogue between research and practice

Claudia V. Angelelli and Holly E. Jacobson

Translation and interpreting are areas of inquiry supported by substantial scholarship. The notion of quality is central to both fields, whether at initial acquisition levels as a formative assessment in educational programs, or at more advanced levels in developing instruments for professional certification, as well as in measuring the quality of translation/interpreting for instruments and processes used for research purposes. Assessment and testing in the two fields is implemented for a number of purposes. Examples include screening applicants for entry into an educational program; providing feedback to students taking a course; testing knowledge and skills at the end of a course of study; carrying out quality assessments in contexts where interpreters play an essential role in achieving interactional goals; certifying professional-level competence in translation or interpreting; determining quality of localization products in the industry, as well as measuring the impact of surveys and other instruments translated for research purposes.

Most of the discussions around theory have focused on quality in theoretical terms, particularly in translation studies. Many of the established theoretical frameworks referred to in the translation literature are based on dichotomies or continua that distinguish between translations that closely adhere to the original linguistic code and more liberal translations that achieve a structure that is less subservient to that of the source text. Nida's (1964) concepts of formal and dynamic equivalence represent one of the first approaches to defining translation quality. His framework calls for determining quality according to the response a translation produces in target readers; that is, the response to the translation by target readers should be equivalent to the response to the original by source-text readers. In turn, Newmark (1982) uses the terms *semantic* and *communicative* translation to refer to a dichotomy that is similar to Nida's formal and dynamic equivalence. Likewise, Toury (1995) established a framework to refer to two types of translations, using *adequacy* to refer to a translation that closely adheres to

Using a rubric to assess translation ability

Defining the construct

Claudia V. Angelelli
San Diego State University

One of the first and most important steps in designing an assessment instrument is the definition of the construct. A construct consists of a clearly spelled out definition of exactly what a test designer understands to be involved in a given skill or ability. This task not only involves naming the ability, knowledge, or behavior that is being assessed but also involves breaking that knowledge, ability or behavior into the elements that formulate a construct (Fulcher 2003) and can be captured and measured by a rubric. Currently, there is no one definition of translation competence and its components that is universally accepted within the academic field of translation studies (Arango-Keith & Koby 2003). Neither is there a rubric that can capture different levels of competency in translation. Instead, there is a continuing debate about how to define translation competence and exactly how its constituent elements are to be conceptualized, broken down, interconnected and measured. This paper reviews the literature from Translation Studies, Testing and Second Language Acquisition and proposes sub-components of a rubric to assess the construct of translation competence.

Introduction

Translation has been characterized as both a process and a product (Cao 1996), more pointedly a very complex process and product. The fact that translation is a multi-dimensional and complex phenomenon may explain why there have been few attempts to validly and reliably measure translation competence/ability. This is evident when comparing the research produced in translation testing with that produced in testing in related fields.

Translation shares some of the same linguistic concerns, such as discourse and grammatical competence in two languages (to name only a few), as the field of Second Language Acquisition. Translation also involves a variety of skills, including analytical skills and strategic skills, which are also present in other fields

such as of Mathematics and others. When comparing the research produced in assessment within the field of Second Language Acquisition or in Mathematics, it is evident that we have not witnessed similar progress in assessment in the field of Translation and Interpreting Studies. This should not be interpreted as if the testing of translation or interpreting were not important enough or interesting enough to be worth the effort. On the contrary, developing a valid and reliable test for translation and interpreting is of paramount importance. Both academe and the industry would benefit enormously from making accurate and sound decisions on translation ability and quality based on meaningful testing.

Valid and reliable procedures for measuring translation or interpreting (or any other construct for that matter) start by posing essential questions about the procedures (Cohen 1994: 6) such as: for whom the test is written, what exactly the test measures, who receives the results of the test, how results are used, etc. Testing for both translation and interpreting share some similarities, specifically in the application of basic principles of measurement. But, because of the differences between these two the remainder of the discussion will focus solely on translation.

The answers to the questions about test procedure guide the test development and cannot be an afterthought. Test development also starts with a clear definition of what is to be measured, i.e. the test construct. Based on a case study of a professional organization certifying translators, this chapter starts by reviewing some of the relevant questions for translation test development. It uses the lens of testing and assessment to investigate the construct of "translation ability" and the potential use of rubrics to measure this construct. In so doing, it reviews how translation competence/ability has been defined by previous research and by professional ideology. Subsequently, it offers a working definition of the construct of translation ability for the specific purpose of translation certification. It argues for the use of rubrics to assess the translation ability of individuals seeking certification. It presents a rubric as a work in progress in the hope of contributing to relevant international discussions on valid and meaningful translation assessment.

1. Initial considerations

In this section I review the key questions in the assessment of translation as they apply to high-stake tests, such as translation certification examinations. The decision-making process of test developers' as to what to assess must be grounded in theory. For the purposes of translation assessment, I am suggesting that conceptualizations of communicative translation (Colina 2003) based on Skopos theory (Nord 1991 and 1997) be broadened to include concepts from cross-cultural communication and communicative competence theories (Bachman 1990; Hymes 1974; Johnson 2001) to allow for decisions regarding what to assess based on broader principles applicable to translation.

1.1 Questions preceding test development

When test developers begin the process of creating a test, they are guided by the following questions (Cohen 1994: 11–48):

- *What* aspects of an individual's translation ability should be assessed?
- *Why* are certain techniques, assessment methods or approaches being used instead of others?
- *How* will the assessment instruments (translation tests) be developed, and how are they going to be validated?
- *When* will the test take place, and how often is the professional organization planning to administer it?
- *Where* will the exam take place and what is the physical environment(s) of the exam?
- *Who* is the intended audience for the test? What information is available about social backgrounds cognitive skills and personal characteristics (diverse or similar) of target audience?
- *For whom* are the results on the translation test intended; for candidates themselves or for organizations which make the exam a requirement?

So far I have presented relevant questions that pertain to the *Wh*-group. Outside *wh*-questions there are other important questions that test developers need to answer as they embark on the test-development process.

1.2 Nature of the test

Among further relevant questions there are those concerning the nature of the test to be developed. Is the test a norm-referenced or a criterion referenced one? This distinction is important since it allows for different things. "A norm-referenced assessment provides a broad indication of a relative standing, while criterion-referenced assessment produces information that is more descriptive and addresses absolute decisions with respect to the goal" (Cohen 1994: 25). The norm-referenced approach allows for an overall estimate of the ability relative to the other examinees. Norm-referenced tests are normed using a group of examinees (e.g. professional translators with X amount of years of experience, or translators who have graduated from translation programs 6 months before taking the

test, etc.). In the criterion-referenced approach, criterion skills or behaviors are determined and then test specifications are written. This approach is used to see if a test taker has met certain objectives or criteria rather than to see how a test taker does compared to another. In addition to the nature of the test, whether is is a criterion-referenced or norm-references, other relevant questions test developers ask are about validity and reliability of the assessment instrument.

1.3 Validity

Traditionally, validity has been present in discussions on testing and test developers have raised questions such as: Is the test measuring what it is supposed to measure? (Kelly 1927; Lado 1961; & Cronbach 1971 in Weir 2005). Additionally validity has been discussed in different types, such as construct validity, content validity and face validity, among others (Bachman 1990; Bachman & Palmer 1996; Fulcher 2003; Messick 1989). As validity is multifaceted and multi-componential (Weir 2005), different types of evidence are needed to support any kind of claims for the validity of scores on a test. In 1985 the American Psychological Association defined validity as "the appropriateness, meaningfulness and usefulness of the specific inferences made from test scores (in Bachman 1990: 243). Therefore, test validity is not to be considered in isolation, as a property that can only be attributed to the test (or test design), nor as an all-or-none, but rather it is immediately linked to the inferences that are made on the basis of test scores (Weir 2005).

As an example, let's consider construct validity. This category is used in testing to examine the extent to which test users can make statements and inferences about a test taker's abilities based on the test results. Bachman and Palmer (1996) suggest doing a logical analysis of a testing instrument's construct validity by looking at the clarity and appropriateness of the test construct, the ways that the test tasks do and do not test that construct, and by examining possible areas of bias in the test tasks themselves. Construct validity relates to scoring and test tasks. Scoring interacts with the construct validity of a testing instrument in two primary ways. Firstly, it is important that the methods of scoring reflect the range of abilities that are represented in the definition of competency in the test construct. Similarly, it is important to ask if the scores generated truly reflect the measure of the competency described in the construct. Both of these questions essentially are concerned with whether or not test scores truly reflect what the test developers intended them to reflect.

The construct validity of a testing instrument can be affected by many of the same factors of the testing situation and the test tasks which create problems in reliability (see below). Therefore, it is important that all aspects of the testing situation and the test itself be considered for possible sources of bias for or against certain candidates. Bias here refers to any factor that may affect test performance that is not a part of the test's construct or testing objectives. The test designers must assure themselves that everything is done to make sure that the setting, instructions, input and expected responses do not end up influencing the test scores. Ideally, the only thing that should influence test scores is the candidate's competence, or lack thereof, as defined by the test's construct. With the question on validity comes the question of reliability.

1.4 Reliability

Reliability is one of the terms commonly used in the assessment field and is also fairly well-known outside of the field itself. It is not uncommon to hear lay persons using the word *reliability* or *reliable* and discussing what they judge reliability to be on a given issue or how reliable something is (e.g. a car or a brand). However, in the field of assessment and testing, the word *reliability* has specific meanings and set methods for its measurement. Primarily, *reliability* is used as a technical term to describe the amount of consistency of test measurement (Bachman 1990; Cohen 1994; Bachman & Palmer 1996) in a given construct. One way of judging reliability is by examining the consistency of test scores. If a person is given the same test at different times, will he or she score more or less the same? If different graders score the same examinee's test, will their scores be similar? If a person takes different versions of the same test, will they have similar scores on both versions? These and other questions reflect aspects of the consistency in test scores which test developers and testing experts are looking at when they examine the reliability of a test. However, reliability is not just about test scores. Creating a reliable test and judging the reliability of an existing test involves looking at the ways in which the consequences of factors outside of what is actually being tested, have been minimized to the greatest extent possible (Bachman & Palmer 1996). Factors of test administration and scoring procedures can be evaluated on the basis of how they might influence the reliability of the test in light of current thinking in the field of testing.

To determine reliability, we can use the questions for making a logical evaluation of reliability as set forth in Bachman & Palmer (1996). The factors impacting reliability are: (1) variation in test administration settings, (2) variations in test rubrics (scoring tool for subjective assessment), (3) variations in test input, (4) variation in expected response, and, (5) variation in the relationship between input and response types.

Both variations in the test setting and the physical conditions under which a test is administered can become possible sources of problems in reliability (Bachman & Palmer 1996). Due to the nature of national or international organizations' site selections, there will inevitably be some variation in test settings. Some test administrations may take place in a quiet office building. Others may take place in a room just off of a busy conference room. There will be natural variations in lighting, noise levels, temperature, and workspace available to a candidate, etcetera. These variations in settings can impact the individual candidate's ability to perform the test task or exercise. Some of these variations are unavoidable. In general, organizations should set minimum guidelines for the settings in which the examination is to be administered to minimize possible variations in performance on the examination due to environmental factors. It is also advisable to have test proctors go over a checklist of environmental factors and note any variance from the "ideal setting" so that these factors may be considered in the candidate's performance. Such data will also help the organization to analyze what, if any, of these factors play a role in variations found among the administration of the test.

When test developers or researchers are focusing on variations in the test protocol, the main question is whether or not there are variations in the way that instructions are given, in the time that is allotted for completion, or in the ways in which the scoring is done, which can influence the scores generated in an unanticipated or undesired way.

Variations in the wording of instructions are sometimes unavoidable for all versions of a certification exam, particularly if they are not all written in the same language. Therefore the question of consistency of test instructions across languages should be posed. At times, this may become a translation problem in and of itself. If the instructions are in fact consistent across languages then there would probably be no threats to the reliability stemming from this aspect of the test rubric. If, however, there is variation in the language and/or phrasing of the instructions on separate versions of the test, there is a possibility of some threat to reliability. Further study of the actual instructions would be needed in order to evaluate these variations for possible threats to reliability.

On the issue of time allotted for the test, care must be taken so that there are no variations in the time allotted due to variation in the performance of test proctors. Therefore, careful proctor training and the use of checklists by proctors are important elements in preventing threats to reliability.

Another area that may affect reliability is the manner in which the candidates' responses are scored. Scoring is ideally done based on fixed and objective criteria. Each instance of scoring by a grader should be similar to other instances of scoring that the same grader performs. This quality is known as *intra-rater reliability*.

Consultation among graders threatens the fixed and objective nature of scoring by threatening the *inter-rater reliability* (i.e. the fact that the same test, using the same scoring criteria and graded by different graders should yield similar results). Graders can pull each other one way or another. This is commonly known as grading by consensus. The most common practice in testing is for such variance in scores to be sent to a third and neutral party for an additional scoring. The averages of all the scores or the two that are in closest agreement are then often used as the final score. The most important factor for the sake of reliability, however, is that each scoring be done completely independently of the others in order to maintain the integrity of grading criteria, the scoring procedure and intra-rater reliability.

A third factor that may affect reliability is the variation in test input. Variation in the input given to test candidates can create problems in reliability when such variation has an unintended consequence on the performance on different parts or versions of a test. Therefore, when testers measure a candidate's translation ability, it is important to look at the ways in which the passages for translation are delivered to the candidates and what the qualities of those passages are. Are they formatted adequately? Are the fonts clear and legible? In order to keep the input consistent, it is advisable that the texts be retyped in a clear and legible font while maintaining as many features of the source material such as headings and paragraph flow, as possible. The desire for authenticity in passages may make the use of copies of passages in the original format attractive. However, this may be unnecessary and possibly create distracting visual variations. This should be avoided to the greatest extent possible. When working with authentic passages from source language materials, controlling the variation of the linguistic features of the passages to a high degree may be difficult. However, for the sake of the reliability of the examination, it might behoove an organization who is certifying candidates in multiple languages to consider which specific linguistic features it is seeking to test and to choose passages based on the degree to which they present those challenges. Sometimes organizations have panels of linguistic experts selecting the passages. Those experts should adhere to criteria designed for the specific selection. As we discussed in the *Wh*-section questions, criteria put forward for passage selection must be specific. For example an organization can give linguists a criterion such as 'make sure that each passage has common translation challenges such as false cognates, etc.' or 'check that passages are of a certain length' or 'select a passage that is generic, and one that pertains to the legal domain.' We may not think these criteria present any problem. However, in the design of a test with maximum reliability, it might be good to have a panel of linguists to analyze these texts for the ways in which they interact with the operational construct of translator competence so that the organization can know exactly which skills within the construct are being tested in each passage. With this information, a group (or

bank) of passages will create a large variety that can be based on the skills tested. Also, any variation in reliability due to the variation in the examination passage can be anticipated and controlled.

An additional type of variation that may affect reliability is the expected response. When organizations require a complete written translation as its only response mode, this aspect of variation has no foreseeable effect on reliability. Some possible areas to be aware of as organizations move into electronic-format tests are issues surrounding consistency. The test should, as much as possible, either be all on computers or all hand-written.[1] Also, the ease of use of any word processing application used for the test should be examined and piloted for possible effects on reliability before implementation. Possible problems in testing electronically are variations in the manner of character entry, variation in editing features, and variation in the visual presentation of text from those encountered in commonly used and accepted professional tools. It is important to consider, that once an electronic test format is made available, it should be either used exclusively, or that measures be taken to minimize possible effects of written versus electronic response formats on test performance and grading across languages or sittings for the sake of maximum reliability.

In terms of variation in the relationship between input and response types, two factors are highly important between the versions and the test tasks: first, the way in which the questions are presented and second, the way a candidate is expected to respond. Again, given the format of translation certification exams, there is no anticipated danger of reliability being threatened by this type of variation. One thing to be cautious about in selecting a passage is the sudden changes in the text type or genre within a piece. For example, does the text suddenly go from narrative to dialogue? If awareness of such a change is being tested, this would be fine. If it is not the skill being tested, such a change may cause unanticipated difficulty for candidates, and this could threaten reliability.

1.5 Test authenticity

Another important aspect of testing is test authenticity. Authenticity is the term that the testing community uses to talk about the degree to which tasks on a test are similar to, and reflective of a real world situation towards which the test is targeted. It is important that test tasks be as authentic as possible so that a strong relationship can be claimed between performance on the test and the performance in the target situation. If test tasks are too different from the situations in which candidates will be employing the competence being tested, the possibility that good or bad performance on the test does not reflect the ability to perform in the target situation increases (Bachman & Palmer 1996). For example, if we are testing for the ability to translate but we require the candidate to discriminate between acceptable and unacceptable translations, there is a high possibility that candidates who are strong in identifying good or bad renditions would succeed. By the same token, there may be a disadvantage for candidates who are better in producing translations rather than identifying good or less than satisfactory renditions. Therefore, it is important that the target situation in which we expect the candidates to perform be clearly defined and that the test tasks mirror that situation as clearly as possible. This issue is particularly relevant for professional associations that grant certification and use in-house tests. In general professional organizations are composed primarily of working professional translators. Members probably know the real world context in which translation competence is applied better than any test developer. However, knowing a situation intimately and defining it clearly for testing purposes are two very distinct things. Definition of a target construct often takes a particular kind of expertise that is different from the expertise of a practitioner. The practitioner is in the midst of the target situation and sometimes fails to notice aspects of the situation merely because they are taken for granted. Much of what translators do, they do automatically, and therefore, unconsciously (Toury 1995).

1. At the time of writing this paper, only the American Translators Association has conducted pilot exams on computers. Although a paper-pencil test poses threats to authenticity, most organizations are still certifying members this way.

In terms of defining the target situation, some ideas about what the primary aspects of the target context entails were set forth previously in the suggested definition of the construct of translation competence. Professional translators deal with a wide variety of source texts that were produced in diverse contexts. They are contracted by clients and agencies with diverse needs. They use a variety of tools and can work individually or as part of teams. All of these aspects form a part of the target-use situation. Although not all of these aspects can be matched with equal efficacy in an examination, it is important that the testing tasks reflect as many of these aspects as possible.

1.6 Task authenticity

When developers are creating a test, another important aspect of task authenticity is the test format, and its impact on the security of the test. When organizations require test takers to produce a translation, this task is reflective of the type of task that a professional will perform in the target situation. The main areas in

which a test task may seem to be inauthentic are the response format, the availability of tools, and the lack of time for editing. Some of these are authenticity problems which are logistically difficult to solve. The handwritten nature of the response format is seen as being fairly inauthentic for most contemporary translation workplaces. However, this is not a simple problem to solve as there are implications for test security and fairness, among others. It is possible, with current technology improvements, to disable e-mail functions temporarily and prevent the exam from leaving the examination room via the internet. Additionally exam proctors can be asked to control the environment and not allow electronic devices such as flash drives into the testing room to avoid downloading exam originals. To increase authenticity, it is important that the computerized test format mirror the tools and applications currently used by professional translators as closely as possible while maintaining test security. It is important that the word processing interface be as similar to the industry standard as possible. Available tools should also be as similar to real world working situations as possible. Since complete internet access could compromise test security (e.g. candidates could e-mail translations to each other during the test), it is understandable that to a certain degree, the format offered to examination candidates would lack authenticity (e.g. current certification exams such as ATA or NAATI are done by hand). However, creative solutions should be sought to minimize this lack of authenticity to the greatest degree possible.

The concept of translation assessment may or may not include translation, editing, and proofreading as separate tasks requiring different skills, and therefore different measurements. If an organization chooses to measure them jointly, this decision needs to be addressed by weighing categories and grading procedures, as well as in the test instructions to candidates. It is only when the test developers have considered all these elements that the test development can begin. Undoubtedly, the process begins with the question asking *what*; which is asking what the test assesses. This requires a clear definition of the test construct.

2. Defining the test construct

A construct consists of a clearly spelled out definition of exactly what a test developer understands to be involved in a given ability. If we are testing an ability to translate, it is important that we first clearly and meticulously define exactly what it is that we are trying to measure. This task not only involves naming the ability, knowledge, or behavior that is being assessed but also involves breaking it down into its constituent elements (Fulcher 2003). Thus, in order to measure a translator's professional ability in translating from one specific language into another, we need to first define the exact skills and sub-skills that constitute a translator's professional ability. In order to design and develop a test that assesses the ability to translate at a professional level, we have to define what the translation ability is. We have to operationalize it. The goal is to consider what type of knowledge and skills (in the broadest sense) might contribute to an operational definition of 'translation ability' that will inform the design and the development of a test of translation competency (Fulcher 2003). That is, we must say exactly what knowledge a translator needs to have and what skills a candidate needs to have mastered in order to function as a qualified professional translator. These abilities cannot be vague or generic.

To illustrate this we look at definitions (operationalizations) of translation competence. One definition of translation competence (Faber 1998) states the following: "The concept of *Translation Competence* (TC) can be understood in terms of knowledge necessary to translate well (Hatim & Mason 1990:32f; and Beeby 1996:91 in Faber 1998:9). This definition does not provide us with specific descriptions of the traits that are observable in translation ability, and therefore it does not help us when naming or operationalizing the construct to develop a test.

To find an example of a definition developed by professional organizations, we can look at the one published by the American Translators Association (ATA). The ATA defines translation competence as the sum of three elements: (1) comprehension of the source-language text; (2) translation techniques; and (3) writing in the target language. In a descriptive article, Van Vraken, Diel-Dominique & Hanlen (2008 http://www.atanet.org/certification/aboutcert_overview.php) define criterion for comprehension of the source text as "translated text reflects a sound understanding of the material presented." The criterion for translation techniques is defined as "conveys the full meaning of the original. Common translation pitfalls are avoided when dictionaries are used effectively. Sentences are recast appropriately for target language style and flow." Finally, evaluation of writing in the target language is based on the criterion of coherence and appropriate grammar such as punctuation, spelling, syntax, usage and style. In this professional organization, the elements being tested (according to their definition) are primarily those belonging to the sub-components of grammatical competence (language mechanics) and textual competence (cohesiveness and style). But while this definition is broader than that of Beeby, Faber, or Hatim and Mason (in Faber 1998), it still does not account for all the elements present in the translation task required by their test.

We could argue that translation involves various traits that are observable and/or visible which include, but are not limited to conveyance of textual meaning, socio-cultural as well as sociolinguistic appropriateness, situational adequacy, style and cohesion, grammar and mechanics, translation and topical knowledge.

These traits contribute to an operational definition of translation ability, and they are essential to the development of a test.

A test can only be useful and valid if it measures exactly what it intends to measure; that is, if it measures the construct it claims to measure. Therefore, for a translation test to be valid, it must measure the correct construct, i.e. translation ability. The first crucial task of the test developer is to define the construct clearly. Once the construct is defined clearly, then and only then can the test developer begin to create a test that measures that construct. Evidently, this process is not linear in the sense that the construct undergoes revisions and modifications, but its definition does need to occur *a priori* (Bachman 1990). As evident from testing principles, a central issue in assessment is construct validity. Establishing construct validity ensures that the right construct is being measured. In the next section we will review how the construct of translation competence has been conceptualized.

3. Review of relevant literature

A good translation is a highly complex activity that involves many diverse areas of knowledge and skill. Therefore, defining translation competence is not an easy task. It is a "dynamic process and it is a human and social behavior" (Cao 1996: 231) that results from experience, training and the feedback effects of client-translator or translator-reader interaction. (Neubert & Shreve 1992: 10 in Cao 1996: 231). Currently, there is no one definition of translation competence and its components that is universally accepted within the academic field of translation studies (Arango-Keith & Koby 2003). In fact, there is considerable debate about how to define translation competence and exactly how its constituent elements are to be conceptualized, broken down and interconnected. Despite this disagreement, the academic discussion about translation competence can be an important aid in helping to define the constructs of what makes a competent and professionally qualified translator.

As Kiraly points out "An empirical description of translation processes implies the possibility of describing what a professional translator has to know and has to do (even if much of what he or she does is subconscious) to produce a high-quality translation." (1995: 13). To begin, Wolfram Wilss (1982 in Kiraly 1995) initially described translation competence as consisting of three primary components which include (a) source language receptive competence coupled with (b) target language reproductive competence operating within (c) a super-competence which reflects the ability to transfer the message from one language to another. This description of translation competence emphasizes that it is not merely enough to transfer a message from one language to another, but rather that there is a need to be strategic about it (Valdés and Angelelli 2003). Presas (2000) helps us further define the idea of Wilss' super-competence by defining what it is that makes translation competence different from bilingual competence. She emphasizes that a competent translator uses specialized linguistic and cultural knowledge to control interference in both the reception of information from the source text and the production of the target text. According to Presas, the competent translator does this in part through making a transfer at the level of meaning rather than at the level of words and phrases between two connected but separate code systems, i.e. languages. However, in order to validly and reliably test these specialized skills and knowledge it is necessary to define them further.

Many contemporary definitions of translation competence view translation as a specialized sort of communication. They define the translator as an individual who is interpreting a text that was written to perform a function in the source language and culture while delivering it into a new form in order to perform a function in the target language and culture (Kiraly 1995; Cao 1996; Neubert 2000; Beeby 2000; Orozco 2000; Adab 2000; Colina 2003). This type of functional approach to translation views translation competence as a specialized type of communicative competence. This concept of a communicative competence comes from the field of Second Language Acquisition (SLA).

Although the fields of SLA and Translation Studies, despite focusing on similar phenomena, have not historically engaged in the sharing of knowledge and theories, greater cross-fertilization between the two has occurred in recent years. The beginnings of this can be seen in more recent works on teaching and testing for translation and interpreting (Angelelli 2000, 2003, 2004b and 2007a and b; Schäffner and Adab 2000; Colina 2002, 2003 and 2008).

SLA theory also interacts with testing theories, especially in reference to testing language abilities and analytical skills (Bachman 1990; Bachman & Palmer 1996; Cohen 1994; Fulcher 2000; Johnson 2001). Therefore, it is important to have an understanding of these theories in communicative competence in order to frame a construct of communicative translation competence that allows us to create a theoretically sound assessment.

Among the most commonly used models of communicative competence in the fields of SLA and language assessment is that proposed by Bachman (1990). His model of communicative competence is divided into organizational competence, which involves the more mechanical aspects of communication and how they are organized; and pragmatic competence, which deals with how language is used in specific situations. Bachman further subdivides organizational competence into grammatical and textual competences. Grammatical competence is composed of the individual's knowledge of the forms and structures of a language.

Textual or discourse competence refers to the way in which sentences and larger chunks of language are woven together to create a coherent and cohesive message. Pragmatic competence is further divided into illocutionary and sociolinguistic competences. Illocutionary, or strategic competence consists of the individual's knowledge of the ways in which language is used to accomplish functions and create impressions and relationships. Sociolinguistic competence is an individual's knowledge of ways of speaking and interacting through language, e.g. politeness, taboos, etc. (Bachman 1990). These different competences are used in any given act of communication. An act of translation, by virtue of being an instance of language use, is a form of communication. Therefore, these communicative competences cannot be disregarded in the construct of translation competence.

For the current discussion on the construct of translation competence, a logical starting point on communicative translational competence is to consider the definitions proposed by Cao (1996), Colina (2003), and PACTE (in Orozco 2000). In Cao's model of translation proficiency, there is a translational language competence (defined similarly to Bachman's language competence), in addition to translational knowledge structures, such as world and specialized knowledge. There also exists a translational strategic competence which is connected to both of the other two, as well as the situational context. Thus, the core of translation competence lies in matching language competence and knowledge structures to the current communicative context (i.e. the translation task). This is achieved through the application of competence in translational strategies in order to execute a communicative translation task. This model helps us to see that translation lies not only in the ability to effectively convey a message between languages but also the ability to do so in a particular context. A translation needs to be both a good rendering of the source text and the proper rendering to meet the needs of the receiver. Being able to produce the right translation in the right context is therefore seen as a part of translation competence.

Colina (2003) defines communicative translational competence as consisting not only of communicative competence in both languages, but also including an element of interlingual and intercultural communicative competence. Colina emphasizes that translation is a communicative interaction in as much as the translator is responsible for the interpretation of source text (ST) meaning and its negotiation and expression in accordance with task specifications, translational conventions, and target language conventions. Thus, the model of translation competence in her work considers the ways in which the context of a translation also operates on the other competences in communicative translation competence. The model that Colina chooses to reflect these views is one put forth by Cao (1996).

Just as Colina's (2003) model adds the element of context to our understanding of translation competence, the PACTE model as outlined in Orozco (2000)

adds the element of methods of achieving communicative translation goals. The PACTE model presents two major competences: transfer competence and strategic competence. Transfer competence is defined as the "the ability to complete the transfer process from the ST to the target text (TT), taking into account the translation's function and the characteristics of the receptor" (Orozco 2000: 199). This competence is further broken down into comprehension competence, the ability to de-verbalize the message and control interference, re-expression competence, and competence in choosing the most adequate method. Transfer competence is seen as being informed by four other competences: communicative competence in two languages, extra-linguistic competence (i.e. world and specialist knowledge), psycho-physiological competence (i.e. using various cognitive, psychomotor and attitudinal resources), and instrumental-professional competence (i.e. the ability to use the tools and apply the norms of the profession). The final element in this model is strategic competence in which all these processes are used in finding and solving problems in the translation process (Orozco 2000).

The standout feature of Orozco's model is the emphasis placed on tools and processes for problem-solving. Competent translators need to be able to find and correct problems in their own translations and processes. It is also important that they are familiar with the tools and standards of their trade. The strategic use of software, on-line residing tools, and more traditional items like dictionaries, are an important part of any translator's work. Even more importantly, knowledge of common practices and standards of professional conduct are also vital to competent translation. These are the tools that translators use to overcome translation problems and, therefore, form a vital part of the competence of a professional translator.

In addition to the contributions of the models of communicative competence and translational language competence, we need to look at the mode in which language is used. There are different modes through which the overall language competence in each language is engaged in the communicative act of translating. The American Council of Teachers of Foreign Languages through the National Standards in Foreign Language Education Project (2006) has defined three primary modes of language use according to the level of contact and interaction between the participants in the act of communication: the interactional mode, the interpretive mode, and the presentational mode. The interactive mode involves situations in which all participants can participate as both the presenter and the audience in conversations. The interpretive mode refers to situations in which the language user is primarily receiving information and the original presenter/writer is not available to clarify questions, similar to what occurs while reading or listening to text in a target language. The presentational mode involves situations in which the language user is primarily presenting information (either orally or in

writing) with limited possibilities to interact directly with the eventual recipient of the message such as writing a document or presenting to a large audience. Of these modes, the interpretative mode (specifically reading) plays a greater role in relation to the translator's access to source text, while the presentational mode (specifically writing for a large readership) plays a greater role in the translator's production of the target text.

Despite the difference in modes, many of the underlying sub-competences are similar. A candidate seeking translation certification must have control and awareness of the various aspects of the source language in order to competently interpret the meaning of the source text. This means viewing the text through the cultural lenses with which it was produced. This also includes: (1) the grammatical aspects of a language which encompass the ways in which words and word parts are combined to form meaning; (2) the textual aspects of the language which include the conventions for how the information in a text is linked together, structured, and presented; and (3) the pragmatic aspects which include the culturally specific limitations on what is said, how it is said, and how these create feelings and establish relationships with the readership and the subject matter. Without an understanding and awareness of the subtleties represented in these different aspects of language, a candidate is not able to fully comprehend either the source text or how these elements will affect the translation task. Similarly, a candidate must have control and awareness over the grammatical, textual and pragmatic aspects of the target language in order to competently produce a target text. These are complex skills and abilities.

In light of the current literature on what constitutes translator competence and communicative translation competence, we turn our attention to the construct as it is defined by available documents from professional organizations granting certification. This definition of the current construct is drawn from both the overt explanations of what is being tested and the implicit priorities set forth in such documents as grading protocols when available.

In a set of ATA documents written to inform the general public about the nature of translation,[2] we can see how the professional organization conceptualizes translation. The ATA through its publication "Translation: Getting It Right," emphasizes the aspect of a translator's strategic (Orozco 2000) and intercultural communication (Colina 2003) competence by mentioning how translators can bridge cultures by encouraging consumers to tell the translator what the translation is for. This shows an emphasis on the pragmatic and functional competences that professional translators possess.

2. At the time of writing this chapter, the author consulted the websites of various professional associations administering translation tests, such as ATA, NAATI, NAJIT. Information on the conceptualization of the test construct was only available at ATA website.

However, the article mentioned earlier by van Vraken, Diel-Dominique and Hanlen about the ATA certification exam does not discuss the purpose of the translation as it mentions the three elements that the exam measures. Those are: (1) comprehension of the source-language text, (2) translation techniques and (3) writing in the target language. The article defines criterion for comprehension of the source text as "translated text reflects a sound understanding of the material presented." The criterion for translation techniques is defined as "conveys the full meaning of the original. Common translation pitfalls are avoided when dictionaries are used effectively. Sentences are recast appropriately for target language style and flow." Finally, evaluation of writing in the target language is based on the criterion of coherence and appropriate punctuation, spelling, syntax, usage and style.

The current three part construct used by the ATA seems to primarily emphasize the reading comprehension, translation ability (not operationalized) and the micro-linguistic elements of translation competence present in writing (e.g. lexicon, grammar and punctuation rather than discourse, cohesion, etc.). The first element of this construct, the comprehension of the source text, combines aspects of both the organizational and pragmatic competences as defined by Bachman (1990). In order to comprehend a source text, a translator must be able to both make sense of the letters, words, phrases, sentences and text as a whole and understand that text in terms of what it means in the original culture. Therefore, in order to make this concept more reliably measurable, we need to break it down further. The second element of translation "technique" fits in with Cao's (1996) translation knowledge structures and Orozco's (2000) transfer competence. This is another example in which, in order for the sub-components of translation "technique" to be more reliably measured, we need to break them down and outline them as specific behaviors and observable aspects of a translated text. The final aspect of this construct, writing in the target language, is focused primarily on the micro-elements of the competence when translating, such as grammar, lexicon and punctuation.

The construct as defined by professional associations (NAJIT, ATA) seems somewhat more limited than the communicative constructs that have been established in the fields of language testing, SLA, and translation studies that take into account more macro elements of cross-linguistic communication such as purpose of the translation, readership, cohesion, register, etc. The language element is quite prominent in the professional organizations' construct and communicative translation competence is not fully represented. The elements being tested under professional organizations definition are primarily those belonging to the sub-components of grammatical competence (language mechanics) and textual

competence (cohesiveness and style) of Bachman's (1990) model. The pragmatic competence defined by Bachman (1990) is only partially represented in the comprehension and rendering of the passage. Candidates are only judged on their understanding of the "full meaning and purpose" (not defined), of the source text and their ability to convey that meaning (again, not defined) in the target text. It is also problematic that comprehension in one language is being tested through production in another language. One could argue that the two could be unrelated. (A better test of comprehension might be to have the candidate summarize the text in the source language, or complete a reading comprehension exercise.) Also, it appears that there is no testing as to whether the candidate can perform the communicative functions necessary to produce a text that is appropriate to the target language readership. This raises the question: is the focused emphasis on grammatical competence appropriate?

Additionally, many current tests (e.g. ATA, NAATI) which certify translators ask for the delivery of a "neutral" translation that mirrors the source text and does not take a particular audience into account. Many times candidates are discouraged from making changes in the style, register and the use of regionalisms although these may be communicatively required in certain translation situations. Brief test instructions (e.g. this text is going to be published in Readers' Digest) designed to save space or time may result in a test-taker having no clear target readership or purpose, which in turn does not allow a candidate to show the best sample of his/her ability and does not allow the grader to make a judgment about the candidate's control of register or regionalisms. Similarly, the lack of a specified audience and function for the translation does not allow for the measurement of illocutionary competence (the ability to achieve functions with language, e.g. apologize, challenge, etc.) and sociolinguistic competence (culture specific references and what is allowed or disallowed in a given context in a given culture). The inclusion of these elements in other materials (e.g. translation brochures such as ATA Translation: Getting it Right) about good-quality translation suggests that they are essential and should be included in the construct of assessment for certification.

4. An expanded framework for an expanded construct

Given what research in the areas of communicative competence, sociolinguistics, translation studies, second language acquisition and testing have shown in addition to what the professional associations granting certification state, how should the construct of communicative translational competence be defined for the specific purpose of a certification examination? What sub-components of this construct are being measured by tests currently used? What sub-components should

be part of a certifying exam? What sub-components can be tested separately? What separate tests, if any, could be used as predictors of success in certification exams? It seems that, partially, associations refer to an operational construct of translation competence that is already functionalist and communicative in nature (e.g. translation brochure ATA Translation: Getting it Right). However, when it comes to defining it (Van Vraken, Diel-Dominique & Hanlen 2008) , the tendency is to focus more on the grammatical and textual competences. Is this a problem? If so, we need to ask ourselves why. While comparing professional associations' ideologies on translation competence to what research in translation studies state, we see a gap. When operationalizing the construct, professional associations tend to have a narrower definition of translation competence, and many times pragmatic and other elements are not included in the construct to be measured.

The operational construct needs to be articulated along similar lines to those used in translation studies in order to capture translation in its entirety and thus properly measure it (Bachman 1990). To this end, I will propose a working definition of the construct of communicative translation competence that includes the communicative elements of Hymes (1974) in addition to Bachman's (1990) frameworks of communication and communicative competence, and the contributions of Cao (1996), Colina (2003) and some of the instrumental elements reflected in the PACTE definition of translation competence. This new measurable construct includes the following sub-components: (1) linguistic competence, (2) textual competence, (3) pragmatic competence, and (4) strategic competence. I do not presume to present this construct as a definite operationalization of translation abilities. This construct is presented as a guide, as a lead to chart directions for research and development in translation assessment, specifically as it pertains to professional associations granting certifications. As research develops, and as we subject this framework to empirical tests, it is likely that it will undergo changes to reflect our collective growing knowledge in the area. Let us look at the subcomponents of the construct.

1. *Linguistic-level competence.* The first sub-component of our construct is the linguistic component defined here in its narrowest sense. Translation is, in many ways, the communicative act of transferring a message between two languages. This activity requires a certain degree of communicative competence in two languages. In all of the models from Translation Studies previously reviewed (Cao 1996; Colina 2003 and Orozco 2000), competence in two languages is an undisputed aspect of translation competence (the same holds true for interpreting – see Angelelli 2003 and 2007b). In language assessment, the dominant models for language competence are the ones set forth by Bachman (1990) and Johnson (2001). Bachman's (1990) model forms the basis for Cao's (1996), Colina's (2003)

Translational Language Competence and Angelelli's model of Language Competence for Interpreting (2003) which is also combined with Johnson's (2001). Cao's notion of organizational competence (1996) includes grammatical competence and textual competence.

Clear grammatical competence plays a vital role in the act of translation. Cao defines this sub component as control of vocabulary (the words of a language), morphology (the way that smaller parts combine to form words), syntax (the way that words combine to form phrases and sentences), and graphemic (writing system) knowledge.

Each of these aspects contribute both to the interpretation of the source text and the production of the target text. A breakdown in any one of these areas can affect the act of translation. Insufficient knowledge of vocabulary can lead to miscomprehension of the source text or failure to successfully communicate meaning in the target text. This competence can be aided through proper and skillful use of dictionaries and other professional resources, but only to a degree. A translator's knowledge of morphology and syntax helps both interpretation and production in the act of translation. A failure to understand the effects that syntax and morphology have on meaning can also lead to incomplete or mistaken understanding of the source text and the production of a difficult or misleading rendition in the target text. Graphemic knowledge, likewise, plays a part in both the interpretation and production aspects of translation. Failure to understand differences in meaning carried by punctuation and diacritical marking can lead to misapprehension of the source text. Lack of knowledge of writing mechanics or misapplication of graphemic knowledge can lead to interference with communication of meaning and difficulty in the comprehension of a target text. Therefore, each of the aspects of grammatical competences is vital to the act of translation and is being assessed either directly or indirectly in any authentic translation task.

To be measured, this linguistic sub-component of translation competence needs to be clearly stated. To assess it during certification exams, for example, one can start by considering a continuum of more to less successful translations and describing what they would look like, how those would reflect more or less mastery of this subcomponent. Table 1 illustrates statements in a possible 5-point-scale to assess the linguistic sub-component of translation competence.

2. *Textual competence.* Textual competence, or the ability to string ideas together as a text, is also a vital part of any act of translation. Within the purview of textual competence, Cao includes cohesive competence, the ability to use linguistic devices to connect sentences, ideas, rhetorical organization competence, and the ability to organize a text in the most appropriate way to achieve its aims in a given community. In fact, Colina (2003) points out that the literature suggests that one

Table 1. Linguistic sub-component (T = translation; TL = target language)

5	T shows a masterful control of TL grammar, spelling, and punctuation. Very few or no errors.
4	T shows a proficient control of TL grammar, spelling, and punctuation. Occasional minor errors.
3	T shows a weak control of TL grammar, spelling, and punctuation. T has frequent minor errors.
2	T shows some lack of control of TL grammar, spelling and punctuation. T is compromised by numerous errors.
1	T exhibits lack of control of TL grammar, spelling and punctuation. Serious and frequent errors exist.

of the skills that separates the professional and the novice translator is the attention to textual clues. A successful translator must understand how a source text is structured internally and the effects that such an organization has on the meaning that the author is creating and communicating. Likewise, the successful translator activates his/her knowledge of similar available devices in the target language to render a similar message and meanings in the target text. This includes such aspects as the creation (by the source text author or the translator) of tenor and tone. This competence involves understanding the rules and conventions of rhetoric and cohesion in both codes well enough to know what meanings are conveyed; through either following or breaking said conventions in the source language, in addition to being able to render similar meaning in the target language, depending on the translation task. In fact, this competence is vital to successful translation both in the interpretative and the presentational modes of using language for communicative purposes (National Standards in Foreign Language Education Project (2006). That is to say, translators will make use of this mode to both read and interpret the source text as well as to produce the target text. Therefore, any authentic translation task will to some extent assess textual competence along with grammatical competence.

Similar to the linguistic sub-component, the textual sub-component of translation competence needs to be clearly stated. Table 2 illustrates a continuum of more to less successful translations by describing how those would reflect more or less mastery of this subcomponent.

3. *Pragmatic competence.* In Bachman's model of communicative competence (1990) adopted by Cao to discuss translation competence (1996) pragmatic competence is subdivided into illocutionary competence and sociolinguistic competence. Even if we could argue that sociolinguistic competence and pragmatic competence are separate, that Cao could have treated them separately, or could

Table 2. Textual sub-component (T = translation; TL = target language)

5	T is very well organized into sections and/or paragraphs in a manner consistent with similar TL texts. The T has a masterful style. It flows together flawlessly and forms a natural whole.
4	T is well organized into sections and/or paragraphs in a manner consistent with similar TL texts. The T has style. It flows together well and forms a coherent whole.
3	T is organized into sections and/or paragraphs in a manner generally consistent with similar TL texts. The T style may be inconsistent. There are occasional awkward or oddly placed elements.
2	T is somewhat awkwardly organized in terms of sections and/or paragraphs or organized in a manner inconsistent with similar TL texts. The T style is clumsy. It does not flow together and has frequent awkward or oddly placed elements.
1	T is disorganized and lacks divisions into coherent sections and/or paragraphs in a manner consistent with similar TL texts. T lacks style. T does not flow together. It is awkward. Sentences and ideas seem unrelated.

have used sociolinguistic competence as an umbrella term, I will follow her classification for the sake of simplicity. Illocutionary competence is the knowledge of how language is used to perform functions (e.g. apologizing, complaining). This competence plays an important role in the act of translation both when the translator approaches the source text as well as when he/she produces the target one. When a translator is approaching and analyzing a source text, this competence allows to discern whether the text is a polemic, a primarily objective report of data, a proposal for action, etc. Likewise, in the production of the target text, the translator makes use of this competence to reproduce those functions in the translation. Under the label of sociolinguistic competence, Bachman includes knowledge of linguistic variations (i.e. dialects, regionalism, and national varieties) and knowledge of cultural reference and figures of speech. Knowledge of variation is important in being able to interpret a source text in a dialect other than the standard. It also may be important in helping the translator understand the particular cultural assumptions that may underlie a source text. These features may also be communicated through cultural references and the use of figures of speech. Culture reference and figures of speech can be an obstacle to successful translation. A successful translator is aware of these elements and is able to resist the temptation to translate them directly and/or find a successful way of communicating their meaning in the target text. In addition to these two aspects of socio-cultural knowledge, it is important to add register (Angelelli 2006, 2004). The degree to which a text creates a social relationship between the author and his reader must be understood and properly rendered in many translation tasks. To provide an extreme example, it would be improper to translate a formal invitation extended to a

foreign dignitary to attend a closing conference ceremony that needs to be RSVP into a colloquial and informal text that informs the reader about a gathering and does not convey the sense of formality and respect intended in the original. Such failure to recognize and communicate the use of register may lead to the intended readership's misunderstanding of the true communicative intent of a document.

Additionally, individuals who take part in a communicative event (and translation is one of them), to use Hymes' terms (1974) need to have knowledge about ways of "doing" and ways of "communicating". Hymes' work on discourse communities and communication, specifically applied to speaking (see Hymes taxonomy 1974) can also be applied to writing, written communication and therefore written translation. To belong to a community of discourse (for example American engineers writing for a German journal on engineering) means that the translator needs to be a competent writer in engineering by using specialized vocabulary and exhibiting mastery of underlying structures and assumptions that are relevant to that specific field of activity. In other words, when translators engage in technical translation for engineers, they need to be able to write as if they belonged to the discourse community of engineers. Even when they do not belong to that discourse community, they have to be perceived by the reader as if they were a part of it or as a "temporary guests" able to navigate it successfully (cf. Angelelli 2000 and the notion of interpreters as temporary guests in discourse communities). In the case of scientific and technical translation, translators need to know the content area of the field. They also need to know the typical renditions of technical terms and ways of discussing different topics in a given field in each of the languages involved. Failure to properly apply such knowledge may lead to an unsuccessful translation that is neither comprehensible nor useful to the target readership.

Translators working in specialized fields must also have enough field knowledge to be able to successfully render terminology and/or concepts that may be new to the field in either language. Similarly, business and legal translation depend on knowledge of governments, legal systems, business organizations and other aspects of these fields that underlie the texts in both cultures. They must be able to communicate these differences while making the target text comprehensible to the recipient. These tasks require a working knowledge of the degrees to which the texts and cultures being mediated converge and diverge. It also requires the ability to make the technical document in one language available to the layperson in another language when necessary. This is all modified by the translator's knowledge about the recipients' knowledge of the various fields involved and cultural differences that exist. Therefore, we need to assess the knowledge of discourse and register as well as the knowledge of the socio-cultural aspects of the language.

Table 3. Pragmatic sub-component (T = translation; TL = target language)

5	T shows a masterful ability to address the intended TL audience and achieve the translations intended purpose in the TL. Word choice is skillful and apt. Cultural references, discourse, and register are completely appropriate for the TL domain, text-type, and readership.
4	T shows a proficient ability in addressing the intended TL audience and achieving the translations intended purpose in the TL. Word choice is consistently good. Cultural references, discourse, and register are consistently appropriate for the TL domain, text-type, and readership.
3	T shows a good ability to address the intended TL audience and achieve the translations intended purpose in the TL. Cultural references, discourse, and register are mostly appropriate for the TL domain but some phrasing or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.
2	T shows a weak ability to address the intended TL audience and/or achieve the translations intended purpose in the TL. Cultural references, discourse, and register are at times inappropriate for the TL domain. Numerous phrasing and/or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.
1	T shows an inability to appropriately address the intended TL audience and/or achieve the translations intended purpose in the TL. Cultural references, discourse, and register are consistently inappropriate for the TL domain. Most phrasing and/or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.

Table 3 illustrates statements of a continuum of more to less successful translations by describing how those would reflect more or less mastery of this subcomponent.

4. *Strategic competence*. The final major aspect of translation competence that is included in this proposed construct is translation strategic competence. This competence has to do with the way in which a translator approaches a translation task and the methods he/she uses to pinpoint and overcome problems within the performance of the translation assignment. According to Orozco (2000), strategies include distinguishing between primary and secondary ideas, establishing conceptual relationships, searching for information, task planning and many others. This competence is where the conscious skill of the translator enters into the translation task. Here is where interference is controlled, where culture is consciously mediated, and where the decision is made to consult resources and determine how they can be properly applied.

Within a translator's strategic competence lies what PACTE (in Orozco 2000) calls instrumental-professional competence. Use of professional tools and standards of behavior is an important part of a translator's ability to be strategic. A translator's knowledge competence can be augmented but not substituted by the use of reference materials and consultation with professionals in the given field.

Proper knowledge of how to use and access these tools, therefore, is also a part of translator strategic competence. Colina (2003) points out that the proper use of translation tools is a factor that differentiates the novice translator from her/his professional counterpart. Inexperienced translators tend to make unskillful use of these tools since they do not possess the benefit of experience to tell them what to accept and what to reject. For the contemporary translator nowadays, these instrumental-professional competences go far beyond the use of dictionaries. The internet, electronic reference materials, CAT tools (computer applications to aid in the process of translation, memory applications, machine-assisted-translation software, as well as word processing programs) are all vital part of the toolkit for today's translator. One must not only know how but when to use each. Competent professional translators must be able to perform a successful web search and be able to identify which sources to accept and reject. They must be able to maintain and properly apply computer-aided translation tools such as translation memory and databases. These skills also include having a sufficient knowledge of the language to be able to successfully accept and reject a word processor's spelling and grammar corrections.

Similarly, the ability to manage human resources is an important part of a translator's strategic competence. A working translator in today's market must know how to obtain necessary information from a manager or client (Fraser 2000). Increasingly, a professional translator must also know how to work in a team (Arango-Keeth & Koby 2003). Due to testing formats and technology limitations, it may not be possible to assess these competences in every assessment task. However, it is important to see them as part of the construct of translation competence and acknowledge whether or not a certification test chooses or does not chose to assess them. Tests may or may not target all of the subcomponents of translation competence. As long as test developers clearly define what their test intends to measure, and justify the reasons for doing so, candidates know in advance what sub-components are included in a test and what sub-components are not.

To some extent, the strategic translation competence is the translator's ability to exercise conscious control over their linguistic, cultural, field, and instrumental knowledge. This competence is involved in choosing which features of the target text to put in the foreground and which to put in the background. It involves choosing between making something explicit (e.g. explaining a point that may be unfamiliar to the target audience) and using other devices to make a point implicit. This competence may also be reflected in the use of self-editing and drafting processes. It is included both explicitly and implicitly in any translation assessment even when the grader only sees the final product that the examinee submits. The application of strategy is only evident in its effect: strategic competence is truly demonstrated in the absence of problematic translations in the final product.

Table 4. Strategic sub-component (T = translation; TL = target language)

5	T demonstrates astute and creative solutions to translation problems. Skillful use of resource materials is evident.
4	T demonstrates consistent ability in identifying and overcoming translation problems. No major errors and very few minor errors are evident. No obvious errors in the use of resource materials are evident.
3	T demonstrates a general ability to identify and overcome translation problems. However, a major translation error and/or an accumulation of minor errors are evident and compromise the overall quality of the translation. Improper or flawed use of reference materials may be reflected in the TT.
2	T demonstrates some trouble in identifying and/or overcoming translation problems. Several major translation errors and/or a large number of minor errors are evident and compromise the overall quality of the translation. Improper or flawed use of reference materials is reflected in the TT.
1	T reflects an inability to identify and overcome common translation problems. Numerous major and minor translation errors lead to a seriously flawed translation. Reference materias and resources are consistently used improperly.

Table 4 illustrates statements of a continuum of more to less successful translations by describing how those would reflect more or less mastery of this subcomponent.

Now that we have operationalized these sub-components of translation ability, we can turn to discussing ways to assess them. Because they are definable, they are also gradable. In the field of testing, it is not uncommon to see subcomponents of a construct scored with a scoring rubric. Many professional associations granting certification do not use rubrics. In the next section I explore the use of a rubric for certifying translators.

5. Using a rubric

Rubrics are commonly used in testing. They allow for a more systematic and holistic grading. A rubric generally contains all sub-components that constitute the construct. It provides descriptive statements of behaviors that candidates may exhibit in a particular sub-component.

Since a scoring rubric can be used to holistically score virtually any product or performance (Moss and Holden 1988; Walvood and Anderson 1998; Wiggins 1998), it makes sense to discuss its feasibility for scoring translation. A rubric is developed by identifying what is being assessed (i.e. translation competence). It implies identifying the characteristics of translation competence, the primary

traits of the product or performance, (i.e. micro-linguistic competence, textual competence, pragmatic competence, strategic competence, etc.) and then delineating criteria used to discriminate various levels of performance, as was done earlier with each of these sub-components. For example, in order to be considered a competent professional translator, an individual must demonstrate sufficient control and understanding of the linguistic features of the source language to successfully comprehend the meaning of a source text appropriate to the translation task. In addition, sufficient control and understanding of the linguistic features and writing conventions in the target language is necessary to successfully produce a target text appropriate to the translation task. This includes: grammatical and mechanical control, control of cohesive and textual devices, control of functional and socio-cultural aspects of the languages, and sufficient relevant world and technical knowledge to successfully complete the translation task. This includes both knowledge of cultural differences in world views and ways of doing things as well as area specific knowledge of institutions, ways of working, professional conventions, concepts and terminology. In addition, a competent professional translator exhibits an ability to identify and overcome problem areas in the performance of a translation task. This includes: application of strategies, use of professional tools and resources, and the ability to work in teams and work with manager and/or clients. Additionally, each of the sub-components carries a certain weight (decided by the association based on their needs and stated on test specifications).

By constructing a scoring rubric, graders can holistically score all the elements that were considered relevant to be included in a test. This assures that in the test, the construct that was intended to be measured is not only measured by the test (as a result of careful development) but it is also scored by graders. This is an important contrast to the point-adding or point-deducting system which is many times used in schools and professional associations. A description of the best work that meets these criteria and the best performance that can be expected will describe the top score. The worst work that can be expected using the same criteria constitutes the lowest acceptable score. Intermediate level work is assigned intermediary scores, and the number of intermediary levels determines the number of levels of a scale. For example, a rubric can have a scale that runs from one to six (e.g. unacceptable translation, inadequate translation, barely adequate translation, competent translation, very competent translation, and outstanding translation), or from one to three (unacceptable, barely acceptable, clearly acceptable) or any other set that is meaningful for the organization that is developing the test.

Table 5 is an example of a five-point-scale rubric that could be used to assess translation ability by professional associations. It was drafted for the American Translators Association as a result of a self-study on their certification exam.

Table 5. Working draft for rubric to assess translation (Angelelli 2006)
T = translation; TL = target language; ST = source text

	Source Text Meaning
5	T contains elements that reflect a detailed and nuanced understanding of the major and minor themes of the ST and the manner in which they are presented in the ST. The meaning of the ST is masterfully communicated in the T.
4	T contains elements that reflect a complete understanding of the major and minor themes of the ST and the manner in which they are presented in the ST. The meaning of the ST is proficiently communicated in the T.
3	T contains elements that reflect a general understanding of the major and most minor themes of the ST and the manner in which they are presented in the ST. There may be evidence of occasional errors in interpretation but the overall meaning of the ST appropriately communicated in the T.
2	T contains elements that reflect a flawed understanding of major and/or several minor themes of the ST and/or the manner in which they are presented in the ST. There is evidence of errors in interpretation that lead to the meaning of the ST not being fully communicated in the T.
1	T shows consistent and major misunderstandings of the ST meaning.
	Style and Cohesion (addresses textual sub-component)
5	T is very well organized into sections and/or paragraphs in a manner consistent with similar TL texts. The T has a masterful style. It flows together flawlessly and forms a natural whole.
4	T is well organized into sections and/or paragraphs in a manner consistent with similar TL texts. The T has style. It flows together well and forms a coherent whole.
3	T is organized into sections and/or paragraphs in a manner generally consistent with similar TL texts. The T style may be inconsistent. There are occasional awkward or oddly placed elements.
2	T is somewhat awkwardly organized in terms of sections and/or paragraphs or organized in a manner inconsistent with similar TL texts. The T style is clumsy. It does not flow together and has frequent awkward or oddly placed elements.
1	T is disorganized and lacks divisions into coherent sections and/or paragraphs in a manner consistent with similar TL texts. T lacks style. T does not flow together. It is awkward. Sentences and ideas seem unrelated.
	Situational Appropriateness (addresses pragmatic sub-component)
5	T shows a masterful ability to address the intended TL audience and achieve the translations intended purpose in the TL. Word choice is skillful and apt. Cultural references, discourse, and register are completely appropriate for the TL domain, text-type, and readership.
4	T shows a proficient ability in addressing the intended TL audience and achieving the translations intended purpose in the TL. Word choice is consistently good. Cultural references, discourse, and register are consistently appropriate for the TL domain, text-type, and readership.

Table 5. (*continued*)

3	T shows a good ability to address the intended TL audience and achieve the translations intended purpose in the TL. Cultural references, discourse, and register are mostly appropriate for the TL domain but some phrasing or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.
2	T shows a weak ability to address the intended TL audience and/or achieve the translations intended purpose in the TL. Cultural references, discourse, and register are at times inappropriate for the TL domain. Numerous phrasing and/or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.
1	T shows an inability to appropriately address the intended TL audience and/or achieve the translations intended purpose in the TL. Cultural references, discourse, and register are consistently inappropriate for the TL domain. Most phrasing and/or word choices are either too formal or too colloquial for the TL domain, text-type, and readership.
	Grammar and Mechanics (addresses micro-linguistic sub-component)
5	T shows a masterful control of TL grammar, spelling, and punctuation. Very few or no errors.
4	T shows a proficient control of TL grammar, spelling, and punctuation. Occasional minor errors.
3	T shows a weak control of TL grammar, spelling, and punctuation. T has frequent minor errors.
2	T shows some lack of control of TL grammar, spelling and punctuation. T is compromised by numerous errors.
1	T shows lack of control of TL grammar, spelling and punctuation. Serious and frequent errors exist.
	Translation Skill (addresses strategic sub-component)
5	T demonstrates able and creative solutions to translation problems. Skillful use of resource materials is evident.
4	T demonstrates consistent ability in identifying and overcoming translation problems. No major errors and very few minor errors are evident. No obvious errors in the use of resource materials are evident.
3	T demonstrates a general ability to identify and overcome translation problems. However, a major translation error and/or an accumulation of minor errors are evident and compromise the overall quality of the translation. Improper or flawed use of reference materials may be reflected in the TT.
2	T demonstrates some trouble in identifying and/or overcoming translation problems. Several major translation errors and/or a large number of minor errors are evident and compromise the overall quality of the translation. Improper or flawed use of reference materials is reflected in the TT.
1	T reflects an inability to identify and overcome common translation problems. Numerous major and minor translation errors lead to a seriously flawed translation. Reference materials and resources are consistently used improperly.

In order to obtain feedback it was presented to ATA graders during the 48th ATA Annual Conference in November 2007. At the time of writing this article, the ATA certification committee had not made a decision to expand or change the sub-components of translation competence listed in their website, nor had they decided on the consideration of this rubric, either partially or in its entirety.

The operational categories selected in the creation of this rubric may not at first glance appear to be equal to the sub-components of the construct defined above. They are however inter-related. Some sub-categories have been collapsed and unified into a single category in order to minimize the number of categories that graders must rate and to facilitate the definition of performance levels (Bachman & Palmer 1996; Cohen 1994). Additionally the terms used in the rubric are more aligned with terminology generally used by graders. The definition of the rubric categories and their justifications are as follows:

Source text meaning is a measure of the extent to which the candidate's response (the target text) reflects or fails to reflect an adequate understanding of the themes and rhetoric of the source text. Appropriate conveyance of meaning is always present in the discourse of professional organizations when they define what the exams are targeting. This is different from language production, although many times the borders between the two areas get blurred. However, meaning is a very indirect measure of the grammatical competence of the candidate in the source language. It is difficult to call this a reliable measure of source language grammatical competence since the difficulty with target language production may also hinder successful communication of major and subtle meanings encoded in the language of the source text. If an organization wanted to measure language comprehension, which may impact the rendering of meaning, a more direct measure of source language comprehension that is not dependent on target language production would be preferable. Nevertheless, a successful communication of the meanings found in the source text may correlate highly with source text comprehension, although this would need to be demonstrated through empirical research. Obvious misinterpretations of the source text as evidenced in the target text may also be seen as possible indicators of problems with source text comprehension.

Target text style and cohesion is seen as being reflective of the candidate's knowledge of the ways in which texts are linked and organized into documents in the given target language genre, or document type within a given communicative setting. Knowledge of genre is seen as an ability to segment the document appropriately into sections and/or paragraphs, such as an introduction, statement of the problem, presentation of findings, discussion, proposals, etc. depending on the purpose and type of document being translated. Knowledge of cohesive devices and the rules for creating a coherent text is reflected in the flow of the document,

the degree to which it seems to form a single unit and how it addresses textual competence.

Situational appropriateness is a measure of the candidate's ability to employ the functional and socio-cultural aspects of the target language in their translation. The functional aspects of language are judged by the degree to which the target text is successful at achieving its intended target language use. The socio-cultural aspects of language are judged in the target text's use of appropriate markers of register, i.e. degree of formality in phrasing and word choice. It addresses pragmatic competence.

Grammar and mechanics is the category which includes spelling, diacritical marks, agreement, punctuation, and other conventions of the writing and grammar of the target language. It addresses linguistic competence in the target language. Together with meaning it is the category most frequently used by professional associations and schools while scoring a translation test.

Translation skill is meant to include the application of strategies to translation problems and the use of resource materials. This category is measured by how well the target text reflects the identification of translation problems and successful solutions to them. It also includes the degree of success or failure in the appropriate use of references in overcoming gaps in language or topic knowledge. (This may appear more clearly in the misapplication of resources.) It addresses strategic competence.

6. Levels of performance of a rubric

The rubric presented above in Section 5 was designed with a high professional standard for certification as its set point. The scale goes from 1 to 5. Number 1 is seen as a true lack of ability. It is imagined that very few candidates will score "1" at this level, given that they should have self-selected out of the examination. A score of "5" is seen as being indicative of particularly outstanding performance. The desired level of performance for certification is seen as being represented by "4" on this scale. Number 3 is seen as the point at which the candidate shows evidence of skill but falls slightly short of the proficiency level desired for certification. A number "2" on the scale represents a deficient performance that is clearly below the level required to perform as a certified professional translator. It is important to point out that organizations which currently grant certification (e.g. ATA, NAATI) focus on certifying candidates in terms of language pairs and directionality. That information is provided by stating "certified in Spanish into English," for example. Other than that, certification does not give any specific information about what types of texts a candidate can translate and/or in which

contexts. Even when some organizations may discriminate between certifying at a professional or paraprofessional level (e.g. NAATI), certification is generic. Therefore, the levels of the rubric simply point to distinct levels of performance that programs may need to show so that test results can be referenced to certification criteria in the event of an examination challenge. It is also believed that the use of this number of performance levels will be easily managed by graders. There is the possibility that more levels would be confusing for candidates and graders while the use of fewer levels would not allow for a clear delineation of competences upon which decisions about certification are made.

Conclusion

This paper explored what translation ability is, and how it may be measured in a meaningful way for the purpose of certifying translators. Certification examinations assess an individual's ability to function as a professional translator. This ability (test construct) can be defined as consisting of the following sub-components: linguistic competence, textual competence, pragmatic competence, and strategic competence. Because important questions are asked before conceptualizing a test, test designers are able to define the test construct based on specifications of professional organizations. As test designers engage in test development and consider what associations deem as important elements to be present in those tests, they also consider how to score them. One possible way to holistically score a test is by using a scoring rubric. Test designers develop rubrics based on the test construct sub-components. Once translation skills have been defined, it is agreed that knowledge of these skills and the ability to successfully apply them form the core of translation ability (test construct). A translation test, however, may not measure them all. It is therefore important to define *a priori* which parts (sub-components) of the construct are examined by the organization test, and then check for validity during and after the development of the test. All tests have consequences. In the case of certification tests or any other high stake test, the consequences for test takers are even more important. This is why extreme care should be taken to develop tests that measure what they are set out to measure (i.e. the construct, in this case translation competence) and that those tests measure translation competence in a valid and reliable way. Clear definitions of constructs as well as validity and reliability considerations constitute the basis from which we need to develop current and future translation assessment examinations.

Limitations and implications

The research presented here should be considered as work in progress. It is an attempt to put forward a way of measuring translation competence in a more systematic way. Until this rubric is put to the test and applied to exams of various language combinations, we cannot begin to discuss its value (or lack thereof) for professional organizations granting certification. In order to do so, we need to see more collaboration between professional organizations and researchers. As has been stated above, the little discussion on translation assessment has been done in the field of translation (and interpreting) and it is still obscured by the tension between theory and practice. Practitioners believe that expertise in testing is obtained by practical experience. Since they may not be aware of the consequences of not developing a test based on testing principles, or of not examining a test for its validity, reliability, authenticity or practicality, they continue testing candidates according to a status quo. In so doing, they are measuring other candidates' performances to the best of their abilities. This, however, may not be enough of a justification.

In sum, as this paper demonstrates, the literature on translation competence and measurement of translation suggests the current need of exploration by mainstream researchers of testing and assessment (i.e. researchers who have expertise in testing, even if they do not work in the field of Translation Studies), as they address the many complex issues surrounding the measurement of translation competence in various language pairs. For applied linguists working with testing and bilingualism, and for translation scholars working on testing, the literature on translation competence and translation measurement offers important insights about the linguistic, psycholinguistic and sociolinguistic characteristics of individual bilingualism, as well as about the testing practices used in the profession. This article concludes with a call for collaboration between practitioners and professional organizations, as well as researchers in translation and researchers in testing. This work has implications for translation teaching and testing, for translation teacher education and translation practice.

References

Adab, Beverly. 2000. "Evaluating Translation Competence". *Developing Translation Competence*, 215–228.

Angelelli, Claudia V. 2000. "Interpreting as a Communicative Event: A Look through Hymes' Lenses." *Meta* 45 (4): 580–592.

Angelelli, Claudia V. 2003. "Connecting World Collaborative Testing Project". Technical Report. Commissioned by Healthy House for a MATCH Coalition and The California Health Partnership and funded by The California Endowment.
Angelelli, Claudia V. 2004a. *Medical Interpreting and Cross-Cultural Communication*. London: Cambridge University Press.
Angelelli, Claudia V. 2004b. *Revisiting the Interpreter's Role: A Study of Conference, Court and Medical Interpreters in Canada, Mexico, and in the United States*. Amsterdam/Philadelphia: John Benjamins.
Angelelli, Claudia V. 2006. "Minding the Gaps: New Directions in Interpreting Studies". *TIS Translation and Interpreting Studies* 1 (1): 41–67.
Angelelli, Claudia V. 2007a. "Validating Professional Standards and Codes: Challenges and Opportunities." *INTERPRETING: International Journal of Research and Practice in Interpreting* 8 (2): 175–193.
Angelelli, Claudia V. 2007b. "Accommodating the Need for Medical Interpreters: The California Endowment Interpreter Testing Project." *The Translator* 13 (1): 63–82.
Arango-Keith, Fanny, and Koby, Geoffrey S. 2003. "Translator Training Evaluation and the Needs of Industry Quality Training." In *Beyond the Ivory Tower: Rethinking Translation Pedagogy*. Brian Bear & Geoffery S. Kobby (eds.) Amsterdam: John Benjamins Publishing Company, 117–135.
ATA Certification Program. 2008. (Active September 29, 2008). (http://www.atanet.org/certification/aboutcert_overview.php)
Bachman, Lyle and Palmer, Adrian. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
Bachman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
Beeby, Allison. 2000. "Evaluating the Development of Translation Competence." In *Developing Translation Competence*. Christina Schäffner & Beverly Adab (eds.) Amsterdam: John Benjamins Publishing Company, 185–198.
Cao, Deborah. 1996. "On Translation Language Competence." *Babel* 42 (4): 231–238.
Cohen, Andrew. 1994. Assessing Language Ability in the Classroom. 2nd edition. Boston: Heinle & Heinle Publishers.
Colina, Sonia. 2002. "Second Language Acquisition, Language Teaching and Translation Studies". *The Translator* 8 (1): 1–24.
Colina, Sonia. 2003. *Translation Teaching From Research to the Classroom a Handbook for Teachers*. Boston: McGraw-Hill Companies.
Colina, Sonia. 2008. "Translation Quality Evaluation: Empirical Evidence from a Functionalist Approach." *The Translator* 14 (1): 97–134.
Faber, Pamela. 1998. "Translation Competence and Language Awareness." *Language Awareness* 7 (1): 9–21.
Fulcher, Glenn. 2003. *Testing Second Language Speakers*. London: Pierson Longman Press.
Johnson, Marysia. 2001. *The Art of Non-Conversation*. New Haven: Yale University Press.
Kiraly, Don C. 1995. *Pathways to Translation: Pedagogy and Process*. Trans. Geoff Koby. Kent: Kent State University Press.
McNamara, Tim, and Roever, Carsten. 2006. *Language Testing: The Social Dimension*. Blackwell Language and Learning Monograph Series. Richard Young, ed. Malden, MA: Blackwell Publishing.

Messick, Samuel. 1989. "Validity". In R. Linn (ed.) *Educational Measurement*. New York: Macmillan, 13–103.
Moss, Andrew, and Holder, Carol. 1988. *Improving Student Writing*. Dubuque, IO: Kendall/Hunt.
NAATI. 2007. *Levels of Accreditation*. http://www.naati.com.au/at-accreditation-levels.html (Active September 20 2008).
National Association of Judiciary Interpreters and Translators. Webpage on National Judiciary Interpreter and Translator Certification (NJITCE). htttp://www.najit.org/ Certification/NAJIT_Cert.html (Active November 20 2008).
National Standards in Foreign Language Education Project. 2006. *Standards for Foreign Language Learning in the 21st Century*. Lawrence, KS: Allen Press, Inc.
Neubert, Albert. 2000. "Competence in Language, in Languages and in Translation." In *Developing Translation Competence*, 3–18.
Nord, Christiane. 1991. Scopos, "Loyalty and Translational Conventions." *Target* 3 (1): 91–109.
Nord, Christiane. 1997. *Translation as a Purposeful Activity: Functionalist Approaches Explained*. Manchester: St. Jerome.
Orozco, Mariana. 2000. "Building a Measuring Instrument for the Acquisition of Translation Competence in Trainee Translators." In *Developing Translation Competence*, 199–214.
Presas, Marisa. 2000. "Bilingual Competence and Translation Competence." In *Developing Translation Competence*, 19–31.
Sawyer, David B. 2004. *Fundamental Aspects of Interpreter Education*. Amsterdam: John Benjamins Publishing Company.
Schäffner, Christina, and Adab, Beverly. (Eds.). 2000. *Developing Translation Competence*. Amsterdam: John Benjamins.
Valdés, Guadalupe, and Angelelli, Claudia V. 2003. "Interpreters, Interpreting and the Study of Bilingualism". *The Annual Review of Applied Linguistics* 23: 58–78.
Walvood, Barbara E. and Anderson, Virginia J. 1998. *Effective Grading*. San Francisco: Jossey-Bass.
Weir, Cyril J. 2005. *Language Testing and Validation. An Evidence-Based Approach*. New York: Palgrave Macmilan.
White, Edward M. 1994. *Teaching and Assessing Writing*. San Francisco: Jossey-Bass.
Wiggins, Grant. 1998. *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA: Jossey-Bass, 184–185.