Improvements to NAATI testing

Development of a conceptual overview for a new model
for NAATI standards, testing and assessment

Report prepared by:

Prof. Sandra Hale
School of International Studies
Faculty of Arts and Social Sciences
The University of New South Wales

Contributing authors:

Ignacio Garcia
(University of Western Sydney),
Jim Hlavac (Monash University)
Mira Kim (University of NSW)
Miranda Lai (RMIT University)
Barry Turner (RMIT University)
& Helen Slatyer (Macquarie University)

For:

The National Accreditation Authority for
Translators and Interpreters (NAATI)

30 November 2012

Acknowledgements

Authors: Sandra Hale
Ignacio Garcia
Jim Hlavac
Mira Kim
Miranda Lai
Barry Turner
Helen Slatyer

Table of Contents

Table of Figures

Executive Summary

This report presents the results of the first phase of the project: "Improvements to NAATI testing. Development of conceptual overview of a new model for NAATI standards, testing and assessment". The project, commissioned by the National Accreditation Authority for Translators and Interpreters (NAATI), consisted of three stages: 1. Review of the literature and consultations with the different stakeholders through focus groups, interviews and questionnaires; 2. The work of five specialist working groups on issues relating to prerequisites to testing, specialisations, testing, assessment and technology and 3. The development of a new conceptual model.

The authors acknowledge NAATI's crucial role in the establishment of Interpreting and Translation as a profession in Australia, its status as an international leader in the accreditation of community interpreters and translators in multiple language combinations and its important relationship with Interpreting and Translation education and training. As part of this integral role, NAATI seeks to reflect on best practice. In response to this proactive imperative this report has been commissioned to review all aspects of the current system that must be addressed in order for NAATI to maintain and strengthen its position as a rigorous accreditation body. The report highlights the need for improvement in the areas of prerequisites to accreditation, validity and reliability of testing instruments, assessment methods and training of examiners. It is worth noting that these shortcomings are not unique to NAATI or to Australia. However, a number of certification bodies around the world are now beginning to address them and we believe it is time for NAATI to do the same. The authors commend NAATI for its willingness to review its practices and implement improvements amidst limitations of resources, logistical challenges and lack of universal support for its role.

The report makes 17 recommendations, which must be viewed within the framework of the new proposed model. Some recommendations will be easier to implement than others, and we acknowledge that any major changes to the current system will require time and adequate resources in order to be implemented.

Recommendations

1. That all candidates complete compulsory education and training in order to be eligible to sit for the accreditation examinations, in accordance with the new suggested model outlined in section 2.3, Table 7.

2. That NAATI produce an information package explaining the meaning of Interpreter and Translator, prerequisites for testing and expectations of potential candidates, including expected levels of language proficiency in English and the LOTE, as outlined in section 2.

3. That NAATI select (or devise) an on-line self-correcting English proficiency test to be taken by potential candidates for a fee, as part of the non-compulsory preparedness stage, as outlined in sections 2.3 and 3.1.

4. That NAATI language panels select (or devise equivalent) on-line self-correcting proficiency tests in the various languages to be taken by potential candidates for a fee, as part of the non-compulsory preparedness stage, as outlined in sections 2.3 and 3.1.

5. That an Advanced Diploma in any discipline (or equivalent) be the minimum pre-requisite for the Generalist accreditation, and a Bachelor's degree in any discipline (or equivalent) or a NAATI approved Advanced Diploma in Interpreting) be the minimum pre-requisite for Specialist accreditations, as outlined in section 2.

6. That the current levels of accreditation be replaced by a Generalist level (for both Interpreting and Translation) and Specialist accreditations for Interpreting, with a Provisional Generalist level with a sunset clause of 2 years, particularly for new and emerging and Aboriginal languages, as explained in section 2.

7. That the following specialisations be established for Interpreter accreditations: Legal, Medical, Conference and Business (see Table 7), with Legal and Medical having priority over the other two, as explained in section 2.3.

8. That NAATI move to computerised translator tests in the first place. Secondly, that test candidates undertaking computerised translator tests be allowed access to the internet while taking the test[1], taking account of security considerations. See section 3.5.2 and section 4.

9. That Interpreting tests be conducted live, as much as possible. Where this is not possible, that candidates be provided with video recorded interactions and that their performance be video recorded for marking. See section 3.5.

10. That Interpreting tests at the Generalist level for both spoken and signed languages include a telephone interpreting component consisting of protocols for identification of all interlocutors, confidentiality assurances and dialogue interpreting only. See section 3.5.1 and section 4.2.1.

11. That a validation research project be conducted to design the new testing instruments for Interpreting and Translation. See section 3.6.

12. That new assessment methods using rubrics (see Table 8) be empirically tested as part of the validation project.

13. That new examiners' manuals be written to reflect the new assessment methods to be adopted.

14. That NAATI review the current composition of examiners' panels to include more graduates of approved courses and fewer practitioners who hold no formal qualifications in Interpreting and Translation. See section 3.7.

15. That examiners undertake compulsory training before being accepted on the panel, and continuous training while on the panel[2]. See section 3.7.

16. That NAATI establish a new Expert Panel, with subpanels for the specialisations, to design the curricula for the compulsory training modules and provide guidelines for the final assessment tasks.

17. That NAATI continue to approve tertiary programs and encourage all applicants to take the formal path to accreditation where such is available for the relevant language combinations.

---

[1] This is being trialled by the American Translators' Association [ATA] and they have signalled their readiness to offer support and technical advice to NAATI working group members in regard to the introduction of logistic protocols and recently-developed software.
[2] For Aboriginal language examiners and possibly other languages of limited diffusion, training may be unrealistic in some languages due to literacy/numeracy considerations. In such cases we recommend that untrained examiners be partnered with a trained examiner, as explained in the report.

1. Introduction

On 8 April 2011, the National Accreditation Authority for Translators and Interpreters (NAATI), published the "Improvement to NAATI Testing. Expressions of Interest" document, seeking applications to conduct the first phase of a three-phase project. Phase 1 of the project was titled: Development of a conceptual overview of the elements of a new model for NAATI's standards, testing and assessment. According to the document, the project's aim would be "…to improve various aspects of NAATI's testing process and related matters", with special emphasis on issues relating to validity, reliability and practicality. This would be the first comprehensive review of the NAATI accreditation system since its inception in 1977.

A team of researchers submitted an expression of interest by the due date, which was accepted and followed by a more detailed research project proposal. The proposal presented a revised structure, with Phase 1 consisting of the exploratory phase of the project, envisaging Phase 2 to be the validation phase and Phase 3 the implementation and trial phase. The NAATI Board approved the research proposal for the first Phase in September 2011. The first Phase of the project commenced in October 2011 and ended in November 2012.

The research team comprised Professor Sandra Hale (University of NSW) as Chief Investigator; Dr Mira Kim (University of NSW), Dr Jim Hlavac (Monash University), Adjunct Professor Barry Turner (RMIT University), Miranda Lai (RMIT University), Dr Ignacio Garcia (University of Western Sydney) as co-investigators; Helen Slatyer (Macquarie University), Professor Claudia Angelelli (San Diego State University, USA), Professor Gyde Hansen (Copenhagen School of Business, Denmark) and Associate Professor Catherine Elder (Language testing research centre, Melbourne University) as consultants, and Associate Professor Jemina Napier (Macquarie University), Dr Michael Cooke (Aboriginal interpreting expert), and Marc Orlando (Monash University) as advisors to the project.

1.1 The project plan

Phase 1 of the project comprised three main stages. In Stage 1, the research team conducted a thorough review of the relevant literature, of current Australian federal and state language and interpreter and translator policies and of the results of the previous NAATI reviews. Consultations with the different NAATI Commonwealth and state government owners were held in the form of focus groups and individual interviews. It should be noted at this stage, that although all states were invited by NAATI to participate, one state and one territory chose not to participate. Based on the reviews and on the feedback received from the NAATI owners, questionnaires were devised to seek further consultation from interpreting and translation practitioners, educators, examiners and agencies.

Stage 2 of the project consisted of the work of five specialist working groups:

1. Group on Rubrics, descriptors and competency-based assessment
2. Group on Technology and interpreting testing
3. Group on Technology and translation testing
4. Group on test reliability
5. Group on pre-requisites and specialisations

Each group was led by one of the researchers who invited experts in each of the relevant areas to participate in the work of each group. Among others, the consultants and advisors also took part in the working groups (See Appendix 1 for group memberships).

Stage 3 of the project consisted of the analysis of the results of the previous two stages, the development of a new conceptual model and a set of recommendations.

Three detailed progress reports were submitted to NAATI at the end of each Stage. The reports were also submitted to the international consultants for their feedback. This final report consolidates the results of each of the three stages, which were already presented in the progress reports, and makes conclusions and recommendations. The report will be organised around three main themes: A conceptual model for a revised accreditation system, including pre-requisites, specialisations and paths to accreditation; Testing, including issues of standards, validity and reliability, test content and delivery and assessment; and recommendations to support and implement such a model. The last section will provide a summary of the recommendations that will be highlighted throughout the report and will suggest practical ways of implementing the recommendations.

1.2 Background

> Testing and accreditation lie at the interface of training and professional work. Australia is at the forefront in the field of T&I accreditation, and there is growing recognition of the importance of using accredited interpreters and translators, but in certain quarters (not least amongst some practitioners themselves) there remain questions as to the need for accreditation and the ability of accreditation tests to determine accurately a candidate's ability to work at a professional standard of quality outside the examination room. For the National Accreditation Authority for Translators and Interpreters (NAATI) to be perceived as a credible testing authority, there is a need for rigorous selection of examiners, workshops for examiners, and an ongoing review of standards, marking guidelines and individual panels. I am pleased to see that NAATI is working on all of these areas, and it deserves commendation for its work in what is often a difficult climate (Wakabayashi, 1996).

Wakbayashi's observation, made sixteen years ago, is to a large extent still true today. Australia has been praised by many for its developments in community interpreting, especially with regards to its nationwide accreditation system, government funded service provision and formal training in multiple language combinations. NAATI is unique in the world for a number of reasons, two of which are paramount: it is a national accreditation body with the laudable aim to accredit in over sixty international languages and forty five indigenous languages[3], and it is owned by the Federal government and all State and Territory governments. For these reasons NAATI has been internationally recognised, as very few countries have managed to have uniform systems that give credentials in so many languages.

As a public accrediting body, NAATI is essentially different from equivalent bodies in other countries, such as those in the US (for e.g. American Translators' Association, ATA) or the UK (for e.g. Chartered Institute of Linguists), which are 'owned' by members of the profession, as will be detailed in our review below. The main reason for the Australian governments involvement in NAATI's establishment and ownership was a response to the high levels of immigration after World War II who spoke languages other than English (Ozolins, 1991).

One of the most important government commissioned reports in relation to Interpreting and Translation in Australia is "The Language Barrier" report, also known as the COPQ[4] report, which recommended the establishment of a national council as an overall "standard setter" for interpreters and translators, working especially in community settings (1977:3), which was later to become NAATI. Thirty five years ago, the report found that "…employers tend to underrate the level of interpreters required, just as they have, over the past twenty five years, underrated the need for interpreters of any kind". The report further comments that the provision of

---

[3] The range of these languages is extremely varied: from major European, Middle Eastern and Asian languages to relatively low-volume developing-country languages such as Dinka, Nepali, or Tetum. By comparison, the CIOL conducts their DPSI tests in over forty languages. The ATA conducts translator testing in seventeen languages, in only eight of which is testing available in both directions.
[4] Committee on Overseas Professional Qualifications

incompetent practitioners "…may lead to the violation of the human and civil rights of those involved" (1977:14). This observation from the cited report highlights two important issues, which NAATI has had to grapple with:

1. To provide the largest possible pool of interpreters and translators in as many languages as are needed by the community and,
2. To ensure that those translators and interpreters meet minimum quality standards that are uniform across the country.

Those two requirements are necessary in order to ensure access and equity among all members of the Australian community regardless of language and cultural background. It is important to note, however, that the two aims stated above may work against each other: the desire to ensure minimum standards will inevitably limit the size of the available pool, especially in some languages. However, pressure from government to ensure the largest possible pool remains strong[5], and NAATI needs to be responsive to that pressure. On the other hand, it is also worth highlighting that these governments' interests in ensuring a large pool of accredited T&Is have generally not been reflected in the amount of funding they have made available to NAATI. NAATI thus has to do its job in an environment of constant financial limitations.

A common concern among the representatives of the Commonwealth and State government owners during the focus group discussion and interviews conducted for this project was the issue of increased costs associated with an improved system and of limited sources of funding. The main concern however, should be the relationship between accreditation and competence. The provision of accredited practitioners who do not meet the adequate levels of expertise to perform their required tasks adequately will only have a negative effect on those receiving the services, on those providing the services and on NAATI's credibility as an accrediting body, but more importantly, as the COPQ report states, will violate the basic human rights of those receiving the services. We do believe, nevertheless, that the current accreditation system has provided a benchmark, which has ensured a certain level of competence, normally distinguishing those with and without accreditation and those with Paraprofessional and Professional accreditation. However, we also strongly believe, that as praiseworthy as the current system is, it has shortcomings (many of which are common to many other similar credentialing bodies around the world) that must be addressed in order to progress to the next level of development. We therefore commend NAATI for their willingness to review the current system and to implement changes for its improvement.

There are multiple factors that contribute to the complex task of ensuring competence and quality. Among these are issues relating to pre-requisites to accreditation, test validity and reliability of the testing instruments and assessment models, and post accreditation checks. As previously stated, NAATI's desire to ensure that there is a concrete link between accreditation and competence has motivated the current review, with the aim to make recommendations to implement changes to the current system.

First of all, pre-requisites to accreditation are related to issues of bilingual competence, education and training and specialisations. One of the most salient shortcomings in the current system is the lack of a requirement for any type of education and training prior to attempting accreditation[6]. The COPQ report recommended very stringent educational requirements for interpreters and translators and stated that "…the linguistic and professional skills involved in

---

[5] This sentiment was prevalent among the representatives of the government NAATI owners during the focus group discussion. While they were all in favour of improvements to the system, they were all adamant about the need for quantity and quality at the lowest possible costs. This is a reality that cannot be ignored, and issues of efficiencies must be seriously considered.
[6] NAATI offers short test preparatory courses (of up to 2 weeks' duration), which are not compulsory. One such course was the New Interpreters Project funded by DIAC.

interpreting and translation normally require education and training at the tertiary level" (COPQ, 1977:3). This recommendation was taken up by NAATI in its course approval system, which continues to operate successfully to date. The original intention was that most accreditations would be through NAATI-approved degrees at universities (for the former Level 3[7]) or diplomas at institutions in the Vocational Education and Training (VET) sector (for the former Level 2[8]). Direct testing by NAATI was made available as a 'back-up' to this system (with a further avenue through recognition of overseas qualifications). This position is still supported by NAATI, as its current Chief Executive Officer stated in a presentation to ASLIA:

> The NAATI ideal is a practitioner who prepares for the profession by gaining tertiary qualifications and who holds the NAATI credential at the appropriate level to show they are practice-ready (John Beever, 2012)[9].

However, because courses in translating and particularly interpreting are staff-intensive and therefore expensive to run, the number of courses available, especially in the Higher Education sector, has tended to decline overall since the 1980s, when most states had a NAATI-approved Bachelor of Arts in Interpreting and Translation. The situation may have been exacerbated by NAATI continuing to test in the same languages for which courses were available (Ozolins, 1998). Three decades later, only one university has retained the undergraduate degree, with the others offering post-graduate awards. The VET sector, however, in response to the decline in undergraduate I&T degrees, began to offer Advanced Diplomas approved by NAATI at the Professional level (former Level 3) in the 1990s, thus blurring the distinction between degrees and TAFE diplomas that existed in the beginning. Just as importantly, the range of languages offered at universities has also tended to be restricted to those with the highest volume of demand, especially languages that have attracted high numbers of international students, such as Chinese, since the late 1990s. On the other hand, training in languages of greatest domestic need (e.g. the 'newly arrived' language communities as well as Indigenous communities in northern Australia) has generally (with the exception of some languages offered at some TAFE colleges, especially in Victoria and South Australia) been limited and ad-hoc, or not available at all. These factors have tended to make direct NAATI testing the *de facto* 'standard option' for the domestic market, so that the majority of local current practitioners would have gained their accreditation by this method. Although NAATI reports that in 2010/2011 (NAATI, 2010-2011), 70% of accreditations were obtained by course completion, that figure can only reflect the languages for which there are formal NAATI approved courses available, with Chinese[10] being the language with the highest number of graduates (mostly international students who return to their country of origin to practise). It is also worth noting that the majority of practising interpreters and translators (59%) who responded to our survey as part of this project, had been practising for more than 5 years, with 45% having over 10 years' experience and that 72% of translators and 66% of interpreters had gained their accreditation by sitting an external NAATI test. This tends to indicate that the current workforce is mostly made up of untrained practitioners, a situation we hope will change in the near future.

When candidates gain accreditation by testing, this is done on the basis of a single relatively short test. This situation where a single test (or a combination of such single tests) can potentially give access to the profession makes NAATI testing a distinctly 'high-stakes' issue. Significantly, this path to accreditation means that even when candidates are successful, there

---

[7] Current Professional Level
[8] Current Paraprofessional Level
[9] "Rediscovering our roots: Shaping our future". Address given by John Beever, NAATI CEO at the ASLIA National Conference August 25, 2012 Adelaide
[10] While no formal figures are kept or available to the public on the ratio of Chinese students to other languages, it is a well known fact among educators that international Chinese students predominate in I&T classrooms. To give an example, the current numbers for the Master's program at the University of New South Wales indicate that 82% of commencing students in Semester 2, 2012 were Chinese.

is no guarantee that they have the level of competence required to operate in the different fields, as many of the competencies are not tested in the current tests (see discussion in 2.1.2.1 below). Furthermore, candidates have usually had no training in issues such as the ethics and practice of the profession, and have no theoretical knowledge to underpin their practice. This 'gap' may partly explain the reason for criticism from various users about the 'poor quality' of some accredited T&Is, as expressed by comments from the government NAATI owners at the focus group discussion and by the I&T agencies who responded to the questionnaire. One of the members of the focus group expressed that they receive frequent feedback from service users that the quality of interpreters and translators can vary enormously between people with the same levels of accreditation, a comment that was also prevalent in the responses to the survey of judicial officers and tribunal members conducted by Hale (2011). An interesting result from our questionnaire to I&T agencies was that government agencies did not give preference to formally qualified practitioners and received the highest number of complaints as compared to private agencies who claimed to give preference to trained practitioners. This practice of ignoring formal educational I&T qualifications in the allocation of work was a common complaint from trained interpreters in previous surveys of Australian practitioners (see Hale, 2011; Ozolins, 2004). We must point out at this stage, however, that there are great differences across courses in terms of duration, content, resources and standards and that not all courses are likely to produce optimum outcomes either. Nevertheless, courses consist of a variety of activities, practical and theoretical content, assessment tasks and practicum opportunities that minimise the level of risk and have a higher chance of assessing individuals' competence levels more comprehensively. One extra layer of quality assurance is provided by NAATI's current monitoring system of approved courses, which we strongly support.

Our survey asked agencies to report on the feedback that they receive from clients on the performance of the practitioners they hire. The types of negative feedback received by agencies on Interpreting, fell into one of the following categories: breaches of the code of ethics (punctuality, impartiality, professionalism); lack of English language competence; lack of management skills and lack of specialist training (medical and legal). The last three are areas that are not currently tested in the NAATI Interpreter examination. The negative feedback on Translation related to incorrect approach (too literal); linguistic issues (grammatical, spelling errors); accuracy of content, register and style and lack of technical skills. These are all issues that could be minimised through language screening and education and training, yet most agencies did not consider training to be of much significance when allocating work. This feedback should of course be taken only as an indication of the perceived deficiencies in the market and can be useful in deciding what measures to implement to improve performance. The positive comments that agencies reported receiving from their clients, on the other hand, were very general in nature: "Very good", "very helpful", "very impressed"; which seems to indicate that the criticism is likely to come from those who are more familiar with interpreting and translation and with what is expected of professional practitioners.

Another limitation of the current accreditation examinations is that in neither translating nor interpreting, even at the Professional level, is the material highly specialised (although it can be situated in specialist areas); in other words, there is no testing of competence in specialised areas such as legal or medical. This aspect differs from some testing conducted overseas, in which specialist areas are specifically tested, using fully authentic texts, such as the Court Interpreting Certification exam conducted in the USA (see Appendix 14). It also differs from a number of formal courses in Australia that have components that specialise in medical, legal, conference and business interpreting already. The results of our questionnaire of practitioners confirmed the need for specialist training. The statement that received the lowest level of agreement by the sample of practitioners who gained their accreditation by sitting a test, was "I was well prepared to interpret in complex settings such as the courtroom, after passing the test". Having completed a formal course appears to have provided participants with greater confidence in all situations, including "I was well prepared to interpret in complex settings such

as the courtroom..." which received the highest level of disagreement from both groups, but was more positively accepted by the trained group. Having generalist practitioners working in highly specialist areas in Australia has led to dissatisfaction on the quality of services, and especially interpreting services in the legal system, where judicial officers and tribunal members have commented that the Professional NAATI accreditation level does not seem to guarantee the level of competence they require of interpreters working in such a specialised field.

Table 1: Quotations from judicial officers on the reliability of current accreditation standards

"NAATI 3 is the benchmark, and we aim for that, although I know it guarantees little in terms of quality"
"The standard of interpreters varies widely – even among those with the same level of NAATI accreditation…"
"My experience over the years is that the rules of qualification as an interpreter are not nearly stringent enough" (in Hale, 2011, p. 14).

As the quotes in Table 1 demonstrate, the judiciary do not seem to have complete confidence in the current accreditation levels as a reliable measure of competence for their purposes. The judiciary, therefore strongly support the introduction of specialist legal interpreting accreditation and compulsory training (see Hale, 2011 for the results of a national survey of judicial officers and tribunal members). Medical practitioners, especially specialists, have also complained about the inadequacy of interpreters working in Sydney (Hale, 2007b).

Another common complaint from users of I&T services concerns interpreters' lack of language proficiency (especially in English)[11]. We believe, therefore that these three pre-requisites to accreditation (adequate bilingual competence, generalist and specialist education and training) are crucial in making the first step to bridging the gap between accreditation and adequate levels of competence.

Secondly, it is also essential to comprehensively research the validity and reliability of the testing instruments (e.g. test tasks, scoring rubrics, rating procedures and the conditions in which they are administered). Over the years, the NAATI tests have been subject to anecdotal criticism over their perceived lack of reliability and validity. This criticism relates, in particular, to the perceived lack of consistent processes in test setting and scoring, both within and across languages, as well as to its inability to assess the competencies required by professional interpreters and translators. This anecdotal evidence was born out by the findings of the Rater Reliability Study undertaken in 2007 by NAATI to investigate claims of variability empirically (Slatyer, Elder, Hargreaves, & Luo, 2008). This research identified discrepancies in inter-rater reliability in some language panels, inter-rater reliability between language panels and problems in variability in some of the test tasks. Intra-rater reliability was generally acceptable. The qualitative findings relating to the study of rater behaviour indicated that some raters were confused in their interpretation of the rating criteria and descriptors and there was disagreement within panels about the relative weighting of errors. The study also found a strong tendency of raters to rate scripts holistically according to a binary pass/fail judgement, adjusting scores to align with their overall impression of the performance, notably in the case of 'borderline' performances. A strong culture of practice was also observed within some language panels, which may lead to a prioritisation of issues relating to the language pair of the panel rather than attending to the achievement of consistency across language pairs.

High-stakes tests such as the NAATI tests should be subjected to regular evaluation, through a rigorous research process, which measures the performance of the tests to ensure that they are fair. Traditionally, interpreting and translation examinations have not been subjected to the same rigour as language proficiency tests, for example, both in Australia and the rest of the

---

[11] This was found in previous research (Hale, 2011), and was corroborated by the results of our current survey of I&T agencies and of the focus group discussion, although Turner & Ozolins (2007) did not find the same results in their national survey.

world. This failing is increasingly being acknowledged around the world as more research in the field of interpreting and translation assessment is being carried out. It should also be noted that the multilingual characteristics of the tests pose a particular challenge in this regard.

Our survey results also indicated dissatisfaction with the adequacy of NAATI examiners, especially by practitioners. Indeed the statement "There should be compulsory training for all NAATI examiners" elicited the highest percentage of agreement from all respondents combined (84.5%). Below are some unsolicited open comments from practitioners about NAATI examiners that further elucidate this perception:

Table 2: Comments from surveyed practitioners on NAATI examiners

"Careful consideration needs to be given to the mix of examiners, especially if there are different dialects or variations in a language".

"CHOOSE EXAMINERS CAREFULLY".

"Thorough screening of examiners".

"Examiners should have ongoing training".

"Examiners themselves should be formally tested and trained".

"NAATI examiners should be retired practitioners who don't see candidates as competition once accredited".

"Perhaps we need some qualified professionals with improved theoretical skills on the examiners' board?"

We therefore argue that tests, marking criteria and descriptors must be empirically designed and validated and examiners adequately trained accordingly, with clear assessment guidelines provided.

The final step towards improving the link between accreditation and competence is to establish a post-accreditation re-validation or re-accreditation system, which NAATI has already begun to implement and we strongly support. However, such a system is meaningless if the original accreditation cannot be relied upon to ensure the competencies required of interpreters and translators to function in the different areas of expertise.

1.3 Policy Review

NAATI of course cannot be held responsible for all complaints about service quality. Many factors that impinge on quality are well beyond NAATI's control, such as issues of policy and policy implementation, working conditions and remuneration. We concur with the statement made by the Working Party that produced *The language barrier* report:

> The Working Party emphasises that its findings and recommendations depend for their effectiveness on the adoption by the Australian and State governments of an occupational classification that gives adequate recognition to the qualifications and contribution of the interpreters and translators at the various levels of skill. There is also an obligation on others using the services of interpreters and translators to recognise that the quality of services provided by tertiary trained personnel calls for commensurate remuneration (COPQ, 1977:4)

Any attempt to improve the current accreditation system must be supported by government policy. General policy principles state their commitment to ensuring access and equity for those members of the community who do not speak English well or at all. For example, Principle 2 of "The people of Australia" states that:

> The Australian Government is committed to a just, inclusive and socially cohesive society where everyone can participate in the opportunities that Australia offers and

where government services are responsive to the needs of Australians from culturally and linguistically diverse backgrounds" (DIAC, 2011, p. 5)

The Access and Equity Framework states the need to provide information in "appropriate languages" and "using interpreters" (DIAC, 2011, p. 14). The Community Relations Commission's Principles of Multiculturalism Act 2000 frames policies relating to the provision of services to migrants in NSW, identifying their objectives as:

> s12(b) access to government and community services that is equitable and that has regard to the linguistic, religious, racial and ethnic diversity of the people of New South Wales

> s12(c) the promotion of a cohesive and harmonious multicultural society with respect for and understanding of cultural diversity

> s13 1(i) to provide (whether within or outside New South Wales) interpreter or other services approved by the Minister (CRC, 2009)

A review of the current language policies around Australia, however, shows a high level of inconsistency across states and departments. Firstly, a number of entities with language or I&T policies do not mention minimum requirements for interpreters and translators, which means that non-accredited interpreters may be engaged. The Family Court of Western Australia, is the most extreme example, specifically stating in its policy that "For undefended divorce proceedings the assistance of a friend to act as interpreter is encouraged and will be sufficient for those proceedings" (Family Court of Western Australia, 2006, p. 1).

Another inconsistency is found in the use of terms such as 'qualified', 'professional', 'accredited', 'credentialled' to mean supposedly the same thing. Many policies state a preference for suitably accredited, qualified or professional practitioners, without specifying what such adjectives mean. For example, the Practice Manual for Tribunals states: "It is always preferable for suitably accredited interpreters to be used rather than family members or friends" (Council of Australasian Tribunals, n.d.); the Australian Government's Department of Health and Ageing recommends using professional interpreters in mental health care, but does not specify what the word professional means other than referring to the Translating and Interpreting Service (TIS) for telephone interpreting (Australian Department of Health and Ageing, n.d.). In some cases, the policy or protocol specify that an interpreter should be NAATI accredited, but the level of accreditation considered as minimum is not mentioned. As an example of this type of policy, the NSW Community Justice Centre's policy stipulates that "any interpreter used must be qualified and accredited", but it also allows for the engagement of community interpreters approved by the Directorate. These are defined as "a person who does not actually hold a current professional interpreting accreditation, but has worked as an interpreter for community organisations and can provide references for their work" (NSW Community Justice Centre, 2009, p. 2). The NSW Workers Compensation Commission puts the onus of ensuring that "individual interpreters are appropriately accredited or recognised" on the interpreter service provider (NSW Workers Compensation Commission, n.d., p. 3). Most policies make the allowance for the engagement of interpreters accredited below the Professional NAATI level (previously Level 3), recognising that this level of accreditation is not available in every language. For example:

> As far as is practicable, interpreters used for court work shall be accredited to a minimum standard of NAATI Level 3. It is acknowledged, however, that for a number of languages Level 2 accredited interpreters are the only ones available and in these circumstances this is acceptable (Federal Magistrates Court, n.d., p. 3).

The above policy makes an interesting exception for Aboriginal languages in remote locations, recognising the particular difficulties posed for access to interpreters for speakers of small language groups:

> Where accredited interpreters cannot be accessed by telephone or on site, suitable members of the community should be engaged as non-accredited interpreters. The rate of payment for non-accredited interpreters is set at 75% of the rate set for accredited indigenous interpreters provided by the Language and Cultural Centre, Alice Springs or other relevant provider of indigenous interpreter services.

> Before the services of a non-accredited interpreter are contracted, staff should be satisfied that the person has a clear understanding of their role and the requirement to have no conflict of interest in the case at hand; to keep the matter confidential and only interpret what is said in the course of the interview or community consultation. They should also agree not to advocate for any party (Federal Magistrates Court, n.d., p. 2).

The difficulty of finding suitably accredited interpreters in many Aboriginal languages is mentioned by many of the reviewed documents, but the Federal Magistrates Court policy is the only one to stipulate a differential pay rate for non-accredited interpreters, thus providing an incentive for interpreters to upgrade their accreditation.

The Federal Court specifies that it "will usually only accept interpreters who are accredited and registered with the National Authority for the Accreditation of Translators and Interpreters (NAATI)" and that it "will generally prefer accreditation to the level known as 'Professional Interpreter'" (Federal Court, n.d., p. 5). The other organisations that stipulate a preference for Professional accreditation, rather than a requirement, are the Migration and Refugee Review Tribunals. Two of the documents reviewed in the legal area make an incorrect assertion in relation to Professional NAATI accreditation, which leads the reader to believe that a Professional accreditation implies training. The Queensland Department of Justice and Attorney General asserts that "Professional interpreters *are trained* to maintain confidentiality, impartiality and accuracy as part of their code of ethics" (QLD Department of Justice and Attorney-General, 2009, p. 34) and the Northern Territory DHLGRS states that "Professional interpreters are bound by a strict code of ethics covering confidentiality, impartiality, accuracy and reliability, *and have completed training and assessment* to certify that they have the level of linguistic competence" [italics added] (NT Department of Housing Local Government and Regional Services, 2011). This is only true for interpreters who have gained their Professional accreditation by training, not for those who have gained it exclusively by testing, but the distinction is not made by the cited policies.

The highest level of requirement recommended by any protocol or policy document is to be found in Queensland for court interpreting. The *Equal Treatment Benchbook,* for example, recognised that NAATI qualifications on their own do not guarantee the specialisation required for court interpreting, and recommends engaging interpreters with Professional accreditation and court experience (Supreme Court of Queensland, 2005, p. 67). Through its Guidelines for working with interpreters (QLD Health Interpreter Service, 2007), the Queensland Health department sets the requirement to give preference to Professional accredited interpreters with health experience. The Department further stipulates that "within each level, there will be a preference for those interpreters that have participated in training about interpreting in a general health setting or a mental health setting" (QLD Health Interpreter Service, 2007, p. 6). South Australia Health establishes the requirement to engage only professional interpreters, defined as accredited by NAATI at any level or recognised (SA Health, 2006, p. 4). Although it does not stipulate the minimum level required, it states that an order of preference must be given to NAATI accreditation levels.

In Western Australia, Government policy which is also reflected in the guidelines provided by the District Court, specify that interpreting services must be provided by "professional interpreters and translators or persons who have completed an accredited interpreting or translating training course in all other situations[12]" (WA Office of Multicultural Interests, 2008, p. 6).

In our view, the confusion and inconsistencies present in the current policies reflect the current inconsistencies in the accreditation system, where trained and untrained practitioners can receive the same level of accreditation. This has also contributed to the confusion regarding the terms 'qualification' and 'accreditation'. We believe that a qualification implies the completion of a formal course of study and accreditation implies the credential awarded by a credentialing authority upon meeting that authority's requirements. In our opinion a streamlined system, where a minimum requirement for some training (albeit short and non-language specific in the case of languages of small diffusion), will apply to *all* accredited practitioners, will provide a much higher benchmark. In a system with compulsory training, where all practitioners will be qualified and accredited, we hope that all state and Commonwealth policies will support NAATI accreditation as the minimum standard. Similarly, we hope that government departments, and in particular the justice system and health care departments will demand the new NAATI accredited specialist interpreters as their minimum requirement. We also believe that the requirement for compulsory training will stimulate the demand for courses, which in turn will lead to their supply, which is currently limited.

2. A conceptual model for a revised accreditation system

Any discussion of testing and standards needs to be framed within the context of a whole accreditation system. In an attempt to gain a better understanding of how Australia's accreditation system compares with others around the world, and how it can be improved to raise the standards of NAATI accredited practitioners, we conducted a review of the literature of translator and interpreter accreditation/certification systems and processes around the world. We must highlight that not all countries have been reviewed in the literature and hence there are many countries that have not been included. We base our review on secondary data, as it would have been far beyond the scope of this project to conduct primary research into accreditation/certification systems.

2.1 Review of accreditation/certification systems around the world

This review is a survey of published information on a group of countries that have been selected due to the points of comparison and contrast that they offer to Australia's practices. Some details on other countries' accreditation conventions have also been gained from information gathered from governmental or professional organisations' websites. This review however, draws substantially from two sources: Stejskal (2005) and Turner and Ozolins (2007). Jiri Stejskal's *Survey of the FIT Committee for Information on the Status of the Translation & Interpretation Profession* which was undertaken for FIT (International Federation of Translators) as a 'state-of-the-affairs' survey, elicited information from 63 organisations from 40 countries. Much of the information is backgrounded by data collected by Stejskal from 2000-2004 on various countries and published in the ATA chronicle. The other source is Barry Turner's and Uldis Ozolins' (2007) *The Standards of Linguistic Competence in English and LOTE among NAATI accredited Interpreters and Translators,* which was a Review commissioned by NAATI. The 2012 "The status of the translation profession in the European Union" Report has also been consulted[13].

---

[12] We note that this policy is currently being revised.
[13] See http://ec.europa.eu/dgs/translation/publications/studies/translation_profession_en.pdf

2.1.1 Different terminology

Since not all countries use the same terminology to refer to the same concepts, it is worth explaining the differences at this point. In Australia, the word 'accreditation[14]' refers to a credential granted to a candidate who has either successfully passed an accreditation examination or has successfully completed a formal Interpreting and/or course (either an approved Australian degree or diploma or an overseas course recognised by NAATI). The same term is used by New Zealand and South Africa. Elsewhere (e.g. US, Canada, parts of Europe), the term 'certification' is used.

There is little difference in the meaning of the two terms 'accredit' and 'certify'. Both can refer to the awarding of (an official) recognition to a person or organisation. However, a distinction between the terms occurs in the labelling of authorities. In North America and increasingly in Europe, the body that issues formal recognition to individuals is a 'certifying body' that 'certifies'. Hierarchically, this body is subordinate to another authority that checks that the 'certifying body' is following required standards in issuing 'certification'. This higher authority is an 'accrediting body' that 'accredits' the certifying body. Thus, authorities equivalent to NAATI in North America are usually termed 'certifying bodies'. One interesting point is that NAATI does not answer to a single higher authority. NAATI answers to the nine governments of Australia, which are the highest authorities in the land and are part owners of it. However, NAATI examiners do not answer to an external authority. Whereas the examinations held at approved NAATI courses and the results awarded are monitored by NAATI examiners, NAATI examiners are not monitored by any external body.

There are other terms that come close to the meaning of 'certify'. In the UK, the term 'chartered' is used with 'linguist' to refer to a member of the professional association. 'Registered' is also used in the UK, in reference to those who have passed the Diploma of Public Service Interpreting. 'Sworn' is a commonly used term, particularly in countries in which the courts were the first or only authority that provided formal recognition of status and skill level.

Stejskal (2005) found that, internationally:

> …the credentialing process occurs under three possible scenarios: certification by a professional association; certification by a government; and certification by an academic institution. Certification by a professional association is strongest in common law countries, whereas certification by a government body is usually employed in civil law countries. Academic programs exist in both civil and common law countries, and are particularly strong in countries where certification is not offered by the government or professional associations' (p. 3).

As we stated above, in Australia, there sometimes seems to be some confusion between the terms 'accreditation' and 'qualification'. We believe that the term accreditation should continue to be used to indicate 'credentialing' from the national accreditation authority. The word 'qualification' should only refer to the completion of a training course. In other words, a person who is accredited by NAATI may possess a number of relevant qualifications. In our new proposed model, all NAATI accredited practitioners will need to have minimum qualifications before becoming accredited (see point 2.3 below).

2.1.2 Comparisons between different countries

2.1.2.1 Australia

---

[14] NAATI may wish to change its nomenclature to align itself with most other countries.

In Australia, the accreditation process for interpreters and translators is administered by NAATI. There are four accreditation levels for translation and interpreting, the titles of which are as follows:

- Paraprofessional Translator / Paraprofessional Interpreter;
- Translator / Interpreter;
- Advanced Translator / Conference Interpreter; and
- Advanced Translator (Senior) / Advanced Interpreter (Senior).

There are five ways to gain NAATI accreditation: 1. by passing a NAATI test; 2. by successfully completing a NAATI-approved translation and/or interpreting course (TAFE diploma, advanced diploma or University undergraduate or post graduate degree); 3. by providing evidence of overseas qualifications recognised by NAATI, 4. through membership of a recognised international association in translating and interpreting (e.g. AIIC) and 5. By providing evidence of advanced standing in translating or interpreting (NAATI website)[15]. NAATI also has a system of 'recognition' in languages for which there are no examination panels and therefore testing is unavailable.

2.1.2.2 Latin America

In Argentina, 'the only way to become a *traductor público* (sworn translator) is through a university degree. In order to be admitted to such a university degree program, prospective students must pass an admission language examination. The university programs vary from four to five years (Stejskal, 2002a, p. 13), and such university degrees are focussed mainly on legal translation. It is not uncommon for sworn translators to be lawyers as well. Sworn translators must be Argentine citizens licensed by a *colegio de traductores públicos* (province based professional registration boards), which have been created through a provincial law passed by province legislature. Registering with such a board allows translators to certify their translations, but only in the particular province where the board exists, and only 5 of the 23 provinces have a board. In the remaining provinces, translators must register at the courts. Thus, the accreditation process for becoming a translator in the legal field in Argentina does not involve sitting a single examination. For areas other than legal (e.g. technical, scientific or literary), there is no registration required, but there are 3 to 4 year university degrees that specialise in non legal interpreting and translation. Argentina is an example of a country with possibly the highest education requirements for translators in the world. Nevertheless, the number of languages is limited to the most popular European languages. Similarly, education for interpreters is also limited and the few available degrees focus on conference interpreting only.

In Mexico, the translation and interpreting industry is not regulated by a specific organisation, though most practitioners hold degrees in languages or in translation/interpreting (Cuevas, 2011). Legal translations must be carried out by sworn translators (*peritos traductores*), certified by the Supreme Court of Justice.

The requirements for becoming a sworn translator are as follows:

- A degree in translation or a related field or, in lieu of a degree, having proof of knowledge related to the field;
- At least five years of professional experience;
- Proof of good character; and
- No criminal record (Cuevas, 2011).

---

[15] http://www.naati.com.au/accreditation.html

The Mexican Organisation of Translators (OMT) offers a certification exam for experienced translators. This is not an official (governmental) accreditation and does not classify as a sworn translator certification. The OMT recommends that candidates hold a degree in translation and have a minimum of three years' experience (Organizacion Mexicana de Traductores, 2011).

In Brazil, the Brazilian Translators Association (ABRATES) has developed an accreditation program that tests professional skills, available to members only. To become members of this association, candidates must produce evidence of graduation in specific and recognised courses, as well as proven work experience. Candidates can choose to take specialist tests in literary, technical, medical-scientific, legal-commercial and general areas. So long as they remain affiliated with ABRATES, the approved candidates can make full use of the accreditation for a period of 10 years (Stejskal, 2001).

2.1.2.3 North America

Under this section we will review the two predominantly English speaking countries in North America, the United States of America and Canada. In the USA, the term used is 'certification' but the definition of this term varies widely from state to state, as there is no national system. It usually involves 'some combination of testing and training in a given industry, such as court or healthcare, and is granted by a recognized certifying body, usually a government entity or professional organization' (Kelly, 2007, p. 32). Universities and colleges that offer interpreting training issue certificates to students, but the possession of such a certificate does not mean that the holder is certified to work in the profession, just that they have passed the requirements for the institution's program; such cases fall outside of the meaning of 'certification' employed here.

Specialist Certifications in Translation and Interpreting

In the USA, medical interpreting 'has progressed from an ad-hoc function performed by untrained, dubiously bilingual individuals to a fledging profession concerned with standards of excellence and ethical practice' (Beltran Avery, 2003, p. 100). The National Council for Interpreting in Health Care was established in 1994. It published the *National Code of Ethics for Interpreters in Health Care* in 2004, but it was not until 2009 that the National Board of Certification for Medical Interpreters launched their process for National Certification. This is not yet a mandatory certificate for medical interpreters, but aims to become the national standard, and encourages hospitals to employ interpreters with this certification. (cf. National Board of Certification for Medical Interpreters, 2011).

Nationally, there are at least ten other interpreter certification programs with a focus on healthcare (cf. Roat, 2006, p. 13). The National Board of Certification for Medical Interpreters may eventually replace these, but a review of their various characteristics and strengths is useful in gaining a more complete picture of the current environment for certification in the US.

With regards to court interpreting, in 1995, the National Centre for State Courts created the National Consortium for State Court Interpreter Certification, a multi-state partnership dedicated to developing court interpreter tests. In response to the low passing rates, some certifying bodies have begun to include interpreter training as part of a certification program, but a university degree for certification is not a requirement (see Kelly, 2007).

In relation to interpreting testing in signed language, the US Registry of Interpreters of the Deaf is much more established than its spoken-language-interpreting counterparts. It has seven tests (reduced from the original 13): the Oral Transliteration Certificate, Certified Deaf Interpreter, Certificate of Interpretation, Certificate of Transliteration, the combined certificate (CI and CT), the Conditional Legal Interpreting Permit-Relay, and the Specialist Certificate: Legal. Five of the

seven are general in nature; the two industry-specific ones are legal. Sign-language interpreters must hold a generalist certificate before they can sit a specialist exam.

In Canada, university graduates are given preference by translation agencies, but there is no legal requirement for translators/interpreters to be certified. Despite this, some professional organisations offer certification, which can mean undertaking an exam to prove expertise, or can involve membership in that organisation, although successful completion of an exam is increasingly a prerequisite for membership in most provincial chapters of the Canadian Translators, Terminologists and Interpreters Council (CTTIC). Below are some examples of the types of certifications offered by individual associations.

The Society of Interpreters and Translators of British Columbia offers a number of certification examinations: Certified Translator, Certified Conference Interpreter, Certified Terminologist and Certified Court Interpreter. Certification examinations are not entry level tests; candidates must be in good standing, must have passed the society's ethics exam, and must comply with (a) or (b) before being able to sit the exam: (a) provide evidence of experience of four years (120,000–440,000 words of translation, depending on language); (b) hold a degree in the study of translation, linguistics, interpretation or language, plus one year of full-time experience.

The Association of Translators, Terminologists and Interpreters of Manitoba offers certification through either portfolio or examination. Certification through examination is arguably less arduous than by portfolio. The criteria for certification (membership) by portfolio are the following: a portfolio that includes written recommendations by two certified members, a copy of diploma/degree/certificate and mark transcripts; a CV, confirmation of length of employment from employers/clients, at least 10 examples of translation work (minimum 500 words each), proof and testimonials from employers/clients that the candidate is the author of said translations, and evidence of regular employment for five years (from 25,000–50,000 words per year, depending on language). Candidates for certification through portfolio must also pass the CTTIC Harmonized Code of Ethics Examination (see Association of Translators Terminologists and Interpreters of Manitoba, 2011).

In general, in Canada, there is separate testing and certification of court interpreters, which is now handled by the national body, the CTTIC. Certification in conference interpreting has been offered in Quebec, but is now being co-ordinated by the CTTIC as well. Otherwise, there is no distinction or grading of 'translator certification' and all candidates sit a 'generalist' examination. The term 'terminologist' is also commonly used in Canada to refer to often government-employed specialists whose job it is to have specialist knowledge and to ensure parity and equivalence in official terms used in the large volume of French<>English translation performed in or for government authorities.

2.1.2.4 Western Europe

Austria has a long tradition of multilingualism, dating from the time of the Austro-Hungarian monarchy. After World War II, Austria set up two large-scale T&I training centres at the universities of Vienna and Graz. Later, a program at the university of Innsbruck was set up. Formal certification does not exist as such, except in relation to court interpreting. Graduates of T&I programs list their qualification after their name and this is the accepted form of 'official certification' that exists in Austria. All recipients of a university qualification are obliged to bear this title with their name in official documents.

For court interpreters, a formal test exists which is open to graduates with evidence of two years' professional work, or to non-graduates with evidence of five years' work. Membership of the Association of Certified Court Interpreters is limited to three years. Renewal is allowed where there is evidence of continuing employment.

While there is no formal certification process in place, other than that for court interpreters, the changing ethnic composition of Austria and new T&I demands has led to a 'two-tiered' provision of T&I services. While there is a good supply of qualified T&I practitioners for English, French, Italian and Spanish, most of these are employed as specialist translators or conference interpreters. The needs of residents in Austria who speak only Albanian, Bosnian, Chinese, Croatian, Kurdish, Russian, Serbian and Turkish are less well addressed. While Bosnian, Croatian, Hungarian, Romanian and Serbian are offered at, at least one of the universities, few graduates are interested in community interpreting due to low remuneration. Recent EU laws which require public institutions to provide T&I services to clients without proficiency in the language of their country or residence have led to a large upsurge in the demand for T&I services. Consequently, similar to Australia, many of the T&I services for speakers of these 'rarer' or 'non European' languages are performed by lay, untrained interpreters.

In Germany, the qualification acknowledged throughout Germany is attained through sitting an exam held by any State Examination Office of a German federal state. Candidates can request papers from previous years before sitting the exam. Another widely recognised system of accreditation is attained through sitting an exam held by a Chamber of Industry and Commerce. To become a member of the Federal Association of Interpreters and Translators, candidates must have completed a translation or interpreting course in Germany or abroad, or have completed an exam at the State Examination Office (Federal Association of Interpreters and Translators, 2011). Germany has some excellent I&T training institutions, but they experience the same shortcomings as Austria. A mismatch occurs between the robust T&I provision for other western European languages and the lack of trained T&I practitioners for other 'migrant' languages spoken in Germany. The latter group is often trained but also recruited without training by municipal authorities – there is usually no formal testing to ascertain skill levels and standards in this group. There is a number of initiatives at the local level in Germany to remedy this unacceptable situation through organised and formalised means.

In Belgium there was no formal certification of translators and interpreters until 2003. In the Belgian Constitution it is stated that any citizen appearing before a court may address that court in the language of his or her choice. A list of 'sworn interpreters' is often kept by courts, drawn up in consultation with the public prosecutor's office, with each court having its own system for the recruitment and accreditation of its interpreters and translators. There is no national register of translators and interpreters, and the titles 'translator' and 'interpreter' are not legally protected. Belgium has a long history of T&I training and there are well-established university centres in Antwerp, Brussels and Gent. Common to other European countries, completion of a university degree (usually at post-graduate level) is still commonly accepted as a benchmark of ability. Practitioners commonly state their qualifications after their names or in advertising or correspondence to demonstrate their level of expertise. The university sector commonly focuses on T&I training in other European languages such as English, German, Italian and Spanish in addition to the two official languages, Dutch and French. Gent has recently increased its repertoire of languages to cover those of recent migrants, e.g. Czech, Russian and Turkish.

The large numbers of speakers of languages other than the four above-mentioned European ones in Brussels and in the northern region of Belgium, Flanders, precipitated the development in 2004, of the Social Interpreting ('community interpreting') test that includes a preliminary language proficiency test (in both languages) followed by 102 hours of compulsory pre-training before the main test which includes sight translation, consecutive interpreting and ethics. The development of the Social Interpreting test is of interest to the Australian context as it is intended to address recently migrated language groups, and at the same time, is predicated by Belgium's long history of T&I training and adoption of some of the features of the UK Institute of Linguists Diploma in Public Service Interpreting test.

In the Netherlands, candidates wishing to become sworn translators and interpreters must 'provide ample evidence to the court that they have a good command of Dutch and the pertinent foreign language, as well as provide a declaration of good conduct' (Stejskal, 2002b, p. 14). 'Ample evidence' differs from court to court. The sworn status is valid throughout the Netherlands and does not have a time limit, though can be recalled if the translator behaves inappropriately or incompetently.

In the United Kingdom, 'the focus is on the certification of *translations* rather than *translators*' (Stejskal, 2002f), and Ireland has a similar situation. In the UK, the Institute of Linguists (hereafter: IoL) is the organisation that co-ordinates and administers the language assessment and the awarding of accredited qualifications to interpreting candidates who pass a test at the end of a long period of non-intensive training and/or preparation. In the case of the IoL Diploma in Public Service Interpreting test (hereafter: DPSI), the final test is given five years after a candidate has fulfilled an initial minimum training requirement, i.e. a candidate has received a 'letter of credit' or 'unit certificate' as the first part of the diploma sequence. Thus, trainees undergo a long 'apprenticeship' but are still required to sit a final examination, which is a pre-requisite for the diploma to be issued. The IoL's DPSI has responded to new language groups that are now resident in Britain and testing is provided in languages such as Bengali, Cantonese, Croatian, Dari, Farsi, Gujurati, Greek, Hindi, Hungarian, Jamaican (sic), Kurdish, Latvian, Lithuanian, Polish, Punjabi, Portuguese (Brazilian), Portuguese (European), Pushto, Romanian, Serbian, Slovak, Somali, Swahili, Tamil, Thai, Tigrinya, Turkish, Ukrainian, Urdu, Vietnamese, as well as the traditionally popular European languages such as French, German, Italian, Russian and Spanish. The Diploma in Interpretation serves as a qualifying examination for membership in the National Register of Public Service Interpreters.

Similarly, the IoL Diploma in Translation consists of tests, which are assessed according to criteria very similar to those of the NAATI test for professional translators. The IoL does not award certification; it co-ordinates training and testing for diplomas. The IoL also has a category or list of practitioners that are termed 'chartered linguists' – practitioners who have gone through a five-year probationary period, must have a university degree, demonstrated expertise in T&I, and have three references. However, an IoL diploma is not an obligatory prerequisite for application to become a 'chartered linguist'.

The Institute of Linguists also serves as an examining body. It offers assessment and accreditation to suit higher-level candidates seeking a professional qualification. Its Diploma in Translation is a mixture of general and specialised translation, and the examinations have a high failure rate. Its Diploma in Public Services Interpreting is available in four options: health, local government, English law and Scottish law.

The other main organisation offering membership in the UK is the Institute of Translation & Interpreting (ITI). This offers different levels of membership to translators and interpreters throughout Europe and in other countries where English is commonly spoken. Levels of membership reflect varying amounts of experience. 'Qualified members' of ITI are not certified themselves, but can certify their translations.

The concept of a 'sworn translator' does not exist in the UK's common law system, but translations must be 'sworn/certified' for various purposes. Such translations have no bearing on the translation quality, but through this process the translators are identified and therefore can be held accountable for their work.

In Ireland there is the Irish Translators' and Interpreters' Association (ITIA), which is working towards standards of certification. Their 'professional members' have to go through the following process to be granted membership:

'Requirements for Professional Translators in the Commercial/Technical Fields

(1): Success in foreign examinations organized by the profession abroad and recognized by ITIA (within a period of five years preceding the application, plus one year of full-time professional experience in the same period); or (2): Award of a translation degree by an Irish third-level institution or similar foreign institution recognized by ITIA (within a period of five years preceding the application, plus one year of full-time professional experience in the same period); or (3): If the applicant is a staff translator, two years of professional experience substantiated by the employer's reference (within a period of five years preceding the application); or (4): If the applicant is a freelance translator, three years of professional experience substantiated by invoices, statements, or other recognized proof of work completed on a commercial basis (within a period of five years preceding the application, where it is estimated that the linguist translated at least 80,000 words in each of the above three years). The applicant has the option of submitting references, or, where discretion will allow, examples of work completed (to be treated in utmost confidence by ITIA). The association also reserves the right to administer a sample translation test. Literary/cultural translators are required to submit a portfolio of work they have had published, broadcast, or produced.

The requirements for Professional Interpreters are slightly different:

For bilateral interpreters: Three years of professional experience, estimated to be at least 40 days of interpretation per year and substantiated by invoices, statements, or other proof of work completed on a commercial basis (within a period of five years preceding the application).

For conference interpreters: Three years of professional experience, estimated to be at least 20 days of conference interpretation per year and substantiated by invoices, statements, or other proof of work completed on a commercial basis (within a period of five years preceding the application). The interpreter has the option of submitting references from clients. The association also reserves the right to administer its own test' (Stejskal, 2002f, p. 13).

2.1.2.5 Northern Europe

Sweden is a European country that has an important similarity with Australia. It was and remains the only European country to have an official policy on multiculturalism. This led to the large-scale provision of Swedish language classes to all immigrants and to the establishment of extensive T&I services, including community interpreting services. The official means of certification is after completion of a degree. The Swedish Legal, Financial and Administrative Services Agency (Kammarkollegiet) is responsible for the accreditation process in Sweden. The organisation sets an 'extremely demanding examination' that consists of a general, a legal and a financial test (Föreningen Auktoriserade Translatorer, 2011). Candidates who pass the examination can become authorised translators (*auktoriserad translator*), an authorisation that must be renewed every five years.

The *auktoriserad translator* title is protected by law and those who have it are subject to statutory rules on secrecy. Only those who have it can become members of the professional organisation *Foreningen Auktoriserade Translatorer* (Föreningen Auktoriserade Translatorer, 2011). Maintenance of a high level of ability and quality over time is a key concern of one of the organisations for translators in Sweden, SFO. For this reason, their admission procedure is strict and places emphasis on continuing education (Stejskal, 2002b, p. 15).

In Finland, candidates must pass a translation exam that has a general and a specialised component. Candidates must reside in one of the member states of the European Union or in another country included in the European Economic Area. The exams are administered by the Translator Examination Board, appointed by the Ministry of Education in conjunction with the Research Institute for the Languages of Finland.

Similarly, in Norway, government authorised translators must pass possess a three year university degree before sitting for their certification translation test (Stejskal, 2002d, pp. 13-14). In contrast with the formal nature of translation testing and certification, a Norwegian Interpreter Certification Examination was established in 1990 to address the need for community interpreting services in Bosnian, Croatian, Russian, Serbian and Spanish. Later, other languages such as Albanian, Arabic, Persian, Somali, Turkish, Urdu were added. The certification examination has been administered and conducted by the Linguistics Department of the University of Oslo since its establishment. This is an example of collaboration between academics, welfare authorities and government departments that were prepared to fund but not organise certification testing. The certification is intended for community interpreting only and the accounts of difficulties in creating training and testing materials for newly-arrived languages will be familiar to those in NAATI testing. Web-based training materials are being trialled as the low success rate and lack of preparedness of many test candidates has alerted testers to the need for comprehensive training before testing (University of Oslo, 2001). We see this as a common thread across all countries: those who do not have a compulsory pre-testing requirement inevitably find that the failure rates are too high and introduce some type of pre-testing education to remedy the situation.

2.1.2.6 Southern and Eastern Europe

As in Germany and Austria, a university qualification is the yardstick for quality in the T&I market in Spain and there are no other formal instances of T&I certification other than that required for court interpreters or 'sworn interpreters'. The term for a sworn interpreter in Spain is *intérprete jurado.* Prior to 1996, the interpretation in court was not recognised as having any legal effect (Stejskal, 2002e). There is now an official examination run by the Ministry of Foreign Affairs and Cooperation Office of Interpreting of Languages. Candidates for the exam must be citizens of a member country of the European Union and must possess a degree in translation and interpreting. Successful candidates become sworn interpreters once their name is published in the Official State Gazette (BOE) (Ministerio de Asuntos Exteriores y de Cooperacion, 2011).

In Ukraine, the Ukraine Translators Association has an accreditation examination and stringent membership requirements. To become full members, freelance translators are admitted after passing the exam, and interpreters should have a minimum of 100 hours of interpreting and client references. The certification procedure involves an exam and the resulting translation is not expected to be highly refined and polished (Stejskal, 2002g, p. 13).

2.1.2.7 Asia

In China, the most authoritative translation and interpreting proficiency credential is the China Accreditation Test for Translators and Interpreters (CATTI). The certificate awarded is called the Translation and Interpretation Proficiency Qualification Certificate of the People's Republic of China. This is the official credential, very similar to NAATI accreditation, and it is incorporated into the national system of professional qualification certificates, though those without certificates can still legally practice translation and interpreting. The certificate is one of the prerequisites for 'translation and interpreting professional and technical posts'. It has four levels, here given from lowest to highest: Level 3 Translator and Interpreter, Level 2 Translator and Interpreter, Level 1 Translator and Interpreter, Senior Translator and Interpreter. Those at the Senior level have to be experienced experts and have the responsibility of mentoring and training new interpreters and translators. At the other end of the scale, Level 3 practitioners have rudimentary skills and can only carry out generalist work. Other accreditations by other organisations include the National Accreditation Examinations for Translators and Interpreters (NAETI), the Shanghai Interpretation Accreditation (SIA) and the Accreditation for Interpreters and Translators (AIT) (Chen, 2009, p. 261).

In Japan there is no governmental involvement in accreditation, and there are multiple certification systems run by a number of organisations. The organisations include:

- Japan Society for Technical Communication (JSTC)
- Japan Translation Association (JTA)
- Japan Translation Federation, Inc. (JTF)
- Honyaku-Jutsumu Kyoiku Gakuin
- National Council of Professional Translators
- Business Education Academy
- Babel Co., Ltd.
- Japan Association for Technical Communication (JATEC)

The most authoritative examination of technical translation skills in Japan is run by the JTA. These are intended for the fields of natural sciences, social sciences and the humanities, and they comprise both a knowledge examination and a technical skill examination. There are four levels:

- Level 4: Practical translation experience in basic language skills, particularly in English, is required.
- Level 3: The applicant must be proficient in comprehension, language expression, and have fundamental specialized knowledge in practical translation.
- Level 2: For persons with more than three years of practical translation experience. Even though the person has not yet attained the level of a professional translator, it is necessary to have a skill level in which the person could become a professional with a small amount of correction.
- Level 1: For people with more than six years of practical translation experience. A mandatory requirement is to have not only a high level of proficiency in English expression, but also to be highly accomplished in written Japanese (Stejskal, 2002c).

Another organisation, Babel Co., offers an entry-level test for translators, 'English Translation Grammar Proficiency Test', and a test designed to evaluate the competence of professional translators, 'Professional Translation Proficiency Test' (PTPT). Each test comprises approximately 1,000 words. The categories of the PTPT include fiction (divided into romance and mystery), non-fiction, subtitles, law- and computer-related texts, and patent specifications. An interesting aspect of this system is that candidates take the test at home (see Stejskal, 2002c).

The Honyaku Kentei (Translator Qualification Examination of JTF) includes fields of specialisation in its certification: Politico-socio-economic; science and technology; finance and securities; medicine and pharmacology; information processing; and patents (see Stejskal, 2002c).

2.1.2.8 South Africa

The profession in South Africa is not currently regulated, but the South African Translators' Institute (SATI) began a system of voluntary accreditation, and only accredited members have voting rights within the institute. The system is widely recognised and is used as a recommendation or prerequisite for job applications by some institutions, but there are practitioners who have not been accredited. Accreditation is only offered at the professional level. The types of accreditation available are:

- Translation
- Sworn translation
- Simultaneous (conference) interpreting
- Language editing
- Terminology
- Corporate accreditation (for language agencies and language offices).

Accreditation of successful test candidates in South Africa is administered and conducted by SATI, which is a non-governmental authority that has limited resources and that is reliant on volunteer labour to maintain its operations. It is not clear what relationship SATI has with the six or seven post-secondary training institutions that offer T&I courses in South Africa. The limited resources of SATI have led to the adoption of en masse features from other programs' testing systems, e.g. the ATA examiners' guide, NAATI testing for languages such as Arabic, Chinese and Russian.

2.1.3 Recent initiative for global harmonisation of national certification and accreditation systems

Recent discussions have been held by an international consortium including members of our research team, to discuss the development of international standards to make interpreter and translator credentials portable. Such a move would require quality assurance measures on the organisations that grant such credentials, such as NAATI. There is currently no international standard that applies to how credentialing organisations around the world assess and award certification. This means that the certification a candidate receives is largely restricted to and recognised in the country that certification was gained. In the global T&I market it is difficult for consumers to assess what a practitioner's certification represents if it was gained elsewhere. There is a need for certification to be 'portable', i.e. for certified practitioners to be able to demonstrate that the certification authority from which certification was gained conforms to internationally set standards on what minimum requirements for certification are. Such a 'meta-standard' for the certification authorities would not necessarily result in a uniform type of testing and assessment structure that each certification authority in each country would have to comply with. Rather, the particular requirements that an international standard would set out would be global and relate to the processes that a certification authority should conform to. Nevertheless, it is a further reason for NAATI to align itself with the more stringent international practices that currently exist in some countries.

2.1.4 Conclusions on international accreditation systems

We learn from the foregoing review that there are generally two major paths to accreditation/certification: one through formal university education for the established European languages, by completing either undergraduate or post graduate degrees which can be up to five years in duration and specialist in nature; the other through testing in the 'rarer' or 'newer' languages of new migrant and refugee communities. The first path is the traditional approach taken by most European and Latin American countries whose main focus has been to train practitioners for international relations. The second path is usually the one taken by immigrant and refugee host countries, whose main focus has been to meet the rapid demand of the language needs of the local immigrant or refugee communities. As migration flows are now increasingly shifting to many more countries around the world, the initially contrasting positions of the two approaches are now converging: the immigrant host countries (such as Canada and the USA) are increasingly seeking to introduce medium-length training courses and university-level courses that precede or incorporate testing with a larger number of benchmarks for specialised T&I services; whereas many of the European countries are slowly starting to address the T&I needs of large immigrant populations through a combination of extending the languages available in existing university T&I programs, but more commonly through local- or regional-level training and testing initiatives. Australia is ahead of many in terms of available

education in what is known as "community interpreting", through NAATI approved courses. We strongly recommend that this situation continue to be strengthened and encouraged as the preferred path to accreditation. The majority of the respondents to our survey (73%) also strongly supported NAATI continuing to approve formal courses. However, the noticeable weakness in the Australian accreditation system lies in the voluntary nature of any training before attempting accreditation.

Some countries do not have certification bodies at all. This is mostly because they have not needed them, as their educational institutions performed the function of providing training and assuring standards. For those with accrediting or certifying bodies, these also differ in nature. In some countries, typically Anglophone countries of the New World and countries in East Asia, there are governmental or semi-official bodies that administer and usually also conduct testing for the awarding of certification (or 'accreditation' or 'registration') to T&I trainees or practitioners who can demonstrate minimum standards of ability and practice. In other countries, professional bodies take the responsibility of awarding the credential; and in others, such as Argentina for legal translators, there is a very highly regulated system where translators complete a formal degree in legal translation (of up to five years), register with a registration board and become government certified.

Another important difference is that in most countries there are generalist and specialist tests and training. These usually relate to court and/or medical interpreting, sometimes also conference interpreting, terminology and/or technical translation.

There are also fundamental differences in the underlying purpose of the certification test. In some countries, certification is granted to experienced practitioners, in other words, it is not an entry-level credential, but a recognition of high standing in the profession. This is sometimes ascertained via the compilation of a dossier/portfolio or evidence of long-standing practice, although these are not the most common avenues to certification. Some countries also have more flexible test delivery options, such as take home exams, permission to use the internet or to conduct on-line tests in the candidate's own time. Pre-testing language screening is also common with some systems. An annotated overview of accreditation/certification procedures in a number of the countries presented above in sections 2.1.2.2 to 2.1.2.8.

In light of the comparison with other countries around the world, we can see two important advantages to the current NAATI system: 1. Its uniformity as a national system, and 2. the availability of testing in many more languages than in other countries, including signed language for interpreting. However, we strongly believe that NAATI could improve in two most important aspects: the requirement for compulsory pre-accreditation education and training and the availability of specialisations. This view was strongly supported by the respondents of our survey, as illustrated by the quotes from two survey respondents in Table 3 below:

Table 3: Comments from survey respondents on the weaknesses of the current system

"The level of competence required at the interpreter/translator level cannot (and should not) be tested by one single exam. This is a ridiculous situation. Currently, we have the ludicrous situation whereby NAATI accreditation at the interpreter/translator level can be achieved by a 2-3 hour exam OR by successfully completing an approved NAATI course and passing the equivalent of a NAATI test at the end" (Survey respondent)

"In Europe it would be unthinkable to let a self-taught interpreter who sat a micky mouse test loose into the general public, and translators undergo lengthy training" (Survey respondent)

It is clear that survey respondents consider that NAATI must incorporate the need for training and specialisations into the accreditation system. We are conscious of the fact that people will continue to practise outside of the accreditation system, especially if the requirements for accreditation are made more stringent. However, we believe that in order for NAATI

accreditation to strengthen its status as a credible and reliable credential, it must only be awarded to those who can adequately prove they have reached the desired standards.

2.2 Results from consultations with interpreting and translation practitioners, educators, examiners and agencies on issues relating to pre-requisites and specialisations

2.2.1 National survey

As mentioned in the Introduction, as part of Phase 1, the team conducted three on-line questionnaires using the Key Survey software, which collected data from three separate groups: "Translation & Interpreting Agencies", "Examiners and Educators", and "Practitioners". Some participants requested paper copies of the questionnaire, which were supplied and later entered into the program. The Network or Snowball sampling technique was used for all questionnaires. The invitation to participate in the survey, containing a description of the project, was sent to a wide distribution list (see Appendix 2 for the full detailed list). The results generated by the Key Survey program were downloaded as Microsoft Excel and SPSS files for further quantitative analyses. The NVivo program was used to assist with the qualitative analyses of the open-ended responses. The questionnaires (see Appendices 3,4 & 5) consisted of three sections: section 1. "Demographic information", section 2. "Behavioural questions" and section 3. "Opinion questions". The results of the different sections are incorporated into the relevant sections of the report. This section will deal with the opinions obtained on issues of pre-requisites to accreditation and specialisations.

2.2.1.1 Demographic information

Twenty one (21) out of 36 agencies completed the questionnaire, a high response rate of 58%. Of the 21 respondents, 11 (52%) were private agencies, 7 (33%) were government, and 5 (24%) identified themselves as 'other'. The majority of the agencies (9 or 43%) were from NSW, which was followed by Victoria (5), Western Australia (3), Queensland (2) and 1 agency from the ACT and Northern Territory. Unfortunately, South Australia and Tasmania were not represented in this sample. There were a total of 95 examiners and educators who responded to the survey from all the Australian states and territories except Tasmania and the Northern Territory. The majority (76) were NAATI panel interpreting/translation examiners with 19 TAFE and University educators. There was one "Other", a professional interpreter and translator who served two terms on the NAATI panel. A further breakdown of the sample showed that 20 respondents were both panel members and educators, ten were University educators only and eight were TAFE educators only. Fifty-seven reported being only NAATI panel examiners. Not unexpectedly, the majority of respondents were from the major states, NSW (44 = 46% of the sample) and Victoria (26 = 30% of the sample). Queensland had eight (8%), ACT six (6%), Western Australia five (5%), and South Australia four (4%). The practitioner sample was the largest. There were 226 respondents of both genders from all the Australian states and territories (68: interpreters, 47: translators, 110: both interpreters and translators, and 1: teacher). The majority of respondents (116) were from NSW, with the next most represented states being Western Australia (41), Victoria (29), then Queensland (16), and South Australia (13). The territories were represented by 8 in the ACT and 1 in the Northern Territory. There was a single respondent (1) from Tasmania.

To compare the results of these three groups, a combined SPSS file of the common variables, namely, the respondents' demographics and their responses to "*Please indicate your level of agreement with the following statements*", was established. As stated above, the numbers of respondents to each survey were 21, 95, and 226, respectively, making a combined total of 342. While responses from NSW dominated with 169, Victoria and WA were well represented with 62 and 49, respectively. Queensland (26), SA (17), and the ACT (15) were better represented than NT (3) and Tasmania (1), as seen in Table 4 below.

Table 4: State and territory of residence

State or Territory of Origin	I & T Agencies	Examiners & Educators	Practitioners	Total
NSW	9	44	116	169
NT	1	0	2	3
QLD	2	8	16	26
SA	0	4	13	17
TAS	0	0	1	1
VIC	5	28	29	62
WA	3	5	41	49
ACT	1	6	8	15
Total	21	95	226	342

A listing of the statements that respondents were asked to rate using a five point Likert scale of '1. Strongly Disagree, 2. Disagree, 3. Neutral, 4. Agree, or 5. Strongly Agree', appears in Table 5.

Table 5: Opinion statements common to all three respondent groups

Interpreters should complete compulsory training before being accredited.
Translators should complete compulsory training before being accredited.
A minimum amount of experience should be mandatory before Interpreters can be accredited.
A minimum amount of experience should be mandatory before Translators can be accredited.
NAATI accreditation should not be necessary if an I&T education program was completed.
NAATI should continue approving training programs that lead to accreditation.
There should be mandatory specialised training for all legal, medical and conference interpreters.
There should be mandatory training for all NAATI examiners.
There should be different types of accreditation according to training and specialisation.
Translators and interpreters should undertake continuous professional development.

Overall, there was support for most statements by the majority of the respondents. Some statements were overwhelmingly supported by over 80% of all respondents. These appear in Table 6 below in pink and include compulsory training for NAATI examiners (84.5% agreement), compulsory training for specialist interpreters in legal, medical and conference settings (84% agreement), continuous professional development for all practising interpreters and translators (81.6% agreement) and compulsory training for interpreters prior to accreditation (81.3% agreement). The next most popular statements, with agreement levels of over 70% were that NAATI should continue to approve training programs (73%) and that translators should also undergo pre-accreditation compulsory training (72%). Over 60% agreement was obtained for two statements: different types of accreditation according to training and accreditation (66.4% agreement) and a minimum amount of experience required of interpreters before being accredited (65.2% agreement). The same statement for translators received slightly less agreement at 58.2%. The statement that received the least amount of agreement was that NAATI accreditation should not be necessary when an Interpreting and Translation formal course has been completed (35% agreement).

Table 6: Combined results (n=342) by level of agreement

	Disagree	Agree
There should be mandatory training for all NAATI examiners	18	289 (84.5%)
There should be mandatory specialised training for all legal, medical and conference interpreters	17	287 (84%)
Translators and interpreters should undertake continuous professional development	18	279 (81.6%)
Interpreters should complete compulsory training before being accredited	24	278 (81.3%)
NAATI should continue approving training programs that lead to accreditation	43	250 (73%)
Translators should complete compulsory training before being accredited	38	246 (72%)
There should be different types of accreditation according to training and specialisation	46	227 (66.4%)
A minimum amount of experience should be mandatory before Interpreters can be accredited	50	223 (65.2%)
A minimum amount of experience should be mandatory before Translators can be accredited	65	199 (58.2%)
NAATI accreditation should not be necessary if an I&T education program was completed	155	120 (35%)

2.2.2 Consultation with Aboriginal Interpreter Service

Due to the low response from the Northern Territory obtained through the national survey, a special consultation was organised with representatives of the Aboriginal Interpreters Service (AIS) in Darwin, Northern Territory. It must be noted at this point that translation is of little concern to the AIS, as the languages it services are mostly oral languages, so all the feedback received was in relation to interpreting accreditation.

Currently the AIS conducts pre-accreditation language screening, training and assessment, and some post accreditation monitoring of their practising interpreters. They have a mentoring system between senior and junior interpreters as well as a tiered system with differential pay scales. For this reason, the group was in favour of compulsory training and language screening before accreditation. They were also very much in favour of the different specialisations for interpreters, especially legal, but also medical and conference interpreting. The legal interpreting specialisation is their main priority and they are currently in the process of devising legal interpreting training modules, in conjunction with the TAFE diploma. Aboriginal interpreters are also often required to interpret in conference-like settings and high-level meetings with government, for which training in conference interpreting would also be very valuable. They welcomed a change to the current accreditation levels to give their interpreters a higher chance of success. They heavily criticised the current professional level examination for not testing the skills that are required of interpreters in professional practice. They stressed, however, the need for them to maintain some flexibility to cater for their interpreters' needs. One example of flexibility would be the adaptation of the requirement for Sight Translation in the examination, which would not be applicable to their languages. When confronted with such situations in the courtroom, for example, Aboriginal interpreters can ask the lawyer or judicial officer to explain or read the document aloud so they can interpret it orally rather than having to read it. They were very much in favour of improvements to the expertise of NAATI examiners, but were also aware

of the extra costs that would be required for training examiners and educators and would like NAATI to fund such extra training if required.

2.3. Suggested conceptual model for an improved accreditation system

Based on the work of the pre-requisites and specialisation working group (see Appendix 1 for a list of group participants), the review of the literature and the feedback received from the consultations with all the different interested parties, as outlined above, we propose the conceptual model for a new accreditation system which appears below as Table 7. All recommendations found in this report are therefore framed within the context of this new conceptual model.

As stated above, we believe that it is no longer appropriate for NAATI to continue to accredit candidates who have not undergone *any* Interpreting and/or Translation training. Such a practice is inconsistent with the existing body of research on the advantages of training (see Berk-Seligson, 1990 / 2002; Cambridge, 1999; Chacón, 2005; Ebden, Carey, Bhatt, & Harrison, 1988) and is strongly rejected by the majority of interested parties in Australia, as evidenced by the results of this project's national survey and other consultations, as well as the results of previous research and previous reviews. The review of certification/accreditation systems across the world also showed that for some types of interpreting and translation, especially for the legal specialisation, there are very stringent educational requirements in place (e.g. Argentina). It also highlighted the tendency towards some type of pre-testing training in the cases where formal education is not available (e.g. community interpreting certification in Belgium). The 1977 COPQ report strongly advocated for compulsory tertiary training for interpreters and translators, as cited above. We agree with the desirability of formal higher education in Interpreting and Translation, as is the practice in many countries for the well established European languages, but understand that such a requirement would be unrealistic in Australia for all languages. We believe, however, that some compulsory interpreting and/or translation training in the way of flexible modules delivered mostly in English is a feasible alternative for the languages for which formal courses are unavailable. We see no valid reason for allowing accreditation without any form of Interpreting and/or translation training. Our recommended model proposes pre-testing training as obligatory and not optional. Candidates can of course present a case for equivalence and not be required to undertake any further training if equivalence is established, as will be explained below.

Further, in almost *all* occupations in Australia, even ones that have been considered 'unskilled' or 'semi-skilled', training is now an *obligatory* condition for employment. For example, security guards (including unarmed ones) cannot gain employment without a Certificate II in Unarmed Guard and Crowd Control[16], which goes for between 3-4 weeks full-time. A personal services assistant in a hospital or clinic (who has no physical contact with patients) requires a Certificate III in Health Services Assistance[17], which is 6 months, full-time at most reputable VET providers. The minimum requirement for any childcare worker is the Certificate III in Children's Services[18], which is also a 6-month, full-time course. The minimum requirement for an integration aide working in a school is Certificate III in Education Support[19], which is a 6-month part-time course. In Victoria, in order to become a taxi driver, non-native speakers of English must do an IELTS or ISLPR test, and all applicants must complete Certificate II in Driving Operations[20], which includes the "Knowledge of Melbourne" test, which has a failure rate of over 70%.

---

[16] http://www.ista.com.au/vic_prs20103-crowdcontroller.asp
[17] http://www.kangan.edu.au/tafe-courses-melbourne-victoria/certificate-iii-in-health-services-assistance-psa/aosc/1723/
[18] http://www.kangan.edu.au/tafe-courses-melbourne-victoria/certificate-iii-in-children-s-services/aosc/1914/
[19] http://www.kangan.edu.au/tafe-courses-melbourne-victoria/certificate-iii-in-education-support/aosc/1821/
[20] http://www.deca.com.au/coursedetail/Car_Driver_Training/Taxi_Driver_Training_Victoria/Certificate_II_in_Driving_Operations_Taxi

These levels of training are *minimum* levels, and in many fields of employment such as childcare and healthcare services, applicants for employment require a level of training that is higher (e.g. Cert. IV or Diploma) as employers now require these further levels of training as a condition for employment. The above examples show that pre-employment training is now almost universal in the Australian labour market. Other professionals, such as nursers or migration agents, which in the past had not required any training, have also moved to compulsory pre-registration university training. We believe it is time for interpreters and translators to adopt a similar stance.

We understand that there may be concerns about access and equity. Access and equity can be seen from two view points – from the point of view of the non English speaker receiving the services and from the point of view of the bilingual person desiring to become accredited. As stated above, we believe that in order for all non English speakers to have equal access to all services there must not only be an adequate supply of interpreters in the required languages, but those interpreters must be competent to perform the required tasks. Providing the service of inadequate interpreters will not fulfil the requirement for access and equity. This supports the requirement for pre-testing training, especially for Interpreting. The second point refers to the means through which a trainee has access to a training course.

Concerns about access to training can perhaps be best addressed through targeted funding and subsidising of courses for particular language communities. For example, in Victoria, state government funding through the Office of Multicultural Affairs and Citizenship, provides bursaries for entry-level (community) interpreters in short courses conducted by Monash University, and bursaries for target language communities (e.g. Dari, Assyrian, Dinka) for training conducted as part of RMIT's Diploma of Interpreting. In Sydney, the University of Western Sydney conducted a fully funded similar course for the languages for which there is a shortage of interpreters (see Hale & Ozolins, forthcoming for more details). More such opportunities, in addition to non language specific programs such as the one offered by Macquarie University, non award courses offered by all universities with NAATI approved I&T courses and other courses that can be designed specifically for this new model, can be presented as a list of different possibilities for training to meet the pre-testing requirement of our new proposed model.

The issue of interpreter supply can be addressed in many different ways, the main one being through improving efficiencies of services, which is of course beyond NAATI's scope. Nevertheless, it is worth noting at this point that current practices show that interpreting services are often not being used in the most efficient ways, with interpreters waiting for hours in waiting rooms or multiple interpreters providing interpreting for different people of the same language combination[21]. The use of technology (such as simultaneous interpreting equipment where two interpreters can interpret for many speakers, or video conferencing facilities to allow interpreters to interpret for remote areas) can be some of the ways demand could be met. In our opinion NAATI should not be concerned with issues of service provision, but with ensuring high standards. We believe that a smaller but better qualified workforce of practitioners (especially specialist interpreters), who will service all of Australia, will lead to a higher volume of better paid work for practitioners, which in turn will justify any extra costs involved with training.

We therefore recommend two pathways to accreditation, both containing a minimum requirement for training and the same accreditation examinations. Such a universal requirement for pre-testing training would consolidate NAATI's position as international leader as an accreditation body.

---

[21] Preliminary results of a research project funded by the Australian Research Council Linkage grant scheme on court interpreting.

The two pathways are outlined below:

1. Accreditation via completion of a formal NAATI approved course of study, either through the VET or Higher Education sectors, as currently instituted. The final NAATI accreditation examinations are administered at the completion of the training and monitored by NAATI, as is currently the case.

2. Accreditation through a staged approach that includes pre-testing compulsory training modules and other pre-requisites as outlined in Table 7 below. The training modules can be delivered by educational institutions but the final NAATI accreditation examinations should be administered by NAATI.

We propose that the current levels of accreditation be changed to only one level for Translation and two levels for Interpreting: a generalist accreditation and specialist accreditations in in the legal, medical, conference and business settings, with priority given to the first two specialisations. The decision to have specialisations in Interpreting only was informed by the international practices as well as by the high level of support for interpreting specialisations but *not* for translation specialisations in the results of our survey and other consultations as well as previous research (Hale, 2011).

These changes would remove all the other current levels as they currently stand, except for Recognition, the recipients of which will also be required to complete the compulsory training modules. However, we propose that all these changes not be applied retrospectively. We recommend that the current holders of Recognitions be encouraged to complete the non-language specific training modules, even if accreditation examinations in their languages are not yet available. Similarly, the current holders of Paraprofessional and Professional accreditations who have not received any training, should also be encouraged to complete the training modules, and later attempt the specialisations. Those NAATI accredited professionals who have already undertaken specialist training in legal, medical, conference and business interpreting (either in Australia or overseas), should be encouraged to attempt the specialist accreditation examinations directly, without the need to undergo further training, unless they wish to do so.

Academic pre-requisites will differ according to the level of accreditation. A minimum of an Advanced Diploma (or equivalent[22]) is to be compulsory for the Generalist accreditations and an undergraduate degree (or equivalent[23], including an Advanced Diploma in Interpreting[24]) for the Specialist accreditations.

Although two parallel systems are proposed, currently approved formal interpreting and translation programs would need to be adjusted to align with changes and improvements to the required contents of training, and to the structure and content of testing instruments and assessment criteria. Later sections of this report will deal with issues of standards, testing and assessment criteria. However, as pointed out in the introduction, we restate at this point, that

---

[22] Equivalence of an Advanced Diploma is established by indicating a combination of other short courses plus recognition of prior learning (RPL). A case for RPL needs to be presented by applicants with supporting documentation to evidence, for e.g. experience in related fields, letters of support from community members. This is a point that will need to be refined before the implementation of the new model.

[23] A Bachelor's degree or equivalent is a common requirement for university post-graduate degrees. Equivalence of a bachelor's degree is established by indicating a combination of other qualifications plus recognition of prior learning (RPL). As per the previous footnote, a case for RPL needs to be presented by applicants with supporting documentation to evidence, for e.g. experience in related fields, letters of support from community members, completion of short courses, etc. This is a point that will need to be refined before the implementation of the new model.

[24] The research team did not reach consensus on this point. Two members of the team advocated for an Advanced Diploma in Interpreting to be considered equivalent to a Bachelor's degree and acceptable as a pre-requisite for the Specialisations. The others insisted on a Bachelor's degree.

any new testing instruments and assessment criteria that are proposed must be subjected to proper validation via a comprehensive validation study in Phase 2. Failure to submit new testing instruments to adequate validation will potentially lead to a new flawed system. Consistent with the guidelines of the International Language Testing Association (ILTA), high stakes test management has the responsibility to provide information, which allows valid inferences to be made. All tests, regardless of their purpose or use, must be reliable. This means that test results must be consistent, generalizable and therefore comparable across time and across settings (ILTA, 2012).

The ILTA Guidelines for Practice, which reflect the principles of the American Psychological Association's Standards for Educational and Psychological Testing, explicitly state that testing bodies have responsibility to provide comprehensive and accurate information to test stakeholders. Some principles relevant to NAATI are outlined below:

Institutions (colleges, schools, certification bodies etc) developing and administering entrance, certification or other high stakes examinations must:

- utilize test designers and item writers who are well versed in current language testing theory and practice.
- publish validity and reliability estimates and bias reports for the test along with sufficient explanation to allow potential test takers and test users to decide if the test is suitable in their situation.
- publish a handbook for test takers which

1. explains the relevant measurement concepts so that they can be understood by non-specialists.
2. reports evidence of the reliability and validity of the test for the purpose for which it was designed.
3. describes the scoring procedure and, if multiple forms exist, the steps taken to ensure consistency of results across forms.
4. explains the proper interpretation of test results and any limitation on their accuracy.

These requirements pre-suppose that testing organisations undertake the relevant research to be able to fulfil their obligations to test-stakeholders.

We also note that special courses will need to be designed and introduced by the different educational institutions to cater for the new training needs of Pathway 2, although, as we have already pointed out, many of the subjects currently offered by the different institutions could be offered to candidates as 'non-award courses'[25] to fulfil their training requirements under this model[26]. The new modules could be delivered by distance as well as face-to-face to cater for candidates in all languages from across Australia.

The proposed conceptual model attempts to bridge the gap that currently exists between trained and untrained practitioners, by ensuring that *all* accredited practitioners meet minimum standards for language proficiency, interpreting and translation competencies and knowledge of interpreting and translation underlying theoretical principles, including issues of professional ethics, in order to make informed choices to underpin their practice. We strongly believe that

---

[25] The new proposed Expert Panel who will write the curricula for the compulsory modules will also establish a list of equivalents with current university and TAFE courses.
[26] For example, most institutions have a theory subject that could be taken by candidates. Some institutions also have specialist subjects such as Legal interpreting, medical interpreting or conference interpreting that could also be taken by candidates.

standards relating to T&I practice that are, at present, untested (for e.g. *introductory and role establishment protocols*, *management skills, simultaneous interpreting* -for interpreters-, *assignment preparation, compilation and organisation of glossaries* -for both interpreters and translators- *use of computer assisted translation software, management and security of translation text files* -for translators, to name just a few) will be enhanced through training at the generalist level. We further contend that specialised training is essential for the different interpreting areas, but especially for court interpreting. The current NAATI Interpreter examination does not test most of the skills and knowledge required of court interpreters (e.g. understanding of the strategic use of questions in examination-in-chief and cross-examination, court protocols, simultaneous interpreting, specialised legal terminology and structures, etc). The tiered model of generalist and specialised testing and accreditation levels will ensure that future training courses conform to current 'minimum' standards of professional level accreditation and the model will ensure that for specialisations, further training needs to be of a level beyond that of the current professional level. Overall, this is a model that provides for clearer and more accessible pathways to specialised testing, and this will lead to an improvement of practitioner standards inasmuch as these further training, testing and accreditation levels are completed. The current short accreditation examinations will be complemented by a minimum set of hours of training and by hurdle tests throughout and at the end of each training module. Accreditation will no longer be seen as just being the result of a one-off accreditation examination, which as we have discussed above, cannot possibly assess all the relevant knowledge and skills required of I&T practitioners.

We anticipate concern from some stakeholders about increasing the level of difficulty for those who currently only hold the Paraprofessional level of accreditation. We argue that although the level of difficulty will increase, candidates will be much better prepared to acquire the necessary knowledge and skills in order to have a higher chance of success than is currently the case. We also argue that stages 0 and 1 will filter out those candidates who should not be attempting accreditation at all due to their lack of the necessary linguistic skills and other relevant knowledge. We also propose a Provisional Generalist accreditation (with a maximum 2 year duration) for those candidates who do not achieve the minimum pass mark for the Generalist examination. Candidates will need to re-sit the examination before the two years are up, after having practised in the field and conducted further training.

The new proposed conceptual model would consist of five stages, with a voluntary pre-stage we have called stage 0. The objective of a stage 0 is to ensure that potential candidates understand the basic requirement for adequate bilingualism before they invest more time and money into progressing any further in the process. This stage would not be compulsory but would be highly recommended to all aspiring candidates. Stage 1 would require at least an Advanced Diploma (in any discipline) or equivalent to ensure that candidates have a minimum level of academic background necessary for the type of skills and competencies required of professional interpreters and translators. For the Specialist levels we recommend that the minimum requirement be a bachelor's degree (in any discipline) or equivalent or an Advanced Diploma in Interpreting[27]. Equivalence can be established in different ways and the details can be agreed on at a later time. However, below are some examples of what may constitute equivalence for a bachelor's degree:

Example 1:

- TAFE Advanced Diploma plus
  - Related professional experience
  - Other professional development courses

---

[27] No consensus was reached on this point, so we offer both options.

Example 2:

- A series of short courses amounting to an equivalence of an Advanced Diploma
- Recommendations from members of the community
- Related professional experience

Stage 2 would comprise the compulsory education modules, which would prepare candidates to sit for the Generalist accreditation examination at Stage 3, but more importantly would provide them with education in the main areas of I&T expertise that will be outlined below (see section 3.2.2.1). Stage 4 would comprise training in the chosen specialisations in interpreting, followed by specialist accreditation examinations at stage 5. Different modules and examinations would be required for Translation and Interpreting. Similarly, different specialist training modules and examinations would be required for the specialist interpreting accreditations. As stated above, graduates of current courses that offer specialisations such as legal, medical or conference interpreting, would be exempt from undertaking the training modules and would be allowed to sit for the specialist accreditation examinations directly. We strongly recommend that, if this new model is adopted, government language policies be amended to reflect the new accreditation system.

Table 7: Proposed conceptual model for an improved accreditation system[28]

STAGE 0
Non-compulsory stage

NAATI accreditation preparedness stage
The onus of meeting this stage will be on the candidate. However, NAATI should clearly state the requirements and expectations of high bilingual proficiency in all linguistic registers. This can be done in the form of an information package posted on the NAATI website. As part of this stage, candidates will be advised to ensure that they meet the minimum language requirements before attempting stage 1, by either passing a language proficiency test administered by NAATI or meeting equivalent criteria as explained below. A clear definition of "accreditation" should also be part of this information package, particularly in establishing the difference between accreditation as a credential and formal qualifications.

***English Proficiency Test for non-native speakers of English and native speakers with limited literacy skills*[29]**	***Equivalence***
On-line test (possibly prepared by NAATI or outsourced) that can be self-administered (for a fee) and autocorrected upon completion (or equivalent as seen in parallel column). An aural/oral version for visually impaired candidates should also be available. The test can be titled 'Are your English skills good enough to attempt NAATI accreditation?' Candidates will see the result of their test upon completion. Candidates will be <u>advised</u> to attempt accreditation **only** if they achieve a minimum mark or above (to be decided). If they do not achieve the minimum mark, they can be advised to pursue TAFE or similar English courses (a list of available courses could be provided), and to re-attempt the English proficiency test in twelve months' time.	Candidates will be advised of what constitutes equivalence, for example: completion of formal studies in English, IELTS 6 or equivalent mark for other standardised English competence tests (e.g. TOEFL, Cambridge, ISLPR, etc).

---

[28] Although the language used in the model is definite (e.g. candidates will …), this is for ease of expression only, as we understand that the model is only a proposal and its implementation will depend on what NAATI decides.
[29] Non-native speaker is an inexplicit term. We suggest this applies for those who learned English after puberty. For a full discussion on the use of the term 'native speaker' see Hale & Basides (2012/13).

LOTE Proficiency for non-native speakers of LOTE

NAATI could commission tests for the required languages to be self-administered on-line for a fee and autocorrected. These tests could also be marketed nationally and internationally to be used for purposes other than accreditation (e.g. admission to a university course).

Candidates will be advised of what constitutes equivalence, for example: completion of formal studies in the LOTE, LOTE standardised language competence tests, etc. For Indigenous languages and other languages of limited diffusion, there may not be such formal competency tests and candidates may need to consult competent speakers of their LOTE to receive informal assessment of their level of competence.

STAGE 1

Advanced Diploma	Equivalence
Candidates MUST hold at least an Advanced Diploma as a pre-requisite to attempting accreditation or show equivalence (see parallel column).	In special circumstances, candidates may not possess a formal Advanced Diploma and can present a case for 'equivalence' by providing evidence of a set of shorter courses, professional experience or other types of education and training as explained above.

STAGE 2

Compulsory training on Interpreting or Translation principles and practice

Candidates will choose either interpreting or translation depending on the type of accreditation sought. There will be two compulsory modules for each discipline: Module 1: *Interpreting **or** Translation principles* and Module 2: *Interpreting **or** Translation skills.* The contents below **are only indicative**, as the final contents will be devised by the proposed Expert Panel in close consultation with the training institutions, and may differ according to language[30].

Interpreting	Translation
Module 1: Interpreting principles *Indicative content:* underlying concepts, principles and theories of Interpreting, including questions on ethics, cross-cultural pragmatics, underlying theories to the practice, applications of research results to the practice, overview of settings, etc. ***Module 2: Interpreting skills*** *Indicative content*: Different modes of interpreting (dialogue, short and long consecutive, simultaneous, sight translation), management skills, note taking skills, protocols, remote interpreting skills, etc.	***Module 1: Translation principles*** *Indicative content:* underlying concepts, principles and theories of Translation, including questions on ethics, applications of research results to the practice, etc. ***Module 2: Translation Skills*** *Indicative content*: Different types of translations, translation briefs, translator notes, translation strategies, uses of technology, revision and editing, documentation, etc.

The bulk of Module 2 may be delivered in English. However, it is recommended that bilingual components where candidates receive formal feedback from bilingual I&T experts be included at least twice during the course.

Candidates can only progress to Stage 3 after passing the hurdle assessment tasks at stage 2.

RECOGNITION

Granted to those candidates for whose language there is currently no accreditation available. In order to receive Recognition they need to have successfully completed stages 1 and 2 above

---

[30]The contents of the training modules need to be flexible enough within a general framework of essential components.

STAGE 3
Generalist Interpreting or Translation Examinations
These examinations would replace both the current Paraprofessional and Professional examinations. The content and format will be decided based on the results of the validation research study conducted in Phase 2 of the project. A Provisional Generalist accreditation with a maximum duration of 2 years can be awarded to candidates who do not reach the minimum pass mark but fall within the next band. This provisional level should only be available to new and emerging languages and Aboriginal languages, unless there are compelling grounds for allowing other languages to be included.
RE-VALIDATION
We propose that practitioners undertake professional development activities in order to maintain their accreditation, as currently being implemented by NAATI. This was supported by the results of our survey.
STAGE 4 Pre-requisites
1. *Successful completion of Stage 3*
2. *A Bachelor's degree (or equivalent – including an Advanced Diploma in Interpreting[31])*

STAGE 4

Compulsory education and training for specialist Interpreting accreditation

As per stage 2, this stage will consist of two modules covering both the theoretical underlying principles and the skills of the chosen specialisation/s (including protocols, techniques, terminology, discourse features, etc).

Practitioners intending to gain specialist accreditation in Interpreting must provide documented evidence of professional interpreting practice for a period of at least ***twelve months*** for those who achieved a Band 4 or 5, and ***18 months*** for those who achieved a band 3[32], before being allowed to sit the Specialist accreditation examination.

Legal Interpreting specialisation	Medical Interpreting specialisation	Business Interpreting specialisation	Conference Interpreting specialisation

Candidates can only progress to stage 5 after passing the hurdle assessment tasks at stage 4

STAGE 5
Specialist Accreditation Examination
The content of each test will differ according to the requirements of the specialisation and will be decided after the validation project in Phase 2.
RE-VALIDATION
We propose that re-validation continue after the attainment of specialisations.

Although we **cannot** at this stage make a firm recommendation on assessment methods and pass marks, we provide a **possible** option in Table 8 below:

---

[31] As stated before, the research team did not reach consensus on this point, with most arguing for a Bachelor's degree at this level.
[32] See explanation of pass marks below

Table 8: Standards and Pass Marks

Accreditation Level	Pass Mark		
	Rubric marking system (Bands)	Academic Grade System equivalent	Percentage equivalent[33]
Provisional Generalist (2 years duration)	3	Credit	65 - 74%
Generalist	4 and 5	Distinction and High Distinction	75% - 100%
Specialist	4 and 5	Distinction and High Distinction	75% - 100%

The system outlined above would align the results with the academic sector and would allow for a Provisional level only for Interpreting with a sunset clause of 2 years' duration. The bands and pass marks align with Angelelli's suggested rubric bands, as cited below in table 12. They also align with the results of our own survey on rubrics (see section 3.3 below).

Having provided a background and proposed a new conceptual model, the next section will discuss issues surrounding testing.

3. Testing

3.1 Language testing

This section provides a brief overview of the best-known proficiency tests for English as well as the CEFR (Common European Framework of Reference for Languages) for European languages. This overview aims to shed light on the types of language tests available to candidates at stage 0 and to also provide some guidance to NAATI in the development of its own language tests, as proposed above. Aspects of language testing can also be taken into account when designing interpreting and translation testing instruments, although we note that interpreting and translation tests have their own very specific characteristics and should not be confused for language proficiency tests.

3.1.1 IELTS

IELTS is a British-based test that is now the most popular measure for English language proficiency for academic and non-academic purposes in the world (with the exception of USA where TOEFL is more widely used). IELTS has a band-scale from 1-9 with the provision of 0.5 scores providing for 17 different gradings. IELTS owes part of its popularity to the descriptors that the testing system provides for each of the 9 bands and for each of the four macro-skills: listening, speaking, reading and writing. IELTS testers look for a mix of specific features, e.g. use of tenses in narrative speech, length and complexity of clauses, use of linking words etc., as well as 'global' features such as pragmatic appropriateness, word-attack, use of speech acts appropriate to situation. The numeric score is based on these in the first instance. Testers are supposed to refer to the descriptors only as a supplementary guide to diagnosis after a preliminary score has been reached. The descriptors provide an outline to testers, candidates and institutions of the various level ratings of the macro-skills.

There are different components used for each of the four macro-skills:

---

[33] Some institutions use slightly different percentages which would need to be taken into account. For example, RMIT uses 60-69% as Credit, 70-79% as Distinction and 80%+ as High Distinction.

Listening: understanding of main ideas and specific factual information; recognising opinions, attitudes and purpose of a speaker; and following the development of an argument.

Speaking: the ability to communicate opinions and information on everyday topics and common experiences and situations by answering a range of questions; the ability to speak at length on a given topic using appropriate language and organising ideas coherently; and the ability to express and justify opinions and to analyse, discuss and speculate about issues

Reading: reading for gist, reading for main ideas, reading for detail; understanding inferences and implied meaning; recognising a writer's opinions, attitudes and purpose; and following the development of an argument

Writing: task achievement/response, coherence and cohesion, lexical resource, grammatical range and accuracy (IELTS n.d.: 6-9).

The IELTS descriptors were first introduced as statements of level to guide candidates as well as examiners. In practice, the descriptors function not only as a guide to potential candidates but to all interested parties – the descriptors outline and disseminate to outside parties the gradings of ability in plain English. It is difficult to establish what role the 'transparency' of the descriptors has played in the overall global success of IELTS, but it can be safely assumed that the descriptors have augmented candidates' and others' notions of what the band marks signify. The descriptors are also regularly reviewed and sometimes edited and adapted. Review of IELTS test procedures and their validity is undertaken by non-interested parties, sometimes with critical conclusions (e.g. Moore & Morton, 2005). IELTS has a large research and test review infrastructure, (e.g. IELTS Research Reports, Studies in English Language Testing and Research Notes) which are all partly funded by IELTS.

3.1.2 TOEFL

The other major international English language proficiency test is TOEFL (Testing of English as a Foreign Language). TOEFL is an American-based test whose results are also accepted at most English-language tertiary institutions. Reflecting the pedagogic philosophies of late twentieth-century America, TOEFL began as a largely error-focussed tool: number and type of errors were calculated for speaking tests and marks deducted accordingly; reading comprehension tests contained multiple choice questions. This approach was normative and allowed for an efficient and speedy marking process. Today, TOEFL's iBT (internet based test) is taken online with a large part of the reading and writing sections constructed so correction can be automated. This is something that could be adopted by NAATI if it decides to deliver its own auto corrected language proficiency tests. TOEFL has adopted descriptors for speaking and writing. Each of these has components and gradings. For speaking, these are: delivery, language use, topic development. For writing, there is only one component: task development.

As explained in our recommended model above, NAATI accreditation candidates would be encouraged to obtain an English language score that is equivalent to the requirement imposed by universities for entry into their formal I&T programs, using any of the standard English language competence tests.

3.1.3 CEFR

The CEFR (Common European Framework of Reference for Languages) was developed as a EU project, funded and administered by the Council of Europe, based in Strasbourg. It is a set of rubrics that describes six levels of language proficiency, from A1 and A2 ('basic user') through to B1 and B2 ('independent user') to C1 and C2 ('proficient user'). Recent fine-tuning to the CEFR allows for a nine-level differentiation: A1, A2, A2+; B1, B1+, B2, B2+; C1, C2. The CEFR also has five rather than four macro-skills: listening, reading, spoken interaction

(pragmatic skills), spoken production and writing. The introduction of the CEFR was precipitated by the need for common terms and benchmarks to apply to speakers' language levels. The mobility of EU citizens within the EU and the harmonisation of higher education institutions led to a need for a common framework which applies to not only the languages of the EU, but to most other European languages and which can be applied to all languages worldwide. The CEFR has established itself as a measure for linguistic proficiency at European institutions of higher education. Non-native speakers are required usually to have a C1 level for admission. This level could be adopted as a guide to candidates to accreditation in the languages covered by the CEFR.

The CEFR has also provided a basis for the European Language Portfolio (ELP). The ELP is a collection of textbooks, information sheets, teaching materials and self-study and self-diagnosis resources so that individuals can ascertain their own CEFR level informally by answering questions about their abilities in their languages. Although the target for the ELP are principally primary and secondary school students, we believe this resource may be used by aspiring interpreters and translators in the relevant languages to ascertain their readiness to attempt the new accreditation process we are recommending in the LOTE. The recently developed social interpreter certification test in Flanders, Belgium has a language proficiency test in both languages, using the CEFR scale. That required level (B2) is one level lower than the C1 scale required from L2 students for entry into European universities. An example of the B2 descriptors for the macro-skills of speaking and listening is provided in Appendix 6.

3.1.4 NFAELLNC

The National Framework of Adult English Language, Literacy and Numeracy Competence (hereafter: NFAELLNC) aims to provide a guide to description, rather than diagnosis of a student's capabilities, and it has its starting point on previous education and vocational experiences, also referred to as 'recognition of prior learning'. The NFAELLNC is a good example of an instrument that offers descriptions for many features, not just language or numeracy proficiency, and could be used to ascertain equivalence to a Bachelor degree in our new proposed model.

The holistic scope of the NFAELLNC can be seen in the break-up of six areas of assessment: task, technology, identity, group, organisation, and community. The first area 'task' relates to particular elicited activities that a test candidate is assessed on. The remaining five areas describe surrounding areas that may be present in attempting an activity: technology (as an instrumental means of fulfilling a task); identity, group and community (relate to socio-psychological and socio-environmental features of performance); organisation (knowledge of administrative and legal/procedural features relevant to an assessment activity). For each of the six scales, there are three levels, from lowest to highest: 'stage 1 – assisted competence'; 'stage 2 – independent competence'; 'stage 3 – collaborative competence'.

The NFAELLNC is not tied to a test. The candidates presenting at entrance tests have very diverse backgrounds and learning experiences and a uniform test would be unworkable. The scale leaves it up to the tester to draw up their own test and to relate testees' performance to the features found in the NFAELLNC scales for diagnoses to occur. The scales are not only for entrance testing, they are also for on-going and exit assessment. A sample of the content of six areas for the highest scale ('stage 3 – collaborative competence') is contained in Appendix 7.

The NFAELNNC can also be used as an example of a scale that focuses not only on task performance but mostly on other, social-interactional and group-based capabilities for diagnosis, which can be relevant to I&T testing. This means that testing scales can seek to describe functions, which are less task-based. In T&I testing these functions can include:

- Dealing with interruptions, crowd control, un-cooperative participants, overlapping speech
- Role relationship
- Check other parties' prior experience working with interpreters
- Knowledge of cultural, rhetorical and social norms that relate to both language communities
- Stress management
- Ability to work independently – freelance
- Ability to work in a team – in-house T&I or freelancer working closely with colleagues or agencies
- Level of general knowledge and current affairs.
- Knowledge of the profiles and workings of future employers
- Research skills – able to locate sources and evaluate their relevance and usefulness to a translation/interpreting task.
- Ability and willingness to further own professional development

Many of these abilities cannot be readily ascertained from a single task and require on-going observation of testees, individually and in multi-party situations. The scale has a continuum of descriptions and once a performance descriptor is satisfied, the testee is judged to have 'achieved' it. There are two grades only: achieved (pass) and partly achieved (fail), with the scale containing only positive descriptions of ability.

3.2 Interpreting and Translation Testing

Translator and Interpreter tests are most commonly designed as performance tests. Performance assessments are a form of direct testing in which an assessment task is provided to the candidate and the candidate's performance on the task is measured according to theoretically-defined criteria. In Interpreting and Translation, it is the product of the performance i.e. the translation or the interpretation, that is measured. Messick (1994) makes the distinction between 'performance assessment' and 'performance and product assessment', a useful additional clarification in the context of translator/translation assessment, where on the one hand we have a shift to empirical methods based on psychometric principles of testing and on the other a theoretical discussion of the sub-components of quality (Angelelli & Jacobson, 2009; Slatyer, et al., 2008).

Performance tests aim to be as 'authentic' as possible, i.e. the tasks and scoring methods aim to replicate real-world assignments. The standard practice in performance test design for translators is one or more translation passages, or for interpreters dialogue and/or consecutive passages, which aim to replicate real-life translations and interpreting assignments. Inevitably, the degree of authenticity of these test items is subject to practical constraints such as the time needed to complete the task, the need to standardize the test instruments and conditions of examination. In order to achieve an acceptable degree of reliability in the test, a compromise is generally sought between the tasks' authenticity and the reliability of the test. An example of an approach to authentic test task design is the NICE test (Norwegian Interpreter Certification Exam), whereby actors (professionals in the field) role-play a semi-scripted scenario in which the candidate plays the interpreter, a practice that is also common in Australian education institutions. The performance is judged live by a jury according to theoretically-derived criteria relevant to professional practice (Mortensen, 2001). While this test has high face, construct and context validity, the potential for variability in the difficulty and delivery of the task is high, impacting on the internal reliability of the test. However, the external reliability is improved by having a panel to judge the performance (multiplying the ratings) and the use of a criterion-referenced system (a set of criteria by which performance is marked, which is less subjective

than holistic judgements), but the cost of running the test is high and therefore its practicability may be low. This example demonstrates the inevitable trade-off between the three fundamental characteristics of tests: validity, reliability and practicality (Bachman, 2000, 2002; Bachman & Palmer, 1996). Our discussion of any amendments to the current NAATI tests is framed within the compromise that needs to be found between these three aspects.

3.2.1 The current NAATI test components

NAATI conducted a Review of its accreditation tests between August 2000 and August 2001. There were meetings and workshops with representatives of the profession. Four groups of experts were formed to look at specific aspects: Translation, Interpreting, Auslan and Indigenous languages. As a result of this Review, a small number of recommendations were implemented in relation to minor changes to the Interpreter and Translator tests which led to the current formats. Below we provide a description of the current NAATI accreditation tests.

3.2.1.1 Translation Tests

NAATI translation tests are general, i.e. they are not set in specialised domains. They require a certain number of words to be translated within a given time period. Candidates must translate into English and into the LOTE at the Paraprofessional level, while tests are in one direction at the Translator level. At both the Paraprofessional Translator and Translator levels, candidates have to answer questions on ethics of the profession. Dictionaries, but not computer/internet access, are allowed.

3.2.1.2 Interpreter Tests (Spoken and Signed)

At the Paraprofessional and Interpreter levels, dialogues are set in a range of settings, many of which reflect situations common in community interpreting (e.g. hospital, welfare, legal, educational) in Australia. For spoken interpreting, candidates must accurately and idiomatically interpret two dialogues in the short consecutive mode, bi-directionally. They must avoid requests for repetitions (more than one may incur penalties). They must also answer two social and cultural awareness questions relating to each of the dialogues. Two further questions relating to the dialogues are about the AUSIT[34] code of ethics, in English and the language other than English (LOTE). For the Professional level, sight translations of two 200-word texts in each language and two long consecutive interpreting passages of up to 150 words at a time (300 words in total length) in both languages, are also part of the examination. In signed language interpreting, in addition to consecutive interpreting, two simultaneous signed interpreting exercises have been included in the test to reflect practices common in signed community interpreting. Although simultaneous interpreting is also a common mode of interpreting for spoken language interpreters, it has not been included in the Interpreter test, despite the recommendation to include it in the 2000/1 Review, mainly due to logistical obstacles.

3.2.2 Interpreter and Translator competencies, skills and related knowledge

In Interpreting and Translation, accreditation aims at ensuring a minimum acceptable standard for practitioners to perform their required tasks. Standards can be measured in many different ways, but most commonly they are judged against the competencies, skills and body of knowledge a practitioner is expected to have. As mentioned above, the NAATI tests have been criticised for failing to measure such competencies, skills and knowledge, which we will call 'skill set' from here on. Interpreting and translation scholars, many of whom have been practitioners themselves, have suggested what they have considered to be the skill set required by I&Ts based on personal experience, theory and/or on research. What appears below is a summary of

---

[34] AUSIT is the Australian Institute for Interpreters and Translators

the skill set identified in the literature on this issue, as well as through consultation with practitioners, educators, examiners and service providers.

3.2.2.1 Interpreting

Interpreting has been described in terms of a process comprising three main components: comprehension, conversion and delivery. Hale (2007b) describes the different skills, competencies and knowledge required of interpreters according to each of the three facets of the interpreting process. At the comprehension level, interpreters require a thorough knowledge of both languages at all levels (lexical, semantic and pragmatic), knowledge of the subject matter and of the particular settings and accompanying discourses. At the conversion level, interpreters require the technical skills, such as mastery of note taking skills and the different modes of interpreting (e.g. consecutive and simultaneous), as well as a thorough understanding of the underlying theories of interpreting to determine the approach to be taken according to the requirements of the setting, the interpreting specialisation, the expected role for the particular assignment and professional ethics. Finally, the delivery phase requires interpreters to be able to reproduce the message processed during the previous two stages into an appropriate form. This entails socio-pragmatic competence, mastery of public speaking skills, ability to produce different registers and to reproduce tone and suprasegmental features of language. At this stage interpreters also need to master management skills in order to coordinate bilingual situations. The current NAATI interpreting examination only assesses some of the skills at the conversion level and ignores the knowledge and skills required at the comprehension and delivery levels.

Kalina (2004) adopts a similar approach but specifically targeted to conference interpreting. She breaks down the interpreter skills into temporally-based factors: the first set of factors refers to factors defined prior to the process, the second to those immediately before and during the interpreting process, the third to those that are "…actual in-process requirements and conditions" and the fourth to post-process factors (p. 126). Table 9 below lists the conference interpreting skills under each of the four sets of factors.

Table 9: Kalina's (2004) taxonomy of Conference Interpreting skills

1. Pre-process prerequisites	3. In-process requirements
skills and competencies	knowledge and presuppositions
contract specifications	conditions of ST presentation
task definition	target language requirements
preparation	interactional competence
2. Peri-process conditions	**4. Post-process efforts**
number of participants	terminological follow-up
working languages	documentation
technical equipment	quality control
booth position	further training
team strength, composition	specialisation
working hours, event duration	adaptation to technical progress
language combinations	
relay quantity/quality	
availability of documents	
information on proceedings	

 (Kalina, 2004, p. 126)

The factors that Kalina (2004) identifies relate to conference interpreting. Court, medical, business or community interpreting (face-to-face and remote interpreting) necessitate an augmentation of these proposed factors. Kalina's list of factors is a formal attempt to comprehensively identify all personal attributes and activities that precede and succeed performance in contrast to assessment or rating scales which include only those factors which are identifiable immediately prior to and during interpreting performance. The notion of standards is more comprehensive than performance, as the establishment of standards seeks to identify *any* relevant feature or activity that condition the candidate's performance. Such standards can usually only be assessed in a course of study, rather than through a single test.

The skills and competencies incorporated into the three facets of the interpreting process by Hale and in the factors identified by Kalina, have been proposed by others also, often classified in terms of linguistic, pragmatic, socio-cultural, occupational and attitudinal attributes. These attributes can be classified according to: 1. pre-training characteristics or aptitude, such as language proficiency levels, ability to paraphrase, 'teachability', demonstrated motivation (Benmaman, 1997; Timarova & Ungoed-Thomas, 2008), and pragmatic and communicative competence (Hale, 2004; Lee, 2008); 2. those that are acquired through education and training, such as advanced listening and comprehension skills (Giovannini, 1993; Sandrelli, 2001), public speaking skills (Hertog & Reunbrouck, 1999; Pochhacker, 2001), advanced interpreting skills (Gentile, Ozolins, & Vasilakakos, 1996), management skills (Bontempo & Napier, 2009; Wadensjö, 1998), knowledge of the context and subject matter (Colin & Morris, 1996), understanding of the goals of the institution where the interpreter is working (Berk-Seligson, 1990 / 2002; Hale, 2004), understanding of the interpreter's role and professional ethics (Edwards, 1995; Mikkelson, 1996), cross-cultural awareness (Chesher, Slatyer, Doubine, Jaric, & Lazzari, 2003), theories that underpin interpreting choices (Roy, 2000; Wadensjö, 1998), knowledge of protocols (Bontempo & Napier, 2009; Jacobson, 2009); and 3. those that can be acquired through practice, such as self-confidence, stress management, ethical workplace behaviour, knowledge of OHS etc. The vast majority of these characteristics require education and training for their acquisition and development and are difficult to assess in a single examination, such as the current NAATI accreditation examination. Most I&T courses assess students in all of these areas throughout their program.

Consultation with stakeholders in Australia has also produced similar lists. In preparation for the development of national qualifications in interpreting and translation for the VET sector, Government Skills Australia (GSA), contracted by the Department of Employment, Education and Workplace Relations in 2008, conducted extensive consultation with I&T practitioners to ascertain their professional skills and competencies. Other aspects of Interpreting that were highlighted by those consulted included ability to prepare for assignments, ability to work as a team and ability to manage multi-party interactions.

Our own survey asked respondents to state the top skills they believed an accreditation examination should be testing. As can be seen in Table 10 below, the results match those that have been presented before.

Table 10: Top Interpreter skills to be tested as expressed by survey respondents

	Educators and Examiners' responses	Practitioners' responses
1	Comprehension (39)	Ethics / Professional conduct / impartiality / confidentiality / role (106)
2	Accuracy /Equivalence / Omissions / Additions / Distortions / Faithfulness / Fidelity (37)	Accuracy /Equivalence / Additions / Distortions / Faithfulness / Fidelity (75)
3	Fluency / Delivery (32)	Cultural knowledge (48)
4	Language proficiency / Competence (26)	Fluency / Delivery (45)
5	Grammar / Structure / Idiomatic (23)	Comprehension / Understanding (40)
6	Ethics / Professional conduct / Impartiality / objectivity (22)	Communication skills / Public speaking (34)
7	Cultural knowledge (21)	Terminology / Vocabulary / Jargon / Lexicon (31)
8	Terminology / vocabulary / jargon (18)	Setting / Subject matter / Context / Field (22)
9	Setting / Subject matter / context / field (18)	Register / Style / Expression (21)
10	Communication skills / Public speaking (17)	Listening / Concentration (19)
11	Register / style /expression (15)	Memory (17)
12	Note taking (13)	Grammar / Idiomaticity (16)
13	Modes / Sight translation / Dialogue / consecutive / simultaneous (13)	Confidence / maturity / assertiveness (14)
14	Memory (12)	Note taking (14)
15	Pronunciation / clear accent (12)	Pronunciation / clear accent (14)

A number of the characteristics listed above reflect those that are currently tested in the accreditation examination, such as note taking skills, ethics, language competence, consecutive interpreting and sight translation. Others, however, are noticeably absent, such as the interpreter's ability to manage the situation when the speakers do not adhere to the expected norms, the interpreter's ability to coordinate turns between speakers, or the interpreter's understanding of his/her role according to the goals of the institution for which they are working. Another important aspect of interpreting that is currently not tested and that is impossible to test in a single generalist examination, is the interpreter's understanding of the theory to inform his/her choices with regards to the approach taken according to the setting and the participants involved. Such theoretical knowledge is also necessary for interpreters to justify their performance when challenged (Baker, 1992; Calzada Perez, 2005; Hale, 2007a), something that is becoming increasingly common, especially in court interpreting. Not surprisingly, few respondents in our survey indicated the need for theory in interpreting and translation tests, although more seemed to indicate this was necessary for translation (see Table 11 below). This lack of appreciation for the theory is likely to be symptomatic of a profession where the overwhelming number of practitioners (and NAATI examiners) lack any formal training in theoretical aspects. In mature professions whose members make decisions that impact on the public (e.g. medicine, law, engineering), skills are developed on the basis of theory. Members of these mature professions are expected to be able to make expert autonomous decisions, to be able to analyse, describe and report upon their professional decision-making using particular terminology that is drawn from theoretical training. We will discuss these issues further when we deal with test design, marking criteria and examiners' competence.

3.2.2.2 Translation

The majority of literature on translation competence to date is theoretical and reflects the disciplinary perspective of the theorist. Text-based linguistic models include Baker's (1992) theory of equivalence, which draws on systemic functional linguistics, pragmatics and cultural studies to propose a bottom-up view of the relationship between source and target text characteristics. Hatim and Mason's (1997) work on communicative models of translation expand these notions of equivalence to include the interactional aspects of the translation act. In contrast, functional models situate the translation act in relation to the purpose (or skopos) of the translation. Nord (1991), expanding on Vermeer's Skopos Theory, proposes a process for analysing the source text as the starting point for translators.

Early models of translator competence (i.e. models of process as opposed to product) include Wilss (1982) who described the act of translating (rather than the characteristics of the translation). Wilss's model comprised three components: source language receptive competence (or the ability to understand the source text), target language reproductive competence (or the ability to express concepts in the target language), and a super-competence, which describes strategic translation competence (or the ability to translate).

While these theoretical models have provided the basis for our understanding of the relationship between source and target text and the role of the translator, they lack empirical evidence. The only comprehensive, empirically supported model of translator competence is the PACTE model (PACTE Group, 2009). The model is the result of over ten years of research which has investigated the construct through a robust, triangulated research design incorporating both product and process components of translator performance. The model describes translator competence according to five interconnected sub-competencies and a psycho-physiological component. The Psycho-physiological components, which are common to professional practice in many fields, include attitudinal (e.g. critical thinking, creativity, etc.), cognitive (e.g. memory, attention, emotion) and psycho-motor skills components. The other five sub-competences are:

- *Bilingual sub-competence*: procedural knowledge and skills related to communication in two languages and comprises pragmatic, sociolinguistic, textual, grammatical and lexical knowledge
- *Extra-linguistic sub-competence*: predominantly declarative knowledge comprising general world knowledge, domain-specific knowledge, bicultural and encyclopaedic knowledge.
- *Knowledge about translation*: predominantly procedural knowledge about translation and aspects of the profession, i.e. knowledge about how translation functions and about professional practice
- *Instrumental sub-competence*: procedural knowledge relating to the documentation resources and information and communication technologies as applied to translation
- *Strategic sub-competence*: procedural knowledge about how to ensure an efficient translation process and to solve problems arising in the translation. Involves, planning, carrying out the translation, evaluating the translation (PACTE Group, 2009, pp. 318-320).

Of these five sub-competences, the first two (Bilingual sub-competence and Extra-linguistic sub-competence) and the psycho-physiological component may reside independently of the translation context. They could be considered to be prerequisites to professional translator education and are often included in screening tests for entry to educational programs to assess the suitability of candidates for the profession. This would be addressed in our proposed stage 0. The last three (Knowledge about translation, Instrumental sub-competence and Strategic sub-competence) comprise the knowledge and skills that are essential for a professional translator. These three sub-competences constitute the professional knowledge and skills that are acquired during the process of education. This would be addressed in our proposed stages

1 and 2, and should also be included as key components of the test construct in translator testing programs whether this be the summative assessment at the end of a course of instruction or a gate-keeping test such as the NAATI test evaluating professional readiness.

The respondents of our survey provided a list of skills to be tested by accreditation examinations as we can see in Table 11 below.

Table 11: Top Translator skills to be tested as expressed by survey respondents

	Examiners and educators' responses	Practitioners responses
1	Language proficiency (81)	Accuracy /Equivalence / Omissions & additions / Faithfulness (84)
2	Register / Style / Expression (53)	Register / Style / Expression / Tone / Naturalness / Idiomaticity / Coherence (71)
3	Translation / Transfer skills (51)	Comprehension / Understanding / Analytical skills (62)
4	Grammar / Structure / Syntax (48)	Grammar / Structure / Syntax / Punctuation (57)
5	Comprehension (23)	Language proficiency/ competence / literacy / Fluency / clarity (44)
6	Accuracy /Equivalence / Omissions / Faithfulness (23)	Terminology / vocabulary / use of dictionaries (38)
7	Cultural knowledge (22)	Cultural knowledge (37)
8	Competence (21)	Writing skills (33)
9	Ethics / Professional conduct (21)	Ethics / Professional conduct / confidentiality / impartiality (33)
10	Terminology / vocabulary (12)	Spelling (19)
11	Research / Theory / Ability to translate according to Skopos/Purpose (12)	Theory / understanding of Purpose / intent / Pragmatics (17)
12	Spelling (9)	Research skills (Preparation) (16)
13	General knowledge (9)	Computers / Technology / localisation (14)
14	Editing / Checking (5)	Translation / Transfer skills (12)
15	Computers / Technology (4)	Editing / Checking / Format (12)

As with the interpreting test, a number of the skills proposed by the survey respondents are currently tested, such as comprehension, accuracy, terminology and writing skills, however, others are absent from the current NAATI examinations (e.g. the use of technology, checking, editing and formatting skills and an understanding of the underlying theories of translation). This report will provide a detailed section on the use of technology for testing later. However, the other issues relating to theory and checking and editing are competencies that are more adequately taught and assessed through education, rather than through an accreditation test. The section below will present an overview of the I&T assessment research and practice around the world.

3.3 Marking systems

3.3.1 Overview of marking systems

Historically, marking systems that are based on pre-conceived criteria and that rely on the deduction or awarding of marks based on these criteria have been termed 'analytic' marking

methods (Eyckmans, Anckaert, & Segers, 2009; Lee, 2008; Turner, Lai, & Huang, 2010). Marking systems that are more 'impressionistic' and which evaluate a candidate's performance in a more 'global', 'intuitive' way, whether examining a test overall or breaking down a test into particular areas, has been termed a 'holistic' marking method (Bontempo & Hutchinson, 2011; Lee, 2009). Descriptors are the usual means for holistic test evaluation. Trial and adoption of descriptors has, in the evaluation of interpreting and translation testing, been advocated by some as a means of providing alternative or supplementary feedback (e.g. Turner, et al., 2010), as a means of verifying and testing the validity of analytic testing (e.g. Turner, et al., 2010; Waddington, 2004) and as a method preferable to analytic testing (e.g. Lee, 2009).

Analytic and holistic testing systems can conform to psychometric testing requirements of validity so they can readily test for activities that a test candidate would undertake in the T&I profession, i.e. the test contains what is required in everyday T&I professional life. However, analytic and holistic testing systems are vulnerable to problems associated with inter-rater reliability, i.e. different testers using the same method and awarding very different marks due to different 'subjective' applications of the marking system. They can also be vulnerable to the problem of intra-rater variation, i.e. the same tester applying the same marking method to the same test but arriving at different scores due to previous tests marked, time of day, level of fatigue etc.

Traditionally, translator and interpreter performance tests have been scored using an error deduction ('points-off' or 'penalty') system, which is the system currently adopted by NAATI for translation tests. The 'points-off' system assesses the product of translation rather than the process or the ability of the translator and takes points off for identified error types. Other marking systems are also being used by different accreditation/ certification bodies. Turner et al. (2010) surveyed 24 different accreditation/certification systems for translators and interpreters currently employed around the world and found that scoring systems are based on one of three designs:

- error analysis/deduction systems (like NAATI's system for translation tests);
- a combination of criterion-referencing (the use of scales of descriptors to describe performance in tests) and error analysis/deduction; and
- criterion referencing (descriptors) with no system of error analysis/deduction (Turner, et al., 2010, p. 12)

A promising recent shift in the field is the development of theoretically derived rubrics, such as those developed by Angelelli (2007, 2009), Lee (2005) and Jacobson (2009). Rubrics-based systems use sets of 'descriptors' (word-pictures) of performance at various levels (typically identified by numbers) to help markers determine the result. Usually, there will not be only one set of such descriptors, but several in order to reflect the various sub-components of the skill being assessed. A marker using a rubrics-based system will still need to identify, at an early stage in the marking process, the various errors that have been made. However, once they have been identified and noted, awarding a level is done on the basis of comparing the observed performance with the various descriptors. The level awarded in each area is determined by selecting the descriptor that most closely matches the observed performance. These sub-components (or dimensions/assessment areas) will usually be determined by carefully analysing the test construct (Angelelli, 2009, p. 38). The descriptors at each level can be determined by looking at various samples of candidate performance that are agreed to be of a given standard, and identifying the observed characteristics. The number of levels available needs to be considered carefully, as too few levels will give insufficient discrimination between very good and very poor performances, while too many levels may simply confuse markers (Angelelli, 2009, p. 44). Pass/fail is usually determined in terms of achieving a specified level in each assessment area (although it is not unusual for a candidate's performance to be uneven across the assessment areas). Pass/fail could be determined at the same level for all

assessment areas, but another possibility is to allow some flexibility, so that being below the specified level in (for instance) one area can still allow a candidate to pass. This is what we are suggesting for the Provisional level.

Another difference between the two major systems presented above (point deduction vs criterion referenced), is the level of directness of the assessment. Performance tests, such as the NAATI test, use the direct assessment method, where the candidate is assessed on the interpreting or translation task directly. Indirect testing, on the other hand, targets the sub-components of the target construct or traits. Examples of this are a vocabulary test for interpreters (Skaaden, 1999) which has good reliability (objective scoring of correct responses) and predictive validity for outcomes of an educational program, but low face validity; or the test designed by Stansfield et al. (1992) for translators, which combines direct and indirect components, thereby balancing validity and reliability; the indirect components (such as multiple choice language items) have higher reliability and the performance items have higher validity.

Interpreting assessment has been the subject of a number of studies. In Australia, Lee (2009) compares analytic scales and holistic scales (rubrics) based on the CISOC (Community interpreting services of Ottawa-Carleton) test. The three bands: accuracy (40%), target language quality (40%); and delivery (20%) were given to nine experienced interpreting examiners to trial. Each of the holistic bands had six scales along it by which examiners could show their assessment. Examiners were also given conventional analytic scales, i.e. scales that examine different components of performance separately with a punitive, point-deduction system for each component. Examiners were asked to rate the same interpreters' performance, firstly using a holistic scale and secondly using an analytic scale. Lee (2009) reports that the examiners initially assumed that their analytic ratings would be more 'accurate' (i.e. more able to closely describe and quantify performance) than the holistic ratings, as this was the convention that they were most used to. However, there was no general dissatisfaction with the holistic scales: "the majority of the raters approved of the [holistic] rating scales proposed by the researcher, and the rating results also pointed to high inter-rater reliability" (Lee, 2009, p. 183). In the end, Lee (2009, p. 193) still cautions that the results are mixed and perhaps inconclusive and recommends that further research be undertaken to further test the validity and reliability of holistic scales. Jacobson (2009, pp. 61-65) argues that rubrics or descriptors constitute a better way to quantify performance for non-linguistic features such as contextualisation cues (e.g. paralinguistic features that signal meaning such as intonation contour, eye gaze, body position) and the professional establishment of the interpreter's role relationship to others (e.g. pre-interaction establishment of the interpreter's role, management of other interlocutors' turn-taking opportunities, management of over-lapping speech and interruptions etc). A sample of Jacobson's rubrics is contained in Appendix 8.

Other practitioner-researchers in Australia have also devised evaluation rubrics for interpreter performance. Bontempo has devised and uses rubrics for the evaluation of (Auslan/English) interpreter performance in general (2009b) and also for specialised situations such as conference interpreting (Bontempo, 2009a) and educational interpreting (Bontempo & Hutchinson, 2011). Bontempo's rubrics are intended as guides not only for testers to employ for on-going or exit testing, but also as a professional development tool for practitioners to evaluate others' performance and for them to reflect on their own. The rubrics contain value-neutral statements or interrogatives and testers are left to award marks as they see the relevant features to be present or absent, e.g. 'Equivalence of message (appropriate for context? Contains textual integrity and fidelity? Is information exchange successful overall?). The rubrics contain four key elements – interpreting aspect, language aspect, interaction/role aspects, professional conduct – and each element provides a mark of up to 5. This breakdown of marks limits cross-linguistic transfer and target language performance to 25% each and awards the remaining 50% of the marks to pragmatic and professional aspects of performance, which differs from more traditional breakdowns of marks that award the majority of marks to the first

two-mentioned elements. A full description of Bontempo's interpreter performance evaluation rubric is contained in Appendix 9. Bontempo's rubrics reflect a school of thought within Interpreting Studies that advocates that a qualitative assessment of performance should equally consider interactional, role-management, pragmatic and professional features. The setting in which Auslan interpreting is often required, namely the educational setting (Ozolins & Bridge 1999, as cited in Napier, 2004, p. 352), has facilitated a strong tradition of pedagogic and practical interchange in the teaching and assessment of sign language interpreting.

Another approach to interpreter assessment is the listener-focussed approach that is less concerned with accuracy of transfer but more concerned with the interpreting quality assessment. Such an approach is proposed by Schjoldager (1995) for conference interpreting assessment, with four progressive scales of performance, from lowest to highest: 1) the listener can understand and bear to listen to the interpreter; 2) the interpreter's use of language is appropriate; 3) the interpreter's rendition is coherent and plausible; 4) the interpreter is a loyal communicator. This scale is augmented by individual comments made by the tester. Another measure of quality is proposed by Pöchhacker's (2001) as containing four features, two of which (1. and 3.) focus on language transfer: 1) accuracy – faithful representation of discourse; 2) adequacy – respecting the formal conventions of the target language; 3) equivalency – the degree the interpretation represents the source-language speaker's interests and intentions; 4) success – communicative interaction for all parties facilitated.

Tiselius (2010) adapts scales from Carroll (1966) that had originally been developed for assessments of the quality of machine translation and applies them to the following criteria in the assessment of conference interpreting: intelligibility and informativeness. The main point of Tiselius's (2010) research is that specialist interpreter examiners and lay people award very similar ratings. Both groups of examiners received transcribed renditions of conference interpreters' interpretations in their first language and, following Tiselius's adapted 6-scale metric for both intelligibility and informativeness, were able to match these in similar ways when reading through the ad verbatim transcriptions of conference interpreters.

For translation assessment, Angelelli (2009) surveyed the recent literature in sociolinguistics, discourse analysis and second language acquisition in order to apply the directions that these disciplines have taken in recent years to the performance of translation testees. In particular, teaching and testing methodologies in the field of second language acquisition have, over the last 25 years in most Western countries, advocated a communicative approach of language use in which a speaker's abilities to functionally communicate with others overrides the importance of grammatical or lexical accuracy. Angelelli (2009, p. 29), describing the current American ATA translation exam, states that "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.)". Without wanting to disregard the importance of grammatical and lexical accuracy, Angelelli seeks to systematically list non-linguistic criteria in translation performance as important and worthy of consideration in assessment.

Angelelli (2009) seeks to measure and to quantify performance through the following descriptive rubrics: source text meaning; style and cohesion (addressing textual sub-component); situational appropriateness (addressing pragmatic sub-component); grammar and mechanics (addressing micro-linguistic sub-component); and translation skill (strategic sub-component). Angelelli (2009, pp. 40-41) proposes descriptive, 5-point rubrics, which are being considered by the ATA for adoption in its marking system. An example of the rubric for 'source text meaning' is given in Table 12 below.

Table 12: Angelelli's proposed 5 point rubrics

	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.

Legend: T = translation; TL = target language; ST = source text.

Angelelli (2009, p. 43) proposes that a mark of 4 or above satisfies a typically required standard: "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 full description of Angelelli's remaining four rubrics is given in Appendix 10.

Turner, Lai, and Huang (2010) conducted a study which compared marking outcomes using the current NAATI system and a rubrics-based system (the DPSI from the UK). In this study, a number of translating test papers from accreditation students at RMIT University were marked by experienced NAATI examiners using both systems (in the case of DPSI, using both blind and non-blind marking), and the results compared. The findings were interesting. There was a strong correlation between 'NAATI' and 'non-blind DPSI' marking across all language groups, and a weaker but still significant correlation between 'NAATI' and 'blind DPSI' marking. In a focus group discussion held with these markers, they stated that they felt they could have benefited from more extensive training in the use of a rubrics-based system before taking part in the study, and supposedly due to their inexperience in the use of rubrics, they tended to prefer the current NAATI system. The results of the 2010 study, however, contrast with the results of a small study on the use of rubrics conducted by our research team for the purpose of this current research project. Two groups were recruited as participants: a group of 'practitioners' or representatives of T&I 'agencies' (N=7); and a group of NAATI 'examiners' and/or 'educators' (N=11). The participants were invited to a short presentation on the background to rubrics-based marking systems and provided with a copy of the set of rubrics proposed by Angelelli (Angelelli, 2009, pp. 40-41), a closely-translated English 'sample translation' of a LOTE text from a previous NAATI Translator test, responses from three translation candidates, with various types of deficiencies, and a rubrics-based grid to score the translations. The participants were then asked to score one of the candidates using the set of rubrics provided, after which they were asked to record their impressions and comments of the process on a questionnaire. The first three questions of the questionnaire were different for the two groups: practitioners & agencies were asked about the potential usefulness of a rubrics-based marking system for those employing or working with T&Is, and examiners & educators were asked about how a rubrics-based system might result in better marking. The remaining questions were the same for both groups, and sought feedback on issues such as preferred 'pass/fail' and 'hurdle' levels, and suggestions for adding or deleting assessment areas. A copy of the questionnaires is found in Appendix 11. In general (and within the limits of the small sample size), there was definite –

but not unanimous – agreement among both groups of participants in favour of using rubrics in marking. Among examiners/educators, there was good agreement that rubrics can provide clearer guidance and be easier to use, and strong agreement that rubrics can encourage markers to take a wider range of factors into account. Among practitioners/agencies, there was strong agreement that a rubrics-based system could be a good basis for determining accreditation, but ambivalence about the usefulness of rubrics-based levels for employment decisions or for reporting results. Both groups expressed a preference for level 4 in a 5-level system as a 'passing' level, and level 3 or lower as being a 'hurdle' level that would preclude a pass.

Kim (2009) advocates a different framework of assessment criteria from the current NAATI criteria of translation examination which include the following: too free a translation, too literal a translation, spelling, grammar, syntax, punctuation, failure to finish a passage, unjustifiable omissions, mistranslations, non-idiomatic usage and insufficient understanding of the ethics of the profession. Based on a systemic functional linguistics (SFL) framework, Kim proposes assessment criteria where points are deducted for different features:

Table 13: Kim's alternative marking criteria

			Lexis	Clause	Text
Major	Experiential	Accuracy	1-2 points	2-3 points	
		Naturalness	1-2 points	2-3 points	
	Logical	Accuracy		1-3 points	
		Naturalness		1-3 points	
	Interpersonal	Accuracy	1-2 points		3-5 points
		Naturalness	1-2 points		3-5 points
	Textual	Accuracy		1-2 points	3-5 points
		Naturalness		1-2 points	3-5 points
Minor	Graphological mistakes such as spelling 0.5 points				
	Minor grammar mistakes that do not impact on meaning 0.5 points				

Kim (2009, p. 135) argues that such an approach which looks at the meaning of the ST and TT is preferable to one that focuses on errors per se: "…the present criteria do not specify possible forms of errors, such as additions, omissions, and inadequate equivalence, because what is important is to judge whether a mistake has something to do with accurate and natural delivery of different aspects of meaning". Kim (2009, p. 150) points to increased student satisfaction in class and also to increased pass rates (from 10% in 2004 to over 60% in 2007) for the NAATI accreditation test (English into Korean) amongst students whose work in class had been assessed to the above grid rather than ones closer to the existing NAATI one.

Another example is Gile's (2004) notion of feedback from trainees or testees about how they approach and perform translation tasks, i.e. an explicit account (without following any required content format) of why and how they translated texts in the way that they did. This kind of feedback, which Gile requires of all his translation students, is termed *integrated problem and decision reporting*. Gile (2004, p. 34) requires of the trainees that they "include full references of sources consulted, and preferably the context in which target-language terms or expressions which they chose were found (generally a sentence, sometimes a whole paragraph)." This focus on students' reflection on the process of translation is common in higher education I&T courses, but absent in performance tests such as NAATI, despite the recommendation for a reflexive component in the accreditation tests from the 2000/1 Review.

3.3.2 International comparisons of current assessment practices

A number of studies have collated data from training institutions, professional associations and/or certification/licensing authorities to provide a wide view of the approaches taken worldwide. Waddington (2001, p. 316) surveyed 52 trainers from 20 universities in Europe and Canada and reports that roughly equal numbers employ a non-punitive holistic method (38.5%) compared to an error analysis approach (36.5%), while the remainder (23%) employ a combination of both approaches. Turner & Ozolins (2007) also collated data from 10 international institutions that conduct testing (a mix of universities and professional associations) and found that the testing authorities generally adopt a punitive approach to marking. Most marking systems were based on point deduction based on the severity and frequency of errors. The American Translators' Association (ATA) test appears more punitive than the current NAATI system with a pass mark of 80% and a rigid method of error allocation. Similar error-focussed approaches are used in the marking of tests conducted by the Brazilian Translators Association, the South African Translators' Institute and the Ukrainian Translators Association (Turner & Ozolins, 2007, p. 44). At the time of collation, Turner & Ozolins (2007, p. 38) report that the ATA was about to adopt a less normative and more holistic approach to marking, as already stated above. Institutions in Norway and the UK Institute of Linguists also report the use of holistically-based rubrics for testing.

Below we provide more detailed descriptions of the I&T tests and marking systems used by a selection of countries with publicly available information.

3.3.2.1 United Kingdom

Examiners for the UK Institute of Linguistics Diploma in Public Service Interpreting (hereafter: DiPSI) award the following grades for particular tasks in the DiPSI: distinction, merit, pass, fail. An averaged mark is arrived at on the basis of a candidate's overall performance in the areas of law, health or local government. Candidates with a 'successful' (i.e. ≥ 50%) average are awarded a 'pass' grade and certification. The IoL DPSI test has three components similar to the components of the NAATI test, e.g. accuracy, delivery and language use for the interpreting test; accuracy/appropriateness of translated text, cohesion and genre conventions, effectiveness of communication for the translating test. The three components are weighted equally and general features recorded on a rubric determine the score that a candidate is awarded (IoLET, 2004, pp. 9-17).

Similarly, the IoL Diploma in Translation consists of tests assessed according to criteria very similar to those of the NAATI tests: candidates receive a percentage score with a 'pass' mark of 60%. A points breakdown is made according to the following three components: comprehension, accuracy and register (50%); grammar (morphology, syntax, etc), cohesion, coherence and organisation of work (35%); technical points relating to spelling, accentuation, punctuation and the transfer of dates, names, figures, etc. (15%) (IoL, 2011). Space is provided on each candidate's mark sheet for examiners to provide comments for each of the three components. The marking systems for the tests in the IoL diplomas are comparable to those used for the NAATI tests (see Appendix 12).

3.3.2.2 United States of America

The American Translators Association (ATA) conducts its own examinations and has a Certification Exam. Passing the exam (minimum pass mark not specified) entitles a candidate to receive certification (ATA, 2011). When marking the exam, examiners provide feedback to candidates on the type and frequency of errors (23 error types are identified with descriptions of each – see Appendix 13). Together with this, candidates receive a rubric or list of descriptors referring to four areas of performance: usefulness/transfer, terminology/style, idiomatic writing and target mechanics. Each rubric contains a description of performance with four benchmarks:

strong, acceptable, deficient, minimal. The rubric's descriptors are short and generalised and list the absence or presence of inadequate features of translation.

Thus, the ATA examination adopts an approach in which performance is measured by both descriptive rubrics and identification of errors made with a refined but complicated system of quantification of error type (ATA, n.d.). At the same time there are reminders to the examiner of 'global' features: e.g. 'Is it intelligible to the target reader?'

As explained above, interpreter certification is conducted according to area of specialisation. For certification with the National Board for Certification of Medical Interpreters, two tests must be passed. The first test is a written exam (in English only) which covers the following: Medical Knowledge – not translation of terms, but understanding of terminology (75% of exam is medical knowledge/background); Roles of the Medical Interpreter; Medical Interpreter Ethics; Cultural Awareness; Medical Specialties; Interpreter Standards of Practice (IMIA, CHIA, NCIHC); Legislation and Regulations (HIPAA, CLAS). The second test is an oral test that is administered at sites across the country using web technology. The test is 40 minutes long and contains 12 mini-scenarios lasting 30 minutes and 2 sight translation passages lasting 10 minutes. The oral test covers the same subject matter as described previously, with all topics based on the job analysis. The pass mark for both tests is 70% and candidates receive "summary results" compiled by a "rater" 2-4 weeks after the test. There are no available details on the marking system employed.

For the certification of court interpreters, pre-test training is typically offered by private providers and the certification exam is offered individually by a nominated authority in each state – sometimes a university institution, sometimes a court authority, sometimes a private enterprise or agency. There is great variation in the training, security screening and formal testing of candidates. For example, for the year 2009 the National Center for State Courts (2009) lists four steps to certification: orientation workshop, security record check, written test and oral test. In some states, all of these four steps are required, in others only some of them (typically the written and oral tests), while in others no testing is planned or available.

In order to gain an insight into a standardised and formally administered test used in the USA, this report takes as an example The Federal Court Interpreter Certification Examination (FCICE) for Spanish/English. The FCICE is a test offered in three languages only: Spanish, Navajo and Haitian Creole. The FCICI consists of a two-phase examination of language proficiency and interpretation performance, consisting of a written examination with multiple-choice answers to test proficiency in grammar and expression in both languages, with a pass mark of 75%. The Oral Examination has a pass mark of 80%. A description of the FCICI test is provided in Appendix 14. The two examinations are administered in alternate years. The first phase of the examination, referred to as the Written Examination, is a multiple-choice test of language proficiency in English and Spanish, and is offered in even-numbered years. The duration of the Written Examination is three hour and fifteen minutes. The second phase is a 45-minute Oral Examination that simulates the work that interpreters do in court, and is offered in odd-numbered years. Candidates must pass the Phase One Written Examination in order to qualify to take the Phase Two Oral Examination (National Center for State Courts, 2011, p. 7). The offering of examinations in alternate years is something NAATI could adopt for languages that have high supplies of practitioners, such as Spanish or Chinese. Candidates in those languages should be encouraged to enrol in formal tertiary courses.

3.3.2.3 Canada

The national organisation, the Canadian Translators, Terminologists and Interpreters Council (hereafter CTTIC) conducts a translation test for certification which contains two texts, each of 175-185 words in length. One text is general in nature and compulsory, the other is a choice of

either a technical/scientific/medical text or an administrative/economic text. No text requires specialised terminological knowledge. Marking proceeds according to the following criteria:

Errors – translation (comprehension, i.e. failure to render the meaning of the original text)
– language (expression, i.e. violation of grammatical and other rules of usage in the target language).

For major mistakes (e.g. serious misinterpretation denoting a definite lack of comprehension of the source language) 10 marks are deducted. For minor mistakes (e.g. unacceptable loan translation) 5 marks are deducted. The pass mark is 70% for each exam. (Further details on the CTTIC guide are in Appendix 15).

No information is available on the marking of interpreting examinations conducted by CTTIC. However, the following details are available about the composition and marking of community interpreter tests at the local level, offered by the city of Ottawa. The CISOC (Community interpreting services of Ottawa-Carleton) has a community interpreting test that contains dialogue interpreting, sight translation and consecutive interpreting with a total test length of 45 minutes (Roberts, 2000). The marking system requires two markers, each one having as their L1 one of the languages of the pair. The dialogue interpreting component of the test has five areas of assessment which are awarded up to 2 marks each. Further details of the marking grid for the CISOC test are contained in Appendix 16. These two tests are similar to our current NAATI translator and interpreter examinations.

Elsewhere in Ontario, the ATIO (Association of Translators and Interpreters of Ontario) has a pathway of certification to becoming a translator, conference interpreter, court interpreter or terminologist through compilation of a dossier which can provide evidence of five years' full-time experience (or two years if applicants hold an bachelor honour's degree, or equivalent, in their occupational category) (ATIO, 2011). In British Columbia, a court interpreting test exists which requires examination of legal knowledge (elicited in written form) and an oral examination. A description of the format and the marking guide appears as Appendix 17.

The Association of Visual Language Interpreters of Canada (AVLIC) revised its certification processes in 2002 and 2004 and (Russell & Malcolm, 2009) are contemplating a new testing procedure, which includes a prerequisite of training with a minimum of two years full-time study and detailed feedback in simulated performance before the certification test. The redesign of the testing procedure included preliminary language testing, firstly in English and secondly in ASL (American Sign Language) before message equivalency was tested. This is similar to what we are proposing for our new model. For the revised testing procedures, the professional association considered on-going, cumulative assessment through portfolio development rather than a test as a means of ascertaining trainees' levels but decided against this due to workload demands (for trainees and testers) and to concerns about the validity and reliability of such a system. Psychometric analysis preceded development of the revised testing procedure to ensure that these features – content and construct validity and inter-rater reliability could be addressed. This is what we propose must happen in Australia before a final decision is made on the new test design and content.

The marking system used in the AVLIC test is not based on error calculation or on a descriptive checklist. Instead, the AVLIC test is based on examiners identifying criteria and "making evidence-based decisions about the consistent representation of those features across all […] test segments" (Russell & Malcolm, 2009). Russell and Malcolm (2009) describe this testing procedure as a qualitative one. The marking system for the test has two sections: sign language criteria (linguistic criteria) and message equivalence criteria (cross-linguistic transfer). The two sections are presented in Appendix 18.

3.3.2.4 Europe

In some European countries with established T&I training centres, usually at university-level, marking practices reflect those of the academic or vocational institution. As explained above, the notion of a 'single testing procedure' for accreditation, recognition or certification generally does not apply in most European countries. Instead, accreditation, recognition or certification, whether tacit (i.e. the permission to list a T&I qualification after one's name) or formal (i.e. being admitted to or seeking permission to practice professionally through formal means other than testing) is gained through course completion and attainment of a specialist T&I academic degree. Further to formal training and qualifications, many European countries have a formal process of 'registration' of practitioners, very often court interpreters and translators.

The European Union has one of the largest and most extensive T&I infrastructures in the world. The focus of T&I performance within the EU is conference and speech interpreting, document, legal and speech translation but candidates for employment at EU institutions must also, in addition to a post-graduate qualification in T&I, pass a test which includes assessed performance of translation or interpreting skills, but also knowledge of EU institutions and areas of responsibility. A description of the test for employment to work as an employed interpreter is provided in Appendix 19. A pass mark of 50% is required, but no details are provided of the marking criteria.

3.3.2.5 Flanders, Belgium

The test and training developed in Flanders is of interest to us because it also prescribes pre-testing training and language screening as we propose for the new model. It is also relevant to Australia because the target group of testees and trainees is community interpreters. The design and content of the test and accompanying training was based on close observation of the practices of the Institute of Linguists (UK) and training institutions in The Netherlands. The language proficiency tests are undertaken before any training commences. Upon successful completion of the proficiency test, a candidate progresses to training, which is developed jointly with testing. The training's curriculum

> "…focuses on training interpreting skills and providing social interpreters with information, practice and discussions on the code of ethics and the contexts in which social interpreters usually work…. The curriculum was reviewed and currently consists of the following elements: a Dutch proficiency admission test, an 18-hour introductory course, an interpreting aptitude test, an 84-hour basic interpreting training module, a certification exam, and a 21-hour remedial training module for candidates who have failed the test" (Vermeiren, Van Gucht, & De Bontridder, 2009, p. 307).

The certification exam consists of consecutive speech interpreting (reproduction), sight translation and dialogue interpreting. The marking system is complicated and prescribes the awarding of up to five marks for fulfilment of criteria such as cohesion, coherence, completeness and presentation, but also the deduction of marks for poor performance. The marking of the exams contains elements of three different marking conventions: analytical assessment of performance through location of errors and mark deduction; descriptors against which performance is measured; and norm-based features such that each speech interpreting task contains numerical references (date and telephone number), names (person organization) and an enumeration of 5 units (e.g. five symptoms of a medical condition). These norm-based features are checked in target speech interpretations. Overall, the marking system is a combination of holistic (descriptor-based) features and analytic and norm-based assessment which are triple-marked by a chairperson, a Dutch language specialist and another-language specialist, all with T&I expertise. Although the exam is preceded by a language proficiency test, accuracy of language (in both languages) is still measured and marked concurrently with assessment of message transfer. The training of graders includes provision for analysis of

spoken performance for these features at the same time, both during the test and when viewing the video recording of it afterwards.

3.3.2.6 South Africa

Testing and 'accreditation' (the term used in South Africa to refer to credentialing candidates' performance standards) is conducted by the South African Translators' Institute (SATI). Tests are available for translation, sworn translation (i.e. court or legal translation with specialised source texts and a higher general threshold for marking than that for the general translation test), interpreting, terminology and language editing. Not only individuals but agencies and businesses may apply for 'accreditation' which is assessed on a case-by-case basis by SATI. An interesting aspect of this test is that it is given to candidates to take home and return in 24 hours, without invigilation. The marking systems employed for the translation, interpreting and South African sign language (SASL) tests are analytic: criteria are provided according to which marks are awarded for performance, without fixed norms.

In the marking of the translation test, it is clear that examiners are instructed to identify errors and assess their severity. The following information is provided on the SASL website in regard to the examination of translation tests:

> Examinations are assessed on the basis of a system of major and minor errors originally drawn up by the American Translators Association. Major and minor errors are defined as follows:
>
> Major errors: Gross mistranslation, in which the meaning of the original word or phrase is lost altogether; omission of vital words or other information; insertion of information not contained in the original; inclusion of alternate translations, where the translator should have made a choice; and any important failure in target-language grammar.
>
> Minor errors: Mistranslation that distorts somewhat, but does not wholly falsify, the intent of the original; omission of words that contribute only slightly to meaning; presentation of alternate translations where the terms offered are synonymous or nearly so; and 'inelegance' in target-language grammar.
>
> To pass the exam a candidate must have:
>
> No more than 20 minor errors in the exam as a whole
> No more than 10 minor errors in any one text
> No more than one major error in the exam as a whole
> No more than six minor errors in a text that contains a major error
> (South African Translators' Institute, 2007)

For spoken interpreting, the assessment criteria emphasise target language performance (i.e. TL vocabulary and register, grammar, idiom and purity are two of the four criteria groups) and the pass mark individually for each criteria group as well as collectively is 80%. For signed language interpreting, pragmatic and professional attributes are also marked in assessment. The marking criteria for South African sign language (SASL) are language skills (vocabulary grammar, idiom, purity), content/message (faithfulness to message, accuracy, clarity), interpreting technique (fluency of delivery, hesitation, backtracking, lag time, irritating habits, eye contact), professional conduct (preparation, knowledge of the topic, behaviour/dress code).

3.3.2.7 Australia

The NAATI marking system is specified in the NAATI Examiners' Manual (EM), which was extensively rewritten in 2005, with minor revisions in 2008. The EM includes a section outlining general principles for marking, with guidelines for marking each specific type of test included in the sections relating to translating and interpreting at each level. As many marking systems do,

as reviewed above, the current marking system is based on error detection, but the basic principle of the subsequent scoring process (including the determination of pass/fail) is what is generally referred to as a 'subtractive' or 'punitive' numerical or deduction system, where the candidate starts with 100%, and marks are then deducted according to the number and seriousness of the errors previously detected. Determination of pass/fail is then based simply on achieving an overall score of 70% or better, although for interpreting tests there are also requirements to achieve 70% minimum in each component (dialogue interpreting, consecutive interpreting, sight translation and questions on ethics and cross cultural issues). This requirement, in essence makes each component worth 100%, thus invalidating the current weightings allocated to each. This is something that we propose changing, so that skills that are found to be more important or more common for interpreters, for example, will receive higher weightings. Such a true weighting system will prevent the current situation where a candidate can obtain, for example, 90% for the dialogue interpreting component and 60% for the consecutive interpreting component and fail overall.

In the current system, the three main aspects to be assessed are:

- accuracy (of conveying the message, both overall and for any given part)
- quality of language (viewed particularly in terms of its contribution to accuracy, not merely for its own sake)
- technique (application of good practices).

According to the EM, accuracy should be given the greatest weighting and technique the least; however, the exact proportions are not specified, although the guidelines for marking translating tests are much more detailed than those for interpreting. In addition to the general guidelines just described, the specific guidelines for marking translating tests include the following:

- Markers need to differentiate between 'general' and 'isolated' errors (the former affecting whole clauses or more, the latter affecting only the immediate word).
- Markers need to penalise errors more severely when they affect accuracy than when they simply offend against lexical or grammatical usage, but the meaning is still clear.
- When a candidate produces a number of 'systemic' errors (usually, ones that indicate a fundamental ignorance of TL lexical or grammatical usage), these can be penalised as many times as they occur. By contrast, a mistranslation of a particular word that might occur multiple times throughout the text can only be penalised a maximum of three times.

In contrast to the relatively specific guidelines for marking translation tests, the guidelines for marking interpreting tests are much less specific. The suggested approach, as currently described in the EM, is as follows:

- Allocate to each segment of the dialogue, in accordance with its length, a proportion of the total available marks.
- After noting the errors made by the candidate, determine what proportion of the 'message' in each segment has been successfully conveyed, and award a mark proportionally for each segment.

At the Professional level, the current guidelines in the EM for marking sight translation and monologue interpreting suggest a somewhat indefinite assessment of how much of the whole 'message' of the text has been adequately conveyed, and awarding a proportional mark accordingly (an admittedly vague and difficult task when looking at the 'message' of a 200-word or 300-word text). If a new system is adopted, a new, improved examiners' manual will need to be produced, although many aspects of the current manual could remain.

Currently markers notify results to NAATI using a proforma, from which selected information is then communicated to the candidate. This proforma provides for:

- recording a numerical score for each part of the test, and overall
- if the candidate has failed, circling one or more letter codes indicating the types of errors that contributed significantly to that result (e.g. A = significant omissions).

Particularly if the candidate has failed, markers are expected to attach a sheet with 'narrative' feedback on the candidate's performance, highlighting areas of particular weakness and perhaps giving some brief examples, as well as suggestions for improvement.

3.4 Conclusions on marking systems

This section has reviewed the literature on marking systems and compared the different practices around the world. There are two types of marking systems that are most popular in translation testing in particuular: the error deduction system and the rubrics based approach. The error-focussed marking systems are becoming less popular, even in America and Canada. The overall weighting of errors as an obstacle to a pass mark is being reduced. Instead, assessment weightings are now more evenly distributed amongst a large number of criteria. Traditionally, T&I testing did not pay a great deal of attention to the socio-interactional nature of interpreter performance. In Translation Studies, marking of translation tests still shows little evidence of Skopos-based analysis of the function that a source text has and how this is rendered in the target text as an assessment criterion. This is starting to change in some of the marking systems that have been examined in this report.

In a rubrics-based approach, descriptors have emerged as a supplementary or replacement means of T&I assessment. Employment of descriptors is popular for many reasons: descriptors appear less punitive; descriptors can combine a number of performance indicators together that are otherwise hard to distinguish individually; some examiners dislike numerical values and descriptors allow a quantification of performance without a particular score (the exception to this is the IoL's DPSI and DiT tests); descriptors are often generalist in content and therefore more able to be adapted or applied to a large range of test formats. Research on the use of descriptors compared to error-calculated systems is scant and preliminary findings indicate that there is little difference in the final scores that examiners award (Lee, 2009; Turner, et al., 2010; Waddington, 2001). However, our most recent study showed a preference for the use of rubrics by a group of Melbourne based practitioners, educators and examiners.

After reviewing all the available marking systems, we believe that the benefits of a rubrics-based system outweigh its potential flaws, and we recommend that NAATI embark on a validation study to construct theoretically derived and empirically tested rubrics for the Australian context. Below we outline the major advantages and potential disadvantages of the rubrics-based marking system:

- They oblige markers to consider a wider range of factors in deciding on the eventual result. For instance, good sets of rubrics direct markers' attention to factors such as register, pragmatics, and the like, which can easily be overlooked in a subtractive marking system.
- Because the descriptors are phrased in terms of the candidate's performance throughout the text, rather than at discrete points, they again oblige the marker to view the candidate's performance holistically, rather than focus on particular errors.
- They encourage markers to identify positive as well as negative aspects of performance.
- Because the descriptors, if well-designed, are generally expressed in non-technical language, they could, if used as part of reporting of results to candidates, give the candidates a more meaningful and more standardised picture of their performance.

- In case a candidate disputes a result, it can be easier (albeit not automatically so) to justify the result by pointing to the descriptor selected and demonstrating how that matches the actual performance.

However, some points need to be noted about the use of rubrics-based systems (although some of these limitations can also be identified in the other marking systems):

- For the rubrics to be valid, the test construct needs to be carefully defined and carefully analysed and the criteria empirically devised (Angelelli, 2009, p. 22).
- Unless the descriptors are carefully worded, there can still be room for varying interpretations of what each level means.
- At least initially, markers used to the current NAATI system may have difficulty dealing with some of the assessment areas that they may not have had to think much about until now (e.g. pragmatics or register). This may particularly be the case, given the discussion above on the backgrounds of many NAATI examiners who have not received any formal education in interpreting and translation studies and are very likely unaware of the theories. Clearly, this draws further attention to the need for extensive training of markers, including trial marking of sample tests followed by inter-comparison amongst panel members.

We must however, highlight that the benefits to be gained from any move to a rubrics-based marking system may be effectively negated (or rendered far less worthwhile than they might otherwise have been) if the overall testing system is not also significantly overhauled. In other words, while the potential benefits of using rubrics are undoubtedly significant, without wider changes to the overall accreditation system NAATI might be at risk of simply 'tinkering around the edges'.

3.5 Test components

We stress at this point that any definitive decision on the new test design, content and weightings must be a result of empirical research. We also stress that our proposed changes are inextricably connected to the new proposed model, which suggests language screening and prescribes pre-testing training. With those two extra requirements in place, the final accreditation examination becomes less important, reducing its level of risk, as candidates will need to pass a number of assessment tasks before reaching the final accreditation examination stage.

Based on our review and on the responses from the national survey, below we provide some suggestions for possible test design, content and weightings, which may need to be changed or adapted as a result of the proposed research project.

3.5.1 Interpreting Generalist test

Skills to be tested

- Dialogue/bilateral interpreting
- Remote/telephone interpreting
- Sight translation (where applicable)[35]
- Consecutive interpreting of oral language likely to appear in community settings, such as information sessions, with repetitions and clarifications permitted[36]

---

[35] For example, in many Indigenous contexts sight translation would only occur in one direction: from English to LOTE, so the examination will need to be adapted to cater for their particular needs.
[36] Two members of the research team were in favour of a level below the Generalist, equal to the current Paraprofessional examination, which only assesses dialogue interpreting for the new and emerging languages and Aboriginal languages. The rest of

- Simultaneous/whispering interpreting
- Management skills

A note on Test authenticity

The lack of authenticity of the interpreting test is a central issue in improving interpreting testing. In addition to the components to be tested that are currently absent from the NAATI test (as outlined above), the current tests lack authenticity in the way they are delivered (via a disembodied tape recording), in the length and structure of the dialogues (which usually lack the many features of spoken discourse and can read more like written texts than oral natural dialogues), and in the penalisation of repetitions or requests for clarification (which is common practice with competent interpreters in a real situations). We therefore, strongly recommend that interpreting examinations be delivered live, that scripts reflect features of spoken language and that candidates be allowed to seek clarification and be assessed on how they manage and coordinate the interpreted situation. Recorded tests cannot assess such crucial aspects of interpreting. Furthermore, the current tests have had, in some instances, the negative effect among some training courses of shifting the focus of the training (training to pass the NAATI test rather than training candidates to become competent interpreters).

We understand that holding live examinations can be logistically difficult and may be impossible in all instances. If live testing is not always possible, we propose that tests be video recorded, so that the candidate can see the participants in the interaction and can stop them when needed. The candidate's performance should also be video recorded for marking, so that the candidate's demeanour and management skills can also be assessed.

NB: Note that all other knowledge and competencies will be assessed in the training.

3.5.2 Interpreting Specialist tests

3.5.2.1 Legal Interpreting specialisation

- Court interpreting (including bidirectional consecutive interpreting of questions and answers, simultaneous interpreting of speech into other language (English and LOTE), sight translation of documents, court protocols such as addressing the Bench, seeking clarification, providing explanations, etc) – this can be done as a simulated scripted test or, in the case of interpreters who are already practising, as observation of a real life case where the candidate works as interpreter.
- Terminology test

NB: Note that all other knowledge and competencies will be assessed in the training.

3.5.2.2 Medical Interpreting specialisation

- Simulated doctor-patient interpreted interaction (dialogue interpreting, sight translation, management skills)
- Simulated mental health interpreting situation (simultaneous spoken or whispered interpreting)
- Terminology test

NB: Note that all other knowledge and competencies will be assessed in the training.

---

the team argued for the discontinuation of the current Paraprofessional level, with the Provisional Generalist examination as a compromise. As consecutive interpreting is not common in Interpreting practice, the consecutive interpreting component is likely to carry a small weight in the new proposed examinations, thus no longer making this component the key cause of failure. Nevertheless, we stress that a validation study will be used to determine the final contents of the accreditation examinations.

3.5.2.3 Conference Interpreting specialisation

- Simultaneous interpreting in a booth of either a mock conference or a real conference

NB: Note that all other knowledge and competencies will be assessed in the training.

3.5.2.4 Business Interpreting specialisation

- Simulated multi-party interpreted interaction

NB: Note that all other knowledge and competencies will be assessed in the training.

3.5.2 Translation test

- Choice of two out of three 250 word texts to be translated into the chosen language
- Accompanying justification of translator's choices, referring to the relevant theory
- Addition of a brief to the candidate, specifying the skopos of the translation task, to determine the approach to be adopted
- Permission to use computers and other digital resources

As per the interpreting tests, the translation test will be complemented by the training and the different hurdle assessment tasks throughout the duration of the modules and upon their completion.

3.6 Issues of validity and reliability

Test validation is the process of generating evidence to support the well-foundedness of inferences about a candidate's performance based on the test scores. Test developers need to provide a clear argument for a test's validity in measuring a particular area of performance to be assessed (Kane 1992 cited in Weir, 2005).

The validity of the test, in general terms, "refers to the appropriateness of a given test or any of its component parts as a measure of what it is purported to measure. A test is said to be valid to the extent that it measures what it is supposed to measure. It follows that the term *valid* when used to describe a test should usually be accompanied by the preposition *for*. Any test then may be valid for some purposes, but not for others" (Henning, 1987, p. 89). This is an important point to remember when testing the validity of the different testing instruments we are proposing (i.e. generalist and specialist tests).

External & internal validity relate to the methods for assessing validity. Internal validity relates to studies of the test content and its perceived impact, external validity (or criterion validity) relates to the relationship between a candidate's test scores and measures of their ability beyond the test. One very important type of validity is 'context validity': the extent to which test tasks compare to real-world tasks undertaken in (in this case) translation and interpreting professional practice (Weir, 2005, p. 19). The current NAATI exams, and in particular the Interpreting exam, seem to be low in context validity, for the reasons outlined above.

It needs to be acknowledged that test developers face additional challenges when designing interpreter and translator tests, as compared to test designers in other more established fields due to a lack of empirically defined and supported models of translator and interpreter competence, and a lack of research into existing tests. The lack of research into accepted models of competence leaves test developers with little more than untested theoretical

frameworks or practitioner experience as the basis for test design, including the design of test passages and scoring rubrics. The lack of a body of research into existing tests means that there are no accepted standards for the validity and reliability of translator and interpreter tests and no tried and tested methods for undertaking this research. Below we review the few research studies into issues of validity and reliability in the field of interpreting and translation.

Clifford draws on Berger and Simon's (1995 cited in Clifford, 2001) list of principles of psychometric evaluation: reliability, equity and utility and adds a fifth principle, comparability, as one which is also important in interpreter assessment. Clifford's (2001, p. 374) descriptions of the principles appear in Table 14 below:

Table 14: Clifford's descriptions of psychometric evaluation principles

Validity	An instrument is valid if it actually measures what it was designed to measure, that is, if it allows the assessor to make inferences about the targeted competency. There are different types of validity: e.g. content, construct, predictive, instructional, consequential (Berger & Simon, 1995).
Reliability	An instrument is reliable if it provides stable results from one administration to another in comparable conditions of use (Berger & Simon, 1995). Reliability may be verified through such techniques as testing and retesting on separate occasions, or using alternate forms of a given assessment (Gipps, 1994).
Equity	Equity is the principle that instructs assessors to be aware of gaps in performance that exist among groups because of differences in familiarity, exposure and motivation on the tasks of interest (Linn, Baker, & Dunbar, 1991).
Utility	An assessment may be valid, reliable and equitable, but high cost or unreasonably elaborate procedures may prevent its use. The utility of an instrument is an indication of how practical it is to use in a given situation.
Comparability	There is comparability in an assessment if it is administered consistently, if there is common understanding of assessment criteria, and if the performance is evaluated fairly (i.e. with the same rubric by all markers; (Gipps, 1994). Interpreter assessments need to demonstrate comparability across interpreters with different working language, across yearly administrations of the assessment, across raters and even between language modes (spoken versus sign language).

For interpreting, content validity applies where a test measures interpreting abilities across a range of scenarios that are determined to be typical and common interpreting situations found across most if not all areas, as we stressed above. For the development of English-sign language tests in Canada, Russell & Malcolm (2009, p. 356) sought to ensure content validity by selecting "test segments created based on community consultation with interpreters and consumers of interpreting services, along with interpreter referral agencies, in order to plan test scenarios that are realistic and reflect the broad range of settings where ASL-English interpreters typically work". The choice and selection process for sourced materials was similar to that employed by Angelelli (2007) in a test designed for medical interpreters. Russell and Malcolm's segments were sent for review and perusal by a number of parties.

Construct validity is apparent where inferences can be legitimately made from testing performance to the theoretical constructs on which the criteria are based. Russell and Malcolm (2009) also sought to ensure construct validity in their testing apparatus. An initial test developed in Canada surveyed a number of theoretical approaches and based its perspective of performance measurement on discourse analysis. One of the most comprehensive investigations into one of the psychometric features of assessment, that of (construct) validity in interpreting testing was undertaken by Clifford (2003) who approached interpreting performance from a discourse perspective and applied three constructs (intelligibility, informativeness, style) to a revised test, given to 15 trainee or practising French-English conference interpreters in Canada. Clifford (2005, pp. 120-122) argues that previously-used tests had been unable to

distinguish between these constructs and that test-takers tended to attain either high or low scores for all of constructs together. His revised test provides a consistent relationship between test scores and statements about the related constructs (Clifford, 2005, p. 127).

Vermeiren et al. (2009, p. 303) add a further sub-category of construct validity, which is the "measure of agreement between the test and the concept or domain it is derived from. In an educational concept, this would be the relevant curriculum (knowledge and skills), in a professional context the professional standard (related to specific knowledge and skills)". Vermeiren et al. (2009) also posit 'predictive validity' as a term referring to a test taker's future performance being predicted from their prior performance. Elsewhere, others term this "maintenance" (Lysaght & Altschuld, 2000, p. 95). NAATI's introduction of a finite time-length for new accreditations, renewed through a process of revalidation, is a move that addresses this psychometric feature of testing, 'predictive validity' or 'maintenance'.

The other psychometric principle that this discussion focuses on is that of reliability. Reliability is usually sub-divided into different types: inter-rater reliability (typically examined in relation to the assessors who use a test and the reliability that a test design can ensure comparable marking outcomes amongst different assessors); 'intra-reliability' (the same assessor marking the same test for different test-takers); test-retest reliability (the same assessor marks the same test according to the same criteria after a sizeable time interval). Russell & Malcolm's (2009) revised sign-language-English testing regime for certification in Canada makes the claim that inter-rater reliability is addressed through training of raters, a high number of raters (6) for each test assessment, with three meeting collectively to mark assessment and three receiving test performance recordings and marking them individually. Vermeiren et al. (2009, p. 304) also advocate for rater training to increase rater reliability for translation marking. Inter-grader triangulation of criterion is also advocated by Vermeiren et al. (2009), who also support repeated testing for every criterion, the application of consistent indexes in assessors' marking and systematic test-retest exercises in training.

Tests therefore need to be both valid and reliable: reliable as tools that can be administered to all candidates and administered and marked in a uniform way, and valid so that they can measure what the test designer wishes to measure. A third important characteristic is that of authenticity, i.e. "the degree to which tasks on a test are similar to, and reflective of a real world situation towards which the test is targeted" (Angelelli, 2009, p. 20). In the case NAATI tests, as discussed above, hand written translation tests or interpreting tests using pre-recorded dialogues are examples artificial situations that do not reflect authenticity.

3.7 Examiner selection and training

Discussion of any changes to marking systems cannot avoid including a discussion of the examiners who would be using a changed system. Currently, the three main groups from which examiners tend to be selected (in order of frequency) are:

- practitioners
- T&I educators
- non-T&I language academics (less often now than in the past).

In some cases (especially when that language community is very small and/or newly arrived), panel members may not belong to any of the above groups, but simply be L1 speakers of the relevant LOTE.

While NAATI prefers language panels to include both L1 LOTE speakers and L1 English speakers, in the case of many languages (not only newly arrived ones) there are hardly any L1 English speakers who have the level needed to be an examiner, and the entire panel is therefore made up of L1 LOTE speakers. For translating LOTE > English, this lack is

compensated for by a panel of 'English markers', but there is no such provision for interpreting tests. Many examiners, especially those who are practitioners, may have a good instinctive sense of what is satisfactory translating or interpreting, but do not have any background in the theory of translating and interpreting (as evidenced by our survey), which may result in a tendency to take an over-literal approach to marking, or focus too much on quality-of-language issues. In addition, many examiners may have little background in the theory of assessment. This can be a particular problem with examiners for some aboriginal languages. One suggestion would be to have such examiners partnered with someone who can assist them to "*translate*" their comments and assessment into standard marks.

The statement "There should be compulsory training for all NAATI examiners" in our survey, received the highest percentage of agreement from all respondents combined (84.5%). This indicates that there is also a clear public perception that the current examiners may not be adequately qualified, thus jeopardising the credibility of NAATI testing. One possible way to improve this, in addition to compulsory training, is a through more rigorous and more transparent recruitment process. Open calls for applications should be widely advertised on a regular basis, with three-year contracts, renewable upon successful review by the panel chair. We believe that the criteria for applicants must also be amended to include Interpreting and Translation formal higher qualifications in addition to NAATI accreditation, where applicable. We believe more examiners should be recruited from the graduates of the existing NAATI approved courses. Such highly qualified examiners can also assist NAATI in training other examiners with no formal I&T education background. Many I&T graduates are practitioners and part time educators. Another benefit to NAATI is that as educators they are familiar with the NAATI examiners' manual already as well as with other assessment practices relevant to the institution for which they work.

At present, there is no consistent formalised training program for examiners. For many years, NAATI was fairly assiduous in organising training workshops at least yearly, and sometimes twice yearly, and continued membership of a panel was (at least officially) conditional on attendance at these workshops. However, because these were often held only in major capital cities (because numbers in the smaller capital cities did not make it viable), and had to be held on fixed dates, not all examiners were able to attend, and some chose not to attend for considerable periods. NAATI was often limited in its ability to enforce attendance because it could not afford to lose panel members in some smaller panels, or to pay them to attend. On the other hand, NAATI has been trialling, and now hopes to implement more widely, a system where members of each panel are brought together in one location to engage in intensive workshopping of test setting and test marking, in response to one of the recommendations in the Cook Report. Some logistical difficulties can be overcome with other methods of delivery such as online training.

The conclusion of all of this is that if any new system is to be introduced examiners must have adequate training.

4. Technology and interpreting and translating testing

This section will present a summary of the main issues that relate to the role of technology in Interpreting and Translation tests. The information was sourced from a review of the literature, the work of the working groups on technology and the results of our survey.

4.1 Technology and translation testing

"Having to sit the exam with pen and paper is nonsense. Translators should be able to sit the exam in the same environment and with the same tools as they would use in real life, that is, a word processor and internet resources. The exam cannot reflect the real ability of the translator if they are restricted to pen,

paper and dictionaries that they might not even use in daily life. I know many great translators who have failed the NAATI exam because of this" (Survey practitioner respondent)

4.1.1 The case for 'computerised' NAATI examinations

Computers have been a part of the office environment since the early 1990s; indeed they have become seemingly indispensable for work, study and leisure. The two-decade period of computer ascendancy has in turn given rise to a new generation that has experienced nothing else – the so-called 'digital natives'. It is therefore hardly surprising that questions should arise about the advantages and advisability of computerising the NAATI translator examination system. Computers also clearly afford practical advantages in data processing and handling that can greatly streamline administrative processes. Exam candidates and back-end administrators will obviously use computers in practically every facet of their working and domestic lives, so it would appear only reasonable and normal for the information revolution to finally make its presence felt in NAATI translator testing. Something of this type of subjective thinking can be glimpsed at least as far back as the 2001/2 NAATI Test Review, which summarises the findings of a lengthy and wide-ranging process commenced in 2000. Thus in the corresponding Executive Summary we find the following recommendation on test formats: "Some respondents made a specific comment that computers and electronic devices must be accepted as standard tools to be used by candidates at Translation tests" (National Accreditation Authority for Translators and Interpreters, 2002, p. 4). There is no corresponding quantitative or objective justification offered, and indeed as recently as 2011 the NAATI Board, in the Call for Expressions of Interest document for this project, identified the obvious administrative advantages but cautioned against pursuing computerisation upon appearances or perceptions alone:

> *New Technology*: With the advent of much new ICT the Board hopes that this will be taken into account in considering the practicality of proposed changes to testing. The Board believes that there are potential benefits in administration, logistics, access, assessment and reduced postage. While the Board is not advocating technological determinism in proposed models as against an evidence-based conceptual framework, it is hoped that the benefits of available and emerging technology will be captured as much as possible. In holding that view the Board notes that it does not want NAATI 'captured' by unique technology that is difficult to maintain and may pose difficulties in access by the typical community of NAATI clients (National Accreditation Authority for Translators and Interpreters, 2010, p. 6 see Current Views of the Board)

The Board warns that the apparent attractions of using new technology in testing cannot ignore the potential drawbacks, including issues of access and equity for candidates and possible technological dependence of NAATI itself. In short, technology must be the servant, and not the master. Caveats and potential pitfalls aside, the call for using computers in examinations seems so strong that the main questions appear to be ones of 'when' and 'how' rather than simply 'if'. The 2001/2 NAATI review already made recommendations for the use of computers in translation testing with the main argument that the current pen and paper tests do not reflect the current practice of translation practitioners, as expressed by the quotation above from one of the respondents to our survey. The 2001/2 Review identified two type of resources that can be available to translators: 1: Dictionaries, Glossaries, Parallel Texts (texts from the same genre or on a similar topic), Terminology databases, On-line and off-line electronic resources, Computers, Software (spell check, grammar check), Internet; and 2. Email, Mobile Phones, Translation Memory. The recommendation from the Review was that only resources under category 1 should be allowed in the NAATI translation test (NAATI Test Review Translators Group, 2001, pp. 16-17).

The Cook Report also strongly advocates the use of computers in NAATI Professional level testing. The authors of the Report offer the following reasons for the need to introduce computerised translation tests:

- It reflects practice in the industry.
- It reflects practice at some levels of NAATI testing.
- It would allow efficiencies of scale in terms of marking time.
- It would allow for the organisation's stated practice of independent double marking to be implemented (Cook & Dixon, 2005).

Cook and Dixon subsequently explore the logistics required to implement computer testing, such as the number of computers that NAATI would need to own or hire, and the technical assistance required to enable access to LOTE scripts while ensuring functionality limits elsewhere - such as Internet resources. On balance, the authors ultimately consider that the advantages gained by doing tests on computers would outweigh the additional upfront expenses, and proceed to suggest ways of implementing it.

The computerisation options suggested in the Cook Report are four in number, wherein NAATI respectively 1) purchases and maintains the computers and associated software; 2) hires computers and testing venues from schools and universities; 3) allows candidates to use own computers (as already implemented in Advanced Level testing), and 4) negotiates with a computer manufacturer to provide candidates with a laptop (including appropriate software) on which to undertake the testing, at a favourable cost. Two recommendations are finally made:

> *Recommendation 1*: In-principle approval be given for the use of computers by candidates presenting for the Professional Level Translation Test, subject to there being a phase-in period of up to five years, where candidates have the choice of handwriting their responses or using a computer to provide their responses.

> *Recommendation 2*: NAATI undertake a feasibility study of the options that are viable to phase in the use of computers for translating tests (Cook & Dixon, 2005, pp. 22-25).

4.1.2 NAATI translator examiners' comments

In addition to the national survey, which elicited some voluntary comments on the use of technology, the group of twelve NAATI markers who gathered in Melbourne for the Rubrics trial, were also informally polled on their thoughts about computer use. While the aims were largely qualitative, certain responses admit some type of quantification and are presented in the table below in Table 15.

Table 15: NAATI Examiners' comments on computer use for translation examinations

NAATI Marker Reponses	Number of markers (n=11)
Tests should be:	
Conducted on a computer	10
Handwritten	1
Candidates have the option to use the keyboard or to hand write it	1
In testing 'on a computer', the hardware should be:	
NAATI owned or leased	6
Candidate allowed to bring their own computer	3
Aids allowed:	

No aids	2
Hardcopy only	3
Dictionaries only (hardcopy and softcopy)	4
Digital aids only (no hardcopy)	3
If digital aids allowed, which ones?	
Spelling & grammar checkers and other aids offered in Word	4
Off-line dictionaries only	5
Dictionaries only, online and off-line	3
Unrestricted (but not TM or MT)	7
Unrestricted (including TM and MT)	2
On choosing a platform to manage the testing process priority should be given to:	
Web-based, with candidate working on the browser, not the hard drive	7
Security, candidates not allowed to keep a copy of test	7
User friendliness	5
Allowing for remote testing	1

The prompts admittedly allowed respondents to support more than one option, but the numbers nonetheless give an insight into relative levels of importance or acceptance. Significantly, the vast majority (ten) seemed in favour of using computers *only* (since an either/or response was also available). Of what might be termed the two dissenters, the one who supported retaining the current handwritten form only felt that "some candidates may not be sufficiently familiarised with computers"[37]. The other favoured implementing computer examinations but also retaining the handwriting option. With regards to resources, four favoured dictionaries only (both hard and softcopy), but if digital aids were allowed then seven favoured unlimited aids such as spelling and grammar checking, and online and offline dictionaries. These responses are consistent with the recommendations of the 2001/2 Review and can be summarised as a list of what would seem to be an acceptable format for Translator testing:

- conducted on computer (NAATI-owned/hired)
- web-based (i.e., via central server, not a local computer hard drive)
- word processing allowed (grammar, spelling etc.)
- hardcopy and electronic dictionaries
- internet, but no Translation Memory or Machine Translation.

Security was regarded overall as more important than user-friendliness, and although web-based exams would make remote delivery possible, that scenario was only accepted by one respondent. We received also a contribution from a translator from Japanese into English who explained that translators working from different scripts, particularly with English as the L1, face additional difficulties which could in part be solved by testing in a digital environment.

4.1.3 Fundamental questions

Based on our brief review above, the main issues concerning the use of computers in Translator testing have been addressed as follows:

---

[37] No reason given, though we might speculate a concern for emerging languages and/or demographically and socio-economically differentiated groups (e.g. age/gender/ethnicity/educational opportunities/country of origin).

1. Should future testing be conducted using a computer instead of being handwritten?
2. Whatever the decision taken on 1, which translation aids should be allowed?
3. Whatever the decisions taken on 1 and 2, how should the process be managed to ensure security and efficiency?

The above assumes that computers must be used in accreditation tests to reflect current professional practice. However, there is a fundamental question that needs to be assessed before deciding on this step: what is the main aim of the NAATI accreditation examination? In other words, does it aim to assess a candidate's basic, core translation skills as a novice translator or does it aim to assess an experienced translator's ability to produce a professional translation product? It appears to us that the first aim could be applicable to the current NAATI accreditation system which does not require any pre-testing training, and that the second aim could more adequately apply to our new suggested accreditation system which will involve compulsory hours of training, including training on the use of translation technologies, as is the case in current formal translation courses. We will therefore continue the discussion in support of the use of computers in accreditation tests and discuss issues relating to logistics and security.

4.1.4 Exam environments

With electronic environments, the number of potential aids is large and increasing. They can be deployed or searched rapidly, and the results or answers can typically be retrieved and imported directly into other digital media. In addition, electronic environments can be local (hard drive or private network, more secure) or online (internet or public network, less secure). Obvious electronic aids of specific interest to language and translation exams would include:

- word processors and text editors in general
- electronic dictionaries and glossaries
- corpus materials, both local and online
- translation applications such as translation memory (TM) and machine translation (MT) software.

Word processors in particular must be classed as an aid and not simply a medium, because they do not only permit *text recording*, but also *text manipulation* (spelling and grammar checking, copying and pasting, storage, etc.).

4.1.5 Exam mediation

In what we will term *Host Mediation*, the exam process is controlled directly by the examining entity itself, and is thus more readily subject to its strictures – including imposition of the test delivery and recording media. This offers the most direct guarantees regarding candidate identity, protection of test materials (non-dissemination), cheating (e.g. restrictions on test aids), and time limits. The Host category would also cover in-house electronic environments, both local and on-line, over which the hosting institution has full control. Notionally it could also accommodate remote situations (e.g. virtual classrooms) provided that it is possible to implement safeguards – such as might be found in secure corporate or defence environments (video links, biometric identification, monitored logins and usage etc) – but the logistics obviously entail great sophistication and expense.

With *Proxy Mediation* the examining entity outsources supervisory control (full or partial) over exam implementation and/or marking to an external agent. For example, Microsoft Certification Programmes employ the services of the Thomson Prometric company (now a subsidiary of Educational Testing Service – ETS); Australian high school examinations are delivered, supervised and recorded in-house but marked externally. There is clearly no bar to a Proxy being engaged to replicate precisely the same strictures and controls as the Host entity would

apply itself, but the process is now removed from the Host, with a corresponding decentralisation of control and accountability.

In *Client Mediation* the candidate ('client'[38]) plays a self-supervisory role that may be more or less significant depending upon the circumstances. Thus, if an institution (or Proxy thereof) examines candidates on-site, but allows them to use their own computers (or simply, say, personal reference materials or aids), this inevitably relinquishes a degree of control to another (highly interested) party. In the case of certain voluntary or non-prescriptive qualifications, candidate autonomy may be extreme (e.g. undertaking exams at home using own equipment, as with SDL Trados Certification), and strict time limits may not always apply (untimed exams). Such self directed tests are also encountered in self-directed studies, correspondence courses, open learning, adult education, and online distance-mode studies. In the global translation marketplace, this is also a common type of test procedure applied to translators seeking to join agency panels, as identified in our review of accreditation systems around the world (see point 2 above). Client mediation thus supposes inherently lower guarantees, posing particular challenges with the copying of material, vetting of candidate identity and restriction of exam aids – although if time limits can be imposed (e.g. online examination with timed logins), these can at least help curtail opportunities to seek unfair advantage.

4.1.6 Use of computers in examinations by other international bodies

Case Study 1 – Chartered Institute of Linguists
As explained above, The Chartered Institute of Linguists (or commonly the Institute of Linguists, IoL) is a UK organisation which offers courses as well as examinations. This body allows the use of computers for examinations but not the use of machine translation programs or access to the Internet. Some examination centres offer IT equipment and facilities that do not contain software with translating programs, but they accept no responsibility for IT facilities that malfunction on the day of the examination. The use of personal laptops or electronic notebooks is not allowed as it is difficult, if not impossible, to monitor the programs installed. For the same reason, electronic dictionaries may not be used. At the time of marking, no distinction is made by examiners between candidates who word-process their scripts and those who handwrite. However, all candidates are advised to present their work well (paying attention to spacing and formatting) and, if handwriting their examination, to write legibly and to ensure that crossings-out and insertions are clear.

Case Study 2 – American Translators Association
As also outlined above, the ATA is a professional association, which offers certification examinations as well as professional development courses, seminars and conferences. The ATA is planning to make keyboarded examinations possible, but currently the examinations are handwritten. No electronic equipment is currently permitted in the examination room. At the certification exam sitting, an exception is made for disabled persons, who may write their examination on a non-memory typewriter. Accordingly, the only current technological concession is aimed at access and equity. Otherwise, the exam environment is strictly controlled, carefully specifying that typewriters, where allowed, must be of non-memory type (thus preventing surreptitious transcriptions of exams or material). Nevertheless, they are currently having discussions on an imminent move to computerised certification examinations.

Their proposed new format combines old-fashioned elements (on-site Invigilators) with ultra-modern ones such as centralised server control, with fully digital exam script delivery and recording. There is apparently mediation through a proxy – Amazon – and it is also not clear whether the exam sitting room will be located at ATA, or some other premises. Candidate

---

[38] The term is not used ill-advisedly, since the modern approach to education tends to view students as paying customers, contrasted against a traditional teacher-disciple paradigm in which students (and examinees) are passive and subordinate.

autonomy is minimal, unless they can somehow bypass server control. At face value, ATA's proposed system appears safe and workable. But as we have repeatedly had occasion to observe, whenever a computerised exam environment is contemplated, security management in all its aspects (candidate identity, exam theft, fraud or manipulation) becomes vulnerable, and the solutions complex – and by extension, expensive. The apparatus, planning and execution for ATA's proposed process are all necessarily more elaborate. For example, allowing a candidate's computer to act as a terminal to a host must entail installing some form of client software and presumably firewalls and malware protection must be in place to prevent malicious circumvention and re-routing. ATA will also provide for a lengthy transition period in which keyboarded and handwritten exams will co-exist, which seems to be a necessary measure for any introduction of technology in translation tests.

4.1.7 Conclusions on technology and translation testing

Translators are constantly honing their resourcefulness through internet searching of dictionaries, glossaries and corpus material. Also thanks to computers, the fruits of such effort – which formerly had to be painstakingly hand-catalogued or even repeated anew - have now been made cumulative through ease of storage and the leveraging of past 'legacy' material held in data-basing and retrieval systems (such as Translation Memory applications). An important feature of practising translators is not simply how much they know, but how much they can find out – in the briefest possible time. It is no exaggeration to say that computers and computer proficiency form the cornerstone of their working environment. By default, a computerised environment is of clear benefit to translators. The question is whether it is desirable to accord them such benefit under examination, and if so, then to what extent.

One argument for imposing limits on exam computerisation is that technological advancement in the digital world is rapid, and in the space of only a few years some new technologies may seem objectively unfair when contrasted against what candidates in earlier years had available. Potential solutions to this difficulty might entail placing a ceiling on permissible exam technology – e.g. word processor only. Whatever the course taken, each approach yields 'frozen' test situations that will become antiquated as the real world moves on, while their unwieldiness of implementation is apparent.

On the other hand, if we accept that resourcefulness is a key attribute of the translator (which was supported by the results of our survey), then a strong case can be made for simply replicating the modern working environment and allowing candidates unfettered access to whatever computerised resources they wish – even Machine Translation, which is now a common adjunct to commercial Translation Memory suites and a tool that is taught in the current formal translation courses. Replicating professional working conditions in turn suggests that candidates might sit exams remotely at home, precisely so they can demonstrate and use whatever resources they have had the perspicacity to acquire. Removal of direct supervision then raises the concern of identity and exam fraud, but this can be lessened if the candidate has already been 'captured' by and is known to the system through having completed the pre-test compulsory training and accompanying tests[39]. We therefore recommend that this be the course taken for translation examinations.

4.2 Role and use of technology in interpreting

In general, an equivalent of the notion of computer-assisted translation does not currently exist for interpreting. Contemporary technology now offers many possibilities for the ready inter-

---

[39] Or consider the case of the Institute of Linguists, which allows candidates five years in which to complete the exam cycle. Apart from simply accommodating schedules of candidates, it adds an important diachronic dimension to their contact with the Institute (one obvious flaw of 'one-shot' accreditation being that the examination constitutes the first and only contact between examinee and examining entity).

lingual transfer of items and texts. Laptops, notebooks, handheld personal digital assistants, together with voice recognition technology now offer 'instantly' translated text to interpreters, sometimes even in spoken form. However, as Donovan (2006), Veisbergs (2007) and Winteringham (2010) conclude, the immediate nature of interpreting makes recourse to textual sources very impractical, if not impossible.

Reference to the use of technology in the practice of interpreting is nevertheless relevant to a discussion on testing, as test design should conform to the requirement of 'authenticity'. The test design and content should reflect conditions and processes that interpreters encounter in their professional lives as practitioners. The sections below will describe the types of interpreting technology available, the types of interpreting assignments conducted through the use of technology, the research conducted in relation to such technology, the types of tests that are conducted using technology and finally the uses of technology to facilitate training.

4.2.1 Telephone interpreting

"Telephone interpreting protocols should be included in NAATI test" (Victorian agency survey respondent)

Telephone interpreting, trialled for the first time in Australia in 1973 (Kelly, 2008, p. 5), now occupies a standard place in the provision of interpreting services not only in Australia, but in most other Anglophone countries of the New World, in western Europe and increasingly in other areas of the world. The market of large telephone interpreting companies such as Language Line is now global and this company now markets its services to customers worldwide. Most telephone interpreting providers are private, although the world's second-largest, Australia's TIS, is still publicly-funded. Telephone interpreting has been widely used in medical/healthcare situations since the 1980s (Hornberger, 1998; Kuo & Fagan, 1999; Lee, Batal, Maselli, & Kutner, 2002; Leman, 1997). In some contexts, telephone interpreting can be the default means of providing interpreting services: one major health provider in Melbourne has adopted, from the start of 2012, a policy of telephone interpreting as the preferred choice for consultations of sixty minutes or less. In a study of his own and others' data, Rosenberg (2007) finds that two-thirds of telephone interpreting assignments were healthcare related and one third commercial. Elsewhere, Chesher et al. (2003, p. 283) report that amongst community interpreters in several countries, the proportion of telephone interpreting is comparable to that of face-to-face interpreting. Kelly's (2008) comprehensive description of logistic, ethical and personal management issues that pertain to telephone interpreting relate not only to all community interpreting settings but to a variety of others such as business and tourism. Ozolins (2011) reports on the world providers of telephone interpreting services, such as US-based Language Line and Cyracom and Manpower Business Solutions (The Netherlands). Many of these telephone interpreting agencies offer testing. The NAATI accreditation Interpreter tests, on the other hand, have never included any aspects of telephone interpreting, and although this skill is covered by some current formal interpreting courses, it does not feature highly in any of them.

4.2.1.1 Telephone Interpreting Test – the Language Line interpreter skills test

The largest single global provider of telephone interpreting services is Language Line (Ozolins, 2011), which primarily offers telephone interpreting services, but also on-site interpreting, translation, as well as training and testing. Language Line offers two training courses: fundamentals of interpreter training and advanced medical training for interpreters. The courses are offered over the phone, with advice provided by instructors about modules and role-plays enacted, while the trainee is left to work with training manuals that focus strongly on terminology. Language Line offers a larger number of tests, all delivered by telephone: language proficiency test (English or LOTE), interpreter skills test, medical certification test and court certification test. The only way to have access to the test is by sitting for it. For this purpose, Dr Hlavac contacted Language Line and booked in for an interpreter skills test for 24

February 2012 (The interpreter skills test is presented as the "entry level assessment for working interpreters" on the Language Line website). Information supplied to test candidates prior to the test set out the length of time, role of examiner (as both the English-speaker and LOTE-speaker), five criteria for assessment (accuracy, listening and retention skills, grammar, knowledge of terminology and interpreting style), paper-copy reference sources allowed, allowances for requests for clarification or repetition. Test candidates are also supplied with a transcript of a model dialogue between an English-speaking healthcare employee and a LOTE-speaking patient. No special telephone equipment is required to attempt the test other than a keyphone telephone. The candidate was contacted on the day of the test by the examiner who explained the format of the test and repeated the protocols expected of interpreter performance – use of first person, recommendation to take notes and allowance for requests for clarification. During the explanation given to the author before the test, the examiner emphasised that requests for clarification are not sanctioned and that the test candidate should ask for repetition or clarification if medically-related information heard was unclear or not retained. Further, the examiner explained that test candidates were welcome to refer to medical dictionaries to clarify the use and translation of medical terms. These instructions differ from those provided to NAATI test candidates who are penalised for two or more requests for repetition and who are not permitted to refer to dictionaries during tests. The protocols for the Language Line test are therefore, in some way, adapted to the situation that successful candidates are likely to find themselves in, should they commence work as healthcare telephone interpreters. In such situations, where an interpreter requires clarification of medical terms or diagnosis, this requirement overrides the general need in interpreting interactions to maintain a normal flow of information exchange. The test itself contains one dialogue only for which the test candidate interprets bi-directionally. There are no sight translation tasks, no speech interpreting, no ethics questions, and no questions on the social or cultural features of either language community. No information is elicited on a test candidate's prior education, details of language acquisition, occupational experience, general aptitude or motivation. No screening procedure is required before admission to the test. The Language Line testing system appears to work from the premise that test candidates are likely to self-select for test admission or be nominated by their employer or other organisation to attempt the test due to linguistic, occupational or other personal attributes that recommend a candidate to consider testing and to gain certification.

The interpreted dialogue in the test lasted eighteen minutes and generally followed the format contained in the sample. The examiner explained that there would be salutations exchanged at the start and at the end of the dialogue and the conclusion of the dialogue would be clearly indicated by the examiner. The examiner also explained that the test would be recorded and that there may be slightly longer pauses between exchanges to allow a clear distinction in the recording between source speech and interpreted turns.

During the test, Dr Hlavac made notes, requested clarification once, and otherwise adopted the role of a professional interpreter who has extensive experience in on-site interpreting and some experience in telephone interpreting. Upon completion of the dialogue, the examiner informed the tester that the test was over and that an assessment of performance would be made on the basis of the recording. Information was not provided on the number of assessors who would evaluate the recording. The assessment would provide the basis for a results report that would be sent to the test candidate within ten working days.

With regards to the macro-level (or psychometric) features of the testing procedure, some features could not be evaluated. The author had no access to other examples of the Language Line interpreter skills test and therefore cross-test consistency and the overall reliability cannot be evaluated. The actual test was congruent to the sample test provided in content, number of words per turn, terminological specialisation, grammatical complexity of utterances within turns, and a variety of turns containing different numbers of key messages. In terms of the test's

relationship to the content and activities required from a healthcare interpreter, these appeared to be congruent and the test conforms to the feature of authenticity.

The Language Line test has a focus which is specific to the means of communication that test-takers are potentially going to use as practitioners (as telephone interpreters) and a focus which is specific to the field of telephone interpreting that its services typically relate to (i.e., healthcare interpreting). The specific focus of the test design accounts for the nature of the test in comparison to the on-site NAATI test. The main features that sufficiently distinguish telephone interpreting from on-site interpreting for the former to justify a separate testing structure are the following:

- Role relationship protocols (e.g. introductions, questions on prior experience in working with interpreters, explanation of interpreter's role, terms of address, directions on length of turn etc.) can be established in verbal form only (for spoken languages).
- Confidentiality and privacy protocols need to be overtly conveyed where visual information is lacking on others present or within earshot.
- Identity and role establishment in multi-party interactions
- Higher frequency of interlocutors using third person as the other interlocutor remains unseen.
- Requests for clarification and repetition
- Strategies to deal with overlapping speech or signing
- Management of interaction.

4.2.2 Video-link and remote interpreting

While many studies appear to use the terms interchangeably, one study (Braun & Taylor, 2011, p. 205) distinguishes between the two, defining video-link or "videoconference interpreting (VCI)" as interpreting which takes place for participants located at two different locations such as a court and a prison. Braun et al. (2011) define remote interpreting as a situation in which speakers of different languages are located in the same place using the services of an interpreter located elsewhere and connected via video. Video-link technology is commonly argued by its advocates and those seeking to cut the cost of physically present interpretation, to be almost identical to physically present interpretation by providing not only the audio but visual presence of the interpreter. A large-scale project was undertaken by Moser-Mercer (2003) for the International Association of Conference Interpreters (hereafter: AIIC), entitled *Remote interpreting and an assessment of human factors and performance parameters*. The human factors that the study relates to are "job design, task analysis and mental workload" while the performance parameters are "logistics and technical components, i.e., equipment and installation, connections and video coverage". In further research, Moser-Mercer (2005) identifies accelerated fatigue and a general feeling of 'disengagement' amongst interpreters who were not provided with views of speakers or their situations.

Video-link interpreting is now also regularly used in prison and remand situations. In one of the few studies to address not only interpreters' but also others' (e.g. court clerk, defence advocate, prisoner) experiences, Fowler (2007) reports serious problems in the acoustic and visual access to source speakers, leading to constant requests for repetition and instances of miscommunication. Despite these findings, remote interpreting via video link is increasingly being used, both in conference and community settings, and it is a mode of interpreting that cannot be ignored. The expansion of video-link and remote interpreting in Europe and in European courtrooms over the last decade precipitated interest in this medium and led to EU funding for the AVIDICUS project, led by Sabine Braun. The aims of the AVIDICUS project were to evaluate the quality of video-mediated interpreting in criminal proceedings and its viability from an interpreter's point of view. The final reports of the AVIDICUS project were published in 2011, containing a list of recommendations (Braun, 2011) and modules for interpreting students,

legal interpreters and legal practitioners (Braun, et al., 2011). The AVIDICUS project also developed modules for delivery to trainee interpreters, practising legal interpreters and legal practitioners. Braun et al. (2011) report that clear majorities of both trainee as well as practising interpreters strongly support training specific to these means of interpreting. The implication of such a finding is that these should be strongly considered as a component of our suggested pre-test training and possibly of the interpreter test itself.

4.2.3 Technology that facilitates simultaneous or consecutive interpreting

In addition to conference interpreting equipment, other technological advances are being trialled to assist interpreters in their work. Digital pens are now being trialled as a means of offering a new mode of interpreting – "Consec-Simul", a consecutive rendering by simultaneously interpreting the source speech digitally recorded by the pen and immediately played back through earplugs , giving the possibility to the interpreter to listen to the source text for a second time (Orlando, 2010). This type of technology is not directly related to interpreter testing as it is currently conceived by most certifying authorities, however, smart pens or other digital recorders with a playback function can be considered as a permitted resource for candidates in interpreter examinations[40].

4.2.4 Survey of practitioners and examiners and their use of technology

This section presents the experiences and attitudes of two groups of key protagonists in Australia towards technological innovations: interpreter practitioners and interpreting examiners. Approximately forty potential participants (both practitioners and examiners) were contacted to participate in the survey. Fifteen informants responded as interpreter practitioners, and eleven as examiners. The sample is too small to be considered representative of the experiences and attitudes of the large number of interpreter practitioners and the dozens of examiners in Australia, but the responses gained from the survey provide some interesting insights of possible experiences and attitudes.

The questionnaire was made available in hard-copy form to potential informants who attended the forum of invited practitioners, agencies and examiners held on 21 February 2012 at RMIT, Melbourne, organised jointly by RMIT and Monash University as two institutions who have working parties for this joint project. Eleven practitioner questionnaires were completed by the attendees, and seven examiner questionnaires were completed by the attendees at this joint forum. The remaining surveys were completed by practitioners and examiners known to the researchers who responded to a global email invitation to participate and who provided their anonymous responses via an electronic Survey Monkey address.

4.2.4.1 Survey results on use of technology by interpreting practitioners

The survey for interpreter practitioners contained eight questions. The questions were constructed to capture informants' responses in regard to technology used for on-site interpreting and experience in telephone, video-link or remote interpreting. Further, attitudinal responses were sought in regard to components of the current interpreting test and the role of technology. In general 60% of the practitioner informants have had no contact with telephone, video-link or remote interpreting. The remaining 40% have experience in telephone interpreting, most of them also with video-link interpreting. Apart from one informant who reports using online dictionaries when telephone interpreting, none of the informants report any other uses of contemporary technology in the performance of interpreting. Laptops and hand-held devices are used for preparation and liaison with clients/agencies but not for verbal or signed inter-lingual transfer. 80% of informants believe that in the future, there will be further technological innovations that will be used in interpreting interactions. For most of those who believe this, their

---

[40] On the topic of digital technology and "Consec-Simul" aka SimConsec, see also Hamidi & Pöchhacker (2007)

belief relates largely to the use of video-link technology, with which a majority of informants had not yet had contact. In other words, informants see the means through which communication between participants in an interpreting interaction as amenable to technological change, but not the activity of inter-lingual transfer as such. There is scepticism that technological innovations will be able to do much more than prepare practitioners for assignments and as an aide for some forms of telephone interpreting where recourse to online sources is considered logistically possible. Overall, while a majority of informants do not report experience with telephone, video-link or remote interpreting, there is a widespread consensus amongst informants that these means of interpretation will become more common and widespread.

4.2.4.2 Survey results from practitioners and examiners on the use of technology for interpreting tests

Practitioners were generally in favour of audio and video recording of candidates' performance. Many, however, preferred audio recording only to protect candidates' anonymity where there is a possibility in smaller LOTE groups that examiners may know the candidates. Others see the importance of video-recording to show candidates' inter-personal skills, demonstration of role relationship, coordination skills and use of paralinguistic markers. For signed interpreting testing, video-recording remains essential.

With regards to remote or distance interpreting, only 20% of practitioners supported this as a good idea. Advantages of remote testing nominated by informants include: lower travel costs and greater access to candidates in remote areas. Some disadvantages included: doubts about transmission quality, requirement to train examiners and conduct pre-test training for candidates. Others also mentioned the difficulty that candidates could have in connecting to the (test) discourse environment and that physical presence is what is required in most interpreting interactions which are on-site.

The examiners were evenly spread on their attitudes towards the benefits of advanced technology for testing. There were neutral and negative responses which cautioned against video-link technology as a communication means for testing, highlighting the following as potential problems: unfamiliar technology as a possible distraction for test candidates, stress and lower performance in the event of technical problems, and doubts that variable bandwidth could ensure good video reception for both candidate and examiner.

In the process of examining candidates' performance and interacting with other examiners, almost all examiners expressed an interest in video-link up and/or online exchange with other examiners and even restricted-access web pages with a repository that stores examiners' reports and allows other examiners access to these reports.

In general, there are very mixed responses to the idea of video-link testing – while many informants can see merit in it through a widening of access to testing for previously disenfranchised groups, others have concerns about the quality and feasibility of video-link/remote testing as a fair and reliable means of testing. These responses also indicate that if video-link/remote testing is considered, it is to be considered as a test that would require pre-test training to familiarise the test candidate not only with the technical equipment that they would use in the test, but the altered discourse and personal protocols that remote communication bears in comparison to face-to-face testing.

4.2.5 Testing of interpreting candidates using technology

Little research has been done on the effectiveness of remote testing in interpreting. In two studies, which follow on from Ko's (2006) earlier research, Chen and Ko (2010; 2011) report on trialling online tests with trainees at the end of a period of online study. These studies, based on the same data sample of online tests, are of great interest as they document the online delivery of all parts of the NAATI professional interpreter exam. Their sample is based on online delivery

of dialogue interpreting, sight translation, consecutive interpreting and ethical questions. The technical specifications of the testing circumstances were specific to the training that preceded testing. A web-based Collaborative Cyber Community (3C) was used for testing purposes (2010, p. 155). Overall, the technology used in Chen & Ko's (2010) test was able to accomplish the requirements of the NAATI test in regard to test delivery and recording of candidate performance for all components of the test. Chen & Ko's 2010 and 2011 studies advocate further development and trialling of online testing and are hopefully likely to continue in this direction themselves. The implication of their study is that remote testing through computers is a realistic possibility and this development should be considered very strongly where NAATI may move to allow the use of computers for interpreting and translation testing.

4.2.5.1 Testing the technology: a case study

As there is no further research into remote interpreting testing, the researchers conducted a simulated interpreter test via video-link between a test candidate and a test administrator who played a dialogue from a NAATI sample professional test and who supplied a document electronically as a sight translation exercise. The simulated test sought to examine the following:

- technical controls available to participants to regulate their appearance and reception, i.e. volume production and reception, zoom functions for visual reception of other participant, seating and microphone protocols, establishment of 'extra-test' protocols in the event of technical problems
- camera coverage of candidate's location with regard to the presence of others in their near proximity
- identity verification and explanation of how the test will be conducted
- quality of audio and video reception between candidate and test administrator
- extraneous influences such as the millisecond delays between receipt of picture and receipt of the accompanying sound and negotiation protocols to avoid "speaking over" the person whose voice continues to be heard for a millisecond or so after speaking has finished from the video image. Echo can also be a feature of some video-link interactions.

The video-link was arranged between two campuses of Monash University, through its Video Conference Services. The system used was Tandberg Edge 95MXP with a transmission speed of up to 4 mg. Video input was provided by a video camera and video output was provided by two televisions. Audio input was provided by a microphone and audio output through speakers located in one of the televisions. A system computer tied together these components, initiating and maintaining the data linkage via the network. The entire simulated test was recorded.

In the simulated test, one screen contained a screen shot of the examiner's computer showing the test items (i.e. tracks from the sample test CD and a Word document for the sight translation task). The other screen followed a voice-activated switch (VAS). This means that the multipoint control unit (usually set up or controlled by technical support staff) adopts a setting that switches the endpoint that can be seen by the other endpoint by the level of one's voice. This setting is important for the record function, because the recorded version of the interaction followed the voice-activated switch protocol – the recording showed in one of the screens the participant that was speaking at that time and switched as each participant finished speaking and the other started speaking. During the actual interaction, the VAS was not operating: on one of the monitors, the participant could continuously see the other participant and could see him/herself in a small box in the top right hand corner of the same screen. The other television screen, as stated, showed the desktop of the computer that the examiner was using to access the dialogues and sight translation document. Both participants were able to establish optimum input and reception of audio and visual features through trialling different microphone positions and seating arrangements. These features are important to establish in pre-test contact. The

dialogue interpreting exercise was led by the test administrator who pressed the play and pause button to regulate audio output. The test candidate was able to receive audio output from the test administrator clearly. The test candidate needed also to signal to the test administrator when each interpreted turn had been completed. This simulated dialogue interpreting exercise was accomplished in a similar fashion to conventional on-site dialogue interpreting testing.

The sight translation exercise was provided to the test candidate by the test administrator as a Word document, on the screen. As an electronic interface, the candidate could not make notes on the document to be sight translated. This means that a test candidate would be required to make notes on a separate document (e.g. their own notepad) which is less convenient. However, it is conceivable that in the future, electronically-delivered texts, such as on hand-held devices, kindles, tablets, etc. could become more common in everyday interpreting practice and an electronic source text as sight translation task. The sight translation task was recorded in the same way as the dialogue interpreting task.

The consecutive interpreting component and ethical questions were not trialled as the recording of these is no different from that of the previous two components. The test was concluded after about forty-five minutes and a request was made to technical support staff to supply both participants with a recording of the interaction.

4.2.5.1.1 Quality of recording from video-link test

Although the video and audio transmitted live between the two video-link points was of high quality, the video and audio recording of the interaction was not of high quality. Monash University technical support staff had alerted the researcher to the fact that recording quality was not high due to the storage memory limitations of the recording facilities. The large volume of a forty-five minute file means that recording from it may occur at a low setting, resulting in a low quality recording. A low quality recording can also result from the recording being optimised for high- to medium-speed download or from problems in data capture. Technical support staff informed us that video-link up equipment is principally designed to maximise real-time video and audio transmission and the same systems do not have the same quality of recording. In any case, the recording that was made available to us was not of a quality that could be supplied to examiners for them to adequately assess a test candidate's performance. The recording contained long video and audio delays of up to three seconds in which both video and audio output were frozen and comprehension of the source dialogue and the interpreted dialogue, were severely distorted. Distortions in both the visual and audio output continued for most of the recording. It has not been possible to establish, through enquires to CISCO (Tandberg), whether current recording facilities can be optimised, to ensure a higher quality of recording. At present, the recording quality that was provided by Video Conference Services at Monash University was not of a standard that allows for adequate video-link as a means for test delivery.

Other synchronous audio video communication systems such as Skype allow for audio video recording of transmissions through programs such as EVAER (Excellent Video and Audio Recorder). However, the quality of SKYPE video and audio output is variable and not of sufficient quality to be a reliable platform for testing. At present, Chen & Ko's (2010) study which was based on the use of a synchronous audio video learning platform with a finite number of participants and with technical specifications that allow synchronous recording without transferral to a recording source file, appears as the best available model where minimum technical specifications can be guaranteed.

4.2.6 Technology that can assist with training

Technological advances in speech recording, virtual learning environments that replicate face-to-face synchronous communication, remote interpreting facilities, independent learning and asynchronous assessment are developments which have extended the range of facilities,

settings and modes that interpreter training can now offer to trainees (cf. Mouzourakis, 2008). These developments greatly assist trainers and trainees alike in simulating the types of interactions that practising interpreters find themselves in. Technology assisted distance education can be considered for our suggested pre-testing compulsory modules, especially for emerging languages for which there are no current courses.

In Australia, Ko (2008) reports teaching interpreting through distance education in a study that compared off-campus teaching to on-campus teaching. The four modes of interaction with off-campus students were (sound-only, multi-group) teleconferencing (for dialogue interpreting, consecutive interpreting and some sight translation), telephone (for consecutive interpreting sight translation in pairs), bulletin board (study materials and texts) and email (general and specific correspondence with students). Ko's (2008) comparison of pre-test, final and independent test results for control groups of on-campus and off-campus trainees showed no significant differences in performance, suggesting that trainees can learn and become skilled in interpreting through distant education with no disadvantage compared to on-campus trainees. Testing was not conducted through distance education in Ko's (2006) study, but in a subsequent one (Chen & Ko, 2010).

In Norway, Skaaden and Wattne (2009) report on teaching interpreting through remote (i.e., distance education) means to 116 students for a twelve-month course on community interpreting. Remote refers here to web-based delivery of teaching materials to trainees and all interaction between instructors and trainees and amongst trainees themselves being web-based. The students had already gone through a screening procedure consisting of a bilingual lexical knowledge test and oral test in which simulated consecutive interpretation of fifteen to twenty short sequences was recorded. Such a remote education course was designed in Norway, which, like Australia, has a thin population spread and long distances are a disincentive to student participation. Time constraints are also considered a motivation.

In the United States, technological advances in remote communication have given rise to a large number of providers of educational training who now use remote means to deliver course content and even to conduct testing. The US National Centre for Interpretation, based at the University of Arizona, has a combined approach with some remote training courses for interpretation, together with on-campus classes. Testing, however, is on-campus only (email correspondence, February 11, 2012).

Other major remote education providers of interpreting training are the providers of telephone interpreting services themselves. The main reason for this is these providers' desire to offer training and testing to trainees who are interested in becoming employees or contracted staff of such agencies. One of the largest telephone interpreting agencies that specialises on healthcare telephone interpreting, US-based Cyracom, offers both training and testing to trainees through remote means. Interpreter Education Online (IEO) offers training programs in general, legal and medical interpreting. It also offers testing in simultaneous and consecutive interpreting and sight translation, with all three offerings available as general tests, tests with legal terminology or tests with medical terminology.

The Medical Interpreting and Translating Institute Online offers three training programs: beginners, intermediate and advanced interpreting of approximately forty hours length in Spanish-English interpreting only. The courses consist of lecture notes sent in hard-copy, videos and a prescribed textbook. There appears to be no test that this educational provider requires other than successful completion of the training, after which trainees are recommended to US-based health providers that require the services of Spanish-English interpreters. The Berkeley Language Institute offers some online training but no testing. Pacific Interpreters offers mostly telephone interpreting services but also on-site interpreting and document translation.

The largest global telephone interpreting agency is US-based Language Line, which services not only North America but all parts of the world where English is one of the languages for which interpreting services are required. Language Line also provides on-site interpreting, document translation, training and testing.

4.2.7 Conclusions on technology and interpreting testing

- Telephone interpreting is now a commonplace and increasingly frequent means of communication for interpreted interactions.
- Telephone interpreting requires particular protocols, relies on different oral/aural input and output from the interpreter in the absence of visual information, requires extra checks from all participants in terms of confidentiality and has different interpersonal and stress management dynamics to manage both source speech or signing and technical features.
- The increased frequency of telephone interpreting justifies its inclusion as a component of NAATI testing at the proposed STAGE 3 level (Generalist Interpreting)
- Video-link (interpreter is located with one participant) or remote (interpreter is separated from both participants) interpreting is currently infrequently used. However, video-link and other remote forms of asynchronous communication are likely to become more frequent in the near future, as reported by most informants.
- Particular protocols relate to video-link and remote interpreting, similar to those for telephone interpreting.
- Digital recording of test performance allows greater ease of distribution and sharing to and between examiners.
- Both practitioner and examiner informants support audio and video recording for spoken language tests. A distinction between the two is important for some small language groups (not necessarily the 'new and emerging' ones) for evaluation of a test candidate's performance without the appearance of the candidate being made available to the examiner. A subsequent video recording is then to be made available to the examiner to evaluate other, non-linguist features of performance.
- Many survey informants support the idea of an electronic repository for recordings of candidates' tests to be stored and accessed from.
- Video-link technology, at present, is unable to secure a sufficiently high level of recording quality for test examination purposes.
- However, in the near future and with the roll-out of the National Broadband Network (NBN), video-link communication is likely to become a very common means of audio-visual communication in a large variety of situations between two or multiple interlocutors. T&I agencies are already investing in what is now costly infrastructure, but the capabilities of the NBN are likely to greatly accelerate the availability and popularity of this mode of communication for interpreted interactions and video-link interpreting (in addition to or as a substitute for telephone interpreting) should be considered as a component of the Generalist Interpreting test in the future.

5. General Conclusions and Recommendations

Below we provide a summary of the recommendations that we have made throughout the report, with implementation suggestions:

Recommendations on pre-requisites to accreditation

> 1. That all candidates complete compulsory education and training in order to be eligible to sit for the accreditation examinations, in accordance with the new suggested model outlined in section 2.3, Table 7.

For the languages for which formal NAATI approved courses are available, candidates should be advised to enrol in such courses as the preferred method of obtaining accreditation. Where candidates' languages are not offered as part of NAATI approved courses, candidates will be directed to follow the staged approach as outlined in the new proposed model. The new model recommends training modules in theory and practice that can be delivered mostly in English and through flexible modes. We acknowledge that there are currently limited opportunities for such training and that should this recommendation be accepted, such training modules will need to be made available before compulsory training is implemented. We therefore recommend that once such training becomes available, no accreditation be gained without having undertaken any Interpreting and/or Translation training.

Implementation suggestions: We propose that NAATI commence the process that leads up to compulsory training by first establishing the Expert Panel (see recommendation 16) to set up the training requirements and establish what constitutes equivalence. We acknowledge that full implementation of this recommendation can take a number of years, but the process can commence within the next year.

> 2. That NAATI produce an information package explaining the meaning of Interpreter and Translator, accreditation vs qualification, prerequisites for testing and expectations of potential candidates, including expected levels of language proficiency in English and the LOTE, as outlined in section 2 above.

Having read such an information package, it is envisaged that those who have misconceptions about Interpreting & Translation will decide not to pursue accreditation. This will ensure that candidates who do not have any chance of success will not waste money and time attempting accreditation. It will also minimise potential complaints about a low pass rate.

Implementation suggestions: We believe this is a recommendation that should be implemented without much delay. This is strongly supported by the results of our research and of previous studies.

> 3. That NAATI select (or devise) an on-line self-rating English proficiency test to be taken by potential candidates for a small fee, as part of the non-compulsory preparedness stage, as outlined in sections 2.3 and 3.1.

There is some controversy over the need to screen for language competence prior to accreditation. For example, Turner & Ozolins (2007) in their survey found no significant concerns over language levels. However, the results of our current study showed that language proficiency continues to be an issue – both for those sitting for the examinations and for those who practice in the field. Some certification bodies overseas also screen for language proficiency before allowing candidates to sit for the certification examination. For this reason, we recommend that language screening be voluntary rather than compulsory, during the preparedness stage. Candidates should be advised against attempting accreditation if they achieve a result lower than a set score (to be decided).

Implementation suggestions: We believe this recommendation is easy to implement (especially if a readily available English proficiency test is adopted), and could also be implemented without delay, in conjunction with recommendation 2.

> 4. That NAATI language panels select (or devise equivalent) on-line self-rating proficiency tests in the various languages to be taken by potential candidates for

a fee, as part of the non-compulsory preparedness stage, as outlined in sections 2.3 and 3.1.

As per recommendation 3, candidates should be advised against attempting accreditation if they achieve a result lower than a set score (to be decided).

Implementation suggestions: This recommendation will require more time than recommendation 3. We believe NAATI could task its language panels to devise its own set of LOTE proficiency tests and recover all costs through a fee to be paid by the candidates. The language proficiency test could also be used by candidates as a means to prove their proficiency for admission to formal Interpreting and Translation courses, also for a fee.

> 5. That an Advanced Diploma (or equivalent) be the minimum pre-requisite for the Generalist accreditation, and a Bachelor's degree (or equivalent, including a NAATI approved Advanced Diploma in Interpreting) be the minimum pre-requisite for Specialist accreditations, as outlined in section 2.

Implementation suggestions: Recommendation 5 can only be implemented in conjunction with recommendation 6.

Recommendations on accreditation

> 6. That the current levels of accreditation be replaced by a Generalist level (for both Interpreting and Translation) and Specialist accreditations for Interpreting, with a Provisional Generalist level with a sunset clause of 2 years, particularly for new and emerging and Aboriginal languages, as explained in section 2.

Implementation suggestions: This recommendation is connected to recommendation 11 for its final implementation. However, the current nomenclature could be changed now to reflect the adoption of this recommendation, with the Paraprofessional level becoming the Provisional Generalist level and the Professional level the Generalist level. We believe that such a nomenclature will more adequately reflect the contents and the purposes of the different accreditation levels.

> 7. That the following specialisations be established for Interpreter accreditations: Legal, Medical, Conference and Business (see Table 7), with Legal and Medical having priority over the other two, as explained in section 2.3.

This recommendation is connected to recommendation 6. The results of our research show strong support for the introduction of specialisations in Interpreting, but very little support for specialisations in Translation.

Implementation suggestions: As recommendation 6, it is also connected with recommendation 11 for its final implementation. However, NAATI already has a conference interpreting specialisation and could start to add the other specialisations to its suite of interpreting accreditations.

Recommendations on testing

In order to improve the authenticity and validity of the NAATI examinations, we recommend the following:

> 8. That NAATI move to computerised translator tests in the first place. Secondly, that test candidates undertaking computerised translator tests be allowed access to the internet while taking the test[41], taking account of security considerations. See section 3.5.2 and section 4.

Implementation suggestions: NAATI could first pilot computerised translator tests with no internet access while exploring security considerations for the use of the internet. The pilot phase could be implemented without much delay.

> 9. That Interpreting tests be conducted live, as much as possible. Where such is not possible, that candidates be provided with video recorded interactions and that their performance be video recorded for marking. See section 3.5.

Implementation suggestions: Some NAATI approved courses already conduct their final examinations live. NAATI could commence testing live for the major languages in the main capital cities where there is a sufficient supply of examiners and organise video recorded examinations for the other languages. This recommendation is connected to recommendations 12 & 13, as the examiners will need a revised assessment instrument to rate the candidates' live performances.

> 10. That Interpreting tests at the Generalist level for both spoken and signed languages include a telephone interpreting component consisting of protocols for identification of all interlocutors, confidentiality assurances and dialogue interpreting only. See section 3.5.1 and section 4.2.1.

Implementation suggestions: This recommendation is connected to recommendation 11, as the validation project will be used to determine the test components.

> 11. That a validation research project be conducted to design the new testing instruments for Interpreting and Translation. See section 3.6.

The validation study will provide empirically based construct definitions to design the components of the test, levels of difficulty of each component, standards, marking criteria and test delivery. Descriptors will need to be empirically defined so that assessment tools can be aligned with them.

Implementation suggestions: The validation study can have a duration of between 1 and 3 years (depending on its scope) and therefore recommendations 6-10 and 12-13 cannot be implemented until the completion of such a project. We suggest such a project be funded by an Australian Research Council Linkage grant, where NAATI becomes the linkage partner.

Recommendations on assessment

> 12. That new assessment methods using rubrics (see Table 8) be empirically tested as part of the validation project.

---

[41] This is being trialled by the American Translators' Association [ATA] and they have signalled their readiness to offer support and technical advice to NAATI working group members in regard to the introduction of logistic protocols and recently-developed software.

> 13. That new examiners' manuals be written to reflect the new assessment methods to be adopted.

Recommendations on examiners

> 14. That NAATI review the current composition of examiners' panels to include more graduates of approved courses and fewer practitioners who hold no formal qualifications in Interpreting and Translation. See section 3.7.

Implementation suggestions: We recommend that an open call for applications for examiners be implemented without delay. We further recommend that examiners serve for a period of 3 years which can be renewed for another three. In the case of languages of small diffusion, the term of office may need to be much longer. This recommendation can be implemented in time for the next round of call for applications.

> 15. That examiners undertake compulsory training before being accepted on the panel, and continuous training while on the panel[42]. See section 3.7 above.

Implementation suggestions: New training will be required if a new system is implemented. In the meantime, NAATI could offer one training session per year for all examiners.

Recommendations for specialist panels

> 16. That NAATI establish a new Expert Panel, with subpanels for the specialisations, to design the curricula for the compulsory training modules and provide guidelines for the final assessment tasks.

The Expert panel should comprise educators from the different NAATI approved courses, whose membership can rotate every five years. This recommendation is consistent with a number of recommendations in the Cook Report. Different Expert sub Panels should be organised for each specialisation, with representatives from the relevant industry/profession as well as from Interpreting (for example, lawyers for the legal specialisation, health care workers for the medical specialisation, etc).

Implementation suggestions: This recommendation could be implemented without delay.

Recommendations for approved courses

> 17. That NAATI continue to approve tertiary programs and encourage all applicants to take the formal path to accreditation where such is available for the relevant language combinations.

---

[42] For Aboriginal language examiners and possibly other languages of limited diffusion, training may be unrealistic in some languages due to literacy/numeracy considerations. In such cases we recommend that untrained examiners be partnered with a trained examiner, as explained in the report.

References

Amtsblatt der Europäischen Union. (2006). Europäisches Amt für Personalauswahl (EPSO) (2006/C 233 A/01). Retrieved 3 Nov, 2011, from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:233A:0003:0017:DE:PDF

Angelelli, C. V. (2007). Assessing Medical Interpreters: The Language and Interpreting Testing Project. *The Translator, 13*(1), 63-82.

Angelelli, C. V. (2009). Using a rubric to assess translation ability. *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue Between Research and Practice, 14*, 13.

Angelelli, C. V., & Jacobson, H. E. (2009). *Testing and Assessment in Translation and Interpreting Studies : A Call for Dialogue between Research and Practice*: John Benjamins Publishing Co.

Association of Translators Terminologists and Interpreters of Manitoba. (2011). Becoming a Member: Certification by Portfolio. from http://atim.mb.ca/en/becomingamember/certificationPortfolio.htm

ATA, [American Translators Association]. (2011). ATA Certification Program, Certification Exam. Retrieved 4 Nov, 2011, from http://www.atanet.org/certification/aboutexams_overview.php

ATA, [American Translators Association]. (n.d.). Flow chart for error point decisions. Retrieved 4 Nov, 2011, from http://www.atanet.org/certification/aboutexams_flowchart.pdf

ATIO, [Association of Translators and Interpreters of Ontario]. (2011). On Dossier Certification. Retrieved 1 Dec, 2011, from http://www.atio.on.ca/services/certification/php

Australian Department of Health and Ageing. (n.d.). A Manual of Mental Health Care in General Practice. Retrieved 16/12/2011, from http://www.health.gov.au/internet/publications/publishing.nsf/Content/mental-pubs-m-mangp-toc~mental-pubs-m-mangp-4~mental-pubs-m-mangp-4-int

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language testing, 17*(1), 1-42.

Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language testing, 19*(4), 453-476.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1): Oxford University Press, USA.

Baker, M. (1992). *In other words: A coursebook on translation*. London: Routledge.

Beltran Avery, M.-P. (2003). Creating a High-Standard, Inclusive and Authentic Certification Process. In L. Brunette, G. Bastik, L. Heinlik & H. Clarice (Eds.), *The Critical Link 3: Interpreters in the Community* (pp. 99-112). Amsterdam/Philadelphia: Benjamins.

Benmaman, V. (1997). Legal Interpreting by any other name is still Legal Interpreting. In S. E. Carr, R. Roberts, A. Dufour & D. Steyn (Eds.), *The Critical Link: Interpreters in the Community* (pp. 179-190). Amsterdam and Philadelphia: John Benjamins.

Berger, M.-J., & Simon, M. (1995). Programmation et évaluation en milieu scolaire: notes de cours. Unpublished manuscript. Université d'Ottawa.

Berk-Seligson, S. (1990 / 2002). *The Bilingual Courtroom. Court Interpreters in the Judicial Process*. Chicago: The University of Chicago Press.

Bontempo, K. (2009a). Conference interpreter performance evaluation rubric [private teaching and assessment resource].

Bontempo, K. (2009b). Interpreter performance evaluation rubric [private teaching and assessment resource].

Bontempo, K., & Hutchinson, B. (2011). Striving for an "A" Grade: A Case Study of Performance Management of Interpreters. *International Journal of Interpreter Education, 3*, 56-71.

Bontempo, K., & Napier, J. (2009). Getting it right from the start. *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue Between Research and Practice, 14*, 247.

Braun, S. (2011). Recommendations for the use of video-mediated interpreting in criminal proceedings. In S. Braun & J. Taylor (Eds.), *Videoconference and remote interpreting in criminal proceedings* (pp. 265-287). Guildford: University of Surrey.

Braun, S., & Taylor, J. (2011). Video-mediated interpreting: An overview of current practice and research. In S. Braun & J. Taylor (Eds.), *Videoconference and remote interpreting in criminal proceedings* (pp. 22-57). Guildford: University of Surrey.

Braun, S., Taylor, J. L., Miler-Cassino, J., Rybińska, Z., Balogh, T. K., Hertog, E., et al. (2011). Training in video-mediated interpreting in legal proceedings: Modules for interpreting students, legal interpreters and legal practitioners. In S. Braun & J. Taylor (Eds.), *Videoconference and remote interpreting in criminal proceedings* (pp. 205-254). Guildford: University of Surrey.

Calzada Perez, M. (2005). Applying Translation Theory in Teaching. *New Voices in Translation Studies, 1*, 1-11.

Cambridge, J. (1999). Information Loss in Bilingual Medical Interviews Through an Untrained Interpreter. *The Translator, 5*(2), 201-219.

Carroll, J. B. (1966). An Experiment in Evaluating the Quality of Translations. *Mechanical Translation and Computational Linguistics, 9*(3 & 4), 55-66.

Chacón, M. J. (2005). Estudio Comparativo de la Actuación de Intérpretes Profesionales y no Profesionales en Interpretación Social: Trabajo de Campo. *Puentes, 5*, 83-98.

Chen, J. (2009). Authenticity in accreditation tests for interpreters in China. *The Interpreter and Translator Trainer, 3*(2), 257-273.

Chen, N. S., & Ko, L. (2010). An online synchronous test for professional interpreters. *Educational Technology & Society, 13*(2), 153-165.

Chesher, T., Slatyer, H., Doubine, V., Jaric, L., & Lazzari, R. (2003). Community-based interpreting. The interpreters' perspective. In L. Brunette, G. Bastin, I. Hemlin & H. Clarke (Eds.), *The Critical Link 3* (pp. 273-291). Amsterdam and Philadelphia: John Benjamins.

Clifford, A. (2001). Discourse theory and performance-based assessment: Two tools for professional interpreting. *Meta: Journal des traducteurs, 46*(2), 365-378.

Clifford, A. (2003). A preliminary investigation into discursive models of interpreting as a means of enhancing construct validity in interpreter certification. Unpublished doctoral dissertation. University of Ottawa.

Clifford, A. (2005). Putting the exam to the test: Psychometric validation and interpreter certification. *Interpreting, 7*(1), 97-131.

Colin, J., & Morris, R. (1996). *Interpreters and the legal process*. Winchester: Waterside Press.

Cook, J., & Dixon, H. (2005). *A review of NAATI administrative processes related to testing including quality control processes.* Sydney: Macquarie University.

Cope, B., Kalantzis, M., Luke, A., McCormack, R., Morgan, B., Slade, D., et al. (1995). *Communication, Collaboration and Culture: The National Framework of Adult English Language, Literacy and Numeracy Competence*. Sydney: Centre for Workplace Communciation and Culture.

Council of Australasian Tribunals. (n.d.). Practice manual for tribunals: 5.4.2. Interpreters, *Chapter 5: Hearings*.

CRC, C. R. C. (2009). *Multicultural Services and Programs*. Retrieved from http://www.crc.nsw.gov.au/multicultural_policies_and_services_program_formally_eaps.

CTTIC [Canadian Translators, Terminologists and Interpreters Council],. (n.d.). *CTTIC Marker's Guide*.

Cuevas, A. (2011). *T&I Labour Market in Mexico*. Sydney: Centre for Translation and Interpreting Research, Macquarie University.

Delisle, J., Lee-Jahnke, H., Cormier, M. C., & Albrecht, J. (Eds.). (1999). *Terminologie de la traduction / Translation terminology / Terminología de la traducción / Terminologie der Übersetzung*. Amsterdam and Philadelphia:: John Benjamins.

DIAC, [Department of Immigration and Citizenship]. (2011). The People of Australia: Australia's Multicultural Policy. 1-16. Retrieved from http://www.immi.gov.au/media/publications/multicultural/pdf_doc/people-of-australia-multicultural-policy-booklet.pdf

Donovan, C. (2006, 30 June – 1 July 2006). *Trends–Where is Interpreting heading and how can training courses keep up?* Paper presented at the Presentation at EMCI conference in London, University of Westminster:"The Future of Conference Interpreting: Training, Technology and Research.

Ebden, P., Carey, O. J., Bhatt, A., & Harrison, B. (1988). The bilingual consultation. *Lancet, 1*(8581), 347.

Edwards, A. (1995). *The practice of court interpreting*. Amsterdam: John Benjamins.

Eyckmans, J., Anckaert, P., & Segers, W. (2009). The perks of norm-referenced translation evaluation. *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue Between Research and Practice, 14*, 73.

Family Court of Western Australia. (2006). Court funded interpreter services policy. July 2006. from http://www.familycourt.wa.gov.au/i/interpreter_services.aspx on 14/01/2010

Federal Association of Interpreters and Translators. (2011). BDU. from http://www.bdue.de/indexen.php

Federal Court. (n.d.). Federal Court Benchbook: 9.3 Interpreters (received from. E. Connolly in email correspondence on 04/02/2010).

Federal Magistrates Court. (n.d.). Interpreter and Translator Policy. from http://www.fmc.gov.au/services/html/interpreters.html on 16/12/2009

Föreningen Auktoriserade Translatorer. (2011). Advice to Customers. from http://www.aukttranslator.se/eng/advice.asp

Fowler, Y. (2007). *Interpreting into the ether: interpreting for prison/court video link hearings.* Paper presented at the Proceedings of the Critical Link 5 conference, Sydney, 11-15/04/2007.

Gentile, A., Ozolins, U., & Vasilakakos, M. (1996). *Liaison Interpreting*. Melbourne: Melbourne University Press.

Gile, D. (2004). Integrated problem and decision reporting as a translator training tool. *The Journal of Specialised Translation, 2*, 2-20.

Giovannini, M. (1993). *Report on the Development of English Language Assessment Tools for Use in Ministry Supported Cultural Interpreting Services for the Settlement and Integration Section*. Toronto: Ministry of Citizenship.

Gipps, C. (1994). *Beyond Testing:: Towards a Theory of Educational Assessment*. London: Falmer Press.

Hale, A., & Basides, H. (2012/13). *The keys to academic English*. South Yarra: Palgrave Macmillan.

Hale, S. (2004). *The discourse of court interpreting. Discourse practices of the law, the witness and the interpreter*. Amsterdam and Philadelphia: John Benjamins.

Hale, S. (2007a). The challenges of court interpreting: intricacies, responsibilities and ramifications. *Alternative Law Journal, 32*(4), 198-202

Hale, S. (2007b). *Community Interpreting*. Hampshire: Palgrave Macmillan.

Hale, S. (2011). *Interpreter policies, practices and protocols in Australian Courts and Tribunals. A national survey*. Melbourne: Australian Institute of Judicial Administration.

Hale, S., & Ozolins, U. (forthcoming in 2014). Monolingual short courses for language-specific accreditation: can they work? A Sydney experience. *The Interpreter and Translation Trainer, 8*(2), TBA.

Hamidi, M., & Pöchhacker, F. (2007). Simultaneous Consecutive Interpreting: A New Technique Put to the Test. *Meta: Journal des traducteurs, 52*(2), 276-289.

Hatim, B., & Mason, I. (1997). *The translator as communicator*. London & New York: Routledge.

Henning, G. (1987). *A guide to language testing: Development, evaluation, research*: Newbury House Cambridge, MA.

Hertog, E., & Reunbrouck, D. (1999). Building Bridges between Conference Interpreters and Liaison interpreters. In M. Erasmus (Ed.), *Liaison Interpreting in the Community* (pp. 263-277). Pretoria: Van Schaik.

Hornberger, J. (1998). Evaluating the costs of bridging language barriers in health care. *Journal of Health Care for the Poor and Underserved, 9*(5), S26-S39.

ILTA, [International Language Testing Association],. (2012). ILTA Guidelines for Practice. from http://iltaonline.com/images/pdfs/ILTA_Guidelines.pdf

IoL, [Institute of Linguists]. (2011). Diploma in Translation. Handbook for Candidates.   Retrieved 4 Nov, 2011, from http://www.iol.org.uk/qualifications/DipTrans/DipTransHandbook.pdf

IoLET, [Institute of Linguists Educational Trust]. (2004). Diploma in Public Service Interpreting.   Retrieved 8 Nov, 2011, from http://www.iol.org.uk/qualifications/IoL-DPSI-Handbook-Apr04.pdf

Jacobson, H. (2009). Moving beyond words in assessing mediated interaction: measuring interactional competence in healthcare settings. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and Assessment in Translation and Interpreting.* Amsterdam: John Benjamins Publishing Co.

Kalina, S. (2004). Quality in interpreting and its prerequisites. In G. Hansen, K. Malmkjær & D. Gile (Eds.), *Claims, changes and challenges in translation studies: selected contributions from the EST Congress, Copenhagen 2001 (Vol. 50). John Benjamins Publishing Company.* (pp. 121-130). Amsterdam/Philadelphia: John Benjamins.

Kelly, N. (2007). Interpreter Certification Programs in the US: Where Are We Headed? *The ATA Chronicle, January*, 31-39.

Kelly, N. (2008). *Telephone Interpreting: A comprehensive guide to the profession*. Clevedon: Multilingual Matters.

Kim, M. (2009). Meaning-oriented assessment of translations: SFL and its application for formative assessment. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and assessment in translation and interpreting studies*. Amsterdam/Philadelphia: John Benjamins

Ko, L. (2006). The need for long-term empirical studies in remote interpreting research: a case study of telephone interpreting. *Linguistica Antverpiensia New Series, 5*(2006), 325-338.

Ko, L. (2008). Teaching interpreting by distance mode: An empirical study. *Meta, 53*(4), 814-840.

Ko, L., & Chen, N. S. (2011). Online-interpreting in synchronous cyber classrooms. *Babel, 57*(2), 123-143.

Kuo, D., & Fagan, M. (1999). Satisfaction with methods of Spanish interpretation in an ambulatory care clinic. *Journal of general and internal medicine, 14*, 547-550.

Lee, J. (2005). Rating interpreter performance. Unpublished MA Dissertation. Macquarie University.

Lee, J. (2008). Rating scales for interpreting performance assessment. *The Interpreter and Translator Trainer, 2*(2), 165-184.

Lee, J. (2009). Toward more reliable assessment of interpreting performance. *The Critical Link, 5*(2009), 171-185.

Lee, L. J., Batal, H. A., Maselli, J. H., & Kutner, J. S. (2002). Effect of Spanish interpretation method on patient satisfaction in an urban walk-in clinic. *Journal of General Internal Medicine, 17*(8), 640-645.

Leman, P. (1997). Interpreter use in an inner city accident and emergency department. *Journal of accident & emergency medicine, 14*(2), 98-100.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Lysaght, R. M., & Altschuld, J. W. (2000). Beyond initial certification: the assessment and maintenance of competency in professions. *Evaluation and Program Planning, 23*(1), 95-104.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Mikkelson, H. (1996). Community Interpreting. An emerging profession. *Interpreting, 1*(1), 125-129.

Ministerio de Asuntos Exteriores y de Cooperacion. (2011). Traductores-Interpretes Jurados'. from http://www.maec.es/es/menuppal/ministerio/tablondeanuncios/interpretesjurados/Paginas/Intrpretes%20Jurados.aspx

Moore, T., & Morton, J. (2005). Dimensions of difference: a comparison of university writing and IELTS writing. *Journal of English for Academic Purposes, 4*(1), 43-66.

Mortensen, D. (2001). *Measuring Quality in Interpreting: A report on the Norwegian Interpreter Certification Examination (NICE)*: University of Oslo.

Moser-Mercer, B. (2003). Remote interpreting: assessment of human factors and performance parameters. *Communicate!* Retrieved from http://www.aiic.net/ViewPage.cfm/article879.htm

Moser-Mercer, B. (2005). Remote interpreting: Issues of multi-sensory integration in a multilingual task. *Meta, 50*(2), 727-738.

Mouzourakis, P. (2008). Remote Interpreter Training – Training for Remote Interpreting?   Retrieved 2 March, 2012, from http://multimedialinguas.wordpress.com/edicoes/ano-i-2010/0001-janeiro/panayotis-mouzourakis-%c2%abremote-interpreter-training-training-for-remote-interpreting%c2%bb/

NAATI. (2010-2011). *NAATI Annual Report 2010-2011*: National Accreditation Authority for Translators and Interpreters Ltd.

NAATI Test Review Translators Group. (2001). Final Summary.

Napier, J. (2004). Sign language interpreter training, testing, and accreditation: an international comparison. *American Annals of the Deaf, 149*(4), 350-359.

National Accreditation Authority for Translators and Interpreters. (2002). *NAATI Test Review (Accreditation, Test Formats and Test Methodology). Executive Summary of Responses to the Discussion Paper by Stakeholders*.

National Accreditation Authority for Translators and Interpreters. (2010). Improvements to NAATI Testing: Expressions of Interest Available from http://www.naati.com.au/pdf/misc/Improvements%20to%20NAATI%20Testing.pdf

National Board of Certification for Medical Interpreters. (2011). Written Exam. from http://www.certifiedmedicalinterpreters.org/written-exam

National Center for State Courts. (2011). Federal Court Interpreter Certification Examination For Spanish/English Examinee Handbook.   Retrieved 5 Nov, 2011, from http://www.ncsonline.org/d_research/fcice_exam/2011approvedbyAO-Online.pdf

National Centre for State Courts. (2009). Consortium for State Interpreter Certification. Survey: Certification Requirements 2009.   Retrieved 5 Nov, 2011, from http://www.ncsonline.org/D_Research/CourtInterp/Res_CtInte_ConsortCertRqmntssurvey2009.pdf

Nord, C. (1991). Scopos, loyalty, and translational conventions. *Target, 3*(1), 91-109.

NSW Community Justice Centre. (2009). 2009 Policy: Use of interpreters. from http://www.lawlink.nsw.gov.au/lawlink/Community_Justice_Centres/ll_cjc.nsf/pages/CJC_publications on 11/03/2010

NSW Workers Compensation Commission. (n.d.). Internal Policy: Provision of interpreter services (received by email from. S. Leatham on 08/12/2009).

NT Department of Housing Local Government and Regional Services. (2011). Working with interpreters. Retrieved 16/12/2011, from http://www.dlgh.nt.gov.au/interpreting/aboriginal_interpreter_service/working_with_interpreters

Organizacion Mexicana de Traductores, A. C. (2011). 'Certificaciones'. from http://www.omt.org.mx/certificaciones.htm

Orlando, M. (2010). Digital pen technology and consecutive interpreting: another dimension in notetaking training and assessment. *The interpreter's newsletter*(15), 71-86.

Ozolins, U. (1991). *Interpreting translating and language policy*. Melbourne: National Languages Institute of Australia.

Ozolins, U. (2004). *Survey of interpreting practitioners. Report*. Melbourne: VITS.

Ozolins, U. (2011). Telephone interpreting: understanding practice and identifying research needs. *The international journal of translation and interpreting research, 3*(2), 33-47.

PACTE Group. (2009). Results of the validation of the PACTE translation competence model.   Retrieved 9 November, 2012, from http://grupsderecerca.uab.cat/pacte/en/content/publications

Pochhacker, F. (2001). Quality assessment in conference and community interpreting. *Meta, 46*(2), 410-425.

QLD Department of Justice and Attorney-General. (2009). Language Services Policy. from http://www.justice.qld.gov.au/__data/assets/pdf_file/0005/33683/DJAG-Lang-Serv-Policy.pdf on 07/01/2010

QLD Health Interpreter Service. (2007). Working with Interpreters: Guidelines.   Retrieved 16/12/2011, from http://www.health.qld.gov.au/multicultural/interpreters/guidelines_int.pdf

Roat, C. (2006). *Certification of Health Care Professionals in the United States. A Primer, a Status Report, and Considerations for National Certification*: The California Endowment.

Roberts, R. P. (2000). Interpreter assessment tools for different settings. *BENJAMINS TRANSLATION LIBRARY, 31*, 103-120.

Rosenberg, B. A. (2007). A data driven analysis of telephone interpreting. In C. Wadensjö, B. Englund Dimitrova & A. L. Nilsson (Eds.), *The critical link 4. Professionalisation of interpreting in the community.* (pp. 65-76). Amsterdam: Benjamins.

Roy, C. (2000). *Interpreting as a discourse process*. New York and Oxford: Oxford University Press.

Russell, D., & Malcolm, K. (2009). Assessing ASL-English interpreters. The Canadian model of national certification. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and Assessment in Translation and Interpreting Studies* (pp. 371-376): John Benjamins.

SA Health. (2006). Language Services Provision: Operational Guidelines For Health Units.   Retrieved 16/11/2011, from http://www.sahealth.sa.gov.au/wps/wcm/connect/e41bb280455fe974a8b8fa8a21f01153/2011+Language+Services+Provision+-+Operational+Guidelines+for+Health+Unitsand+coversheet.pdf?MOD=AJPERES&CACHEID=e41bb280455fe974a8b8fa8a21f01153&CACHE=NONE

Sandrelli, A. (2001). Teaching Liaison Interpreting. Combining tradition and innovation. In I. Mason (Ed.), *Triadic Exchanges. Studies in Dialogue Interpreting* (pp. 173-196). Manchester: St. Jerome.

Schjoldager, A. (1995). An exploratory study of translational norms in simultaneous interpreting: methodological reflections. *Hermes, 14*, 65-87.

Skaaden, H. (1999). Lexical knowledge and interpreter aptitude. *International Journal of Applied Linguistics, 9*(1), 77-97.

Skaaden, H., & Wattne, M. (2009). Teaching interpreting in cyberspace. The answer to all our prayers? In R. de Pedro Ricoy, I. Prerez & C. Wilson (Eds.), *Interpreting and translating in public service settings* (pp. 74-88). Manchester: St. Jerome.

Slatyer, H., Elder, C., Hargreaves, M., & Luo, K. (2008). *An investigation into rater reliability, rater behaviour and comparability of test tasks*. Sydney: Access Macquarie, Macquarie University.

South African Translators' Institute, S. (2007). Translation Accreditation.   Retrieved 9 Dec, 2011, from http://translators.org.za/sati_cms/index.php?frontend_action=display_text_content&content_ind=1761

Stansfield, C. W., Scott, M. L., & Kenyon, D. (1992). The measurement of translation ability. *The Modern Language Journal, 76*(4), 455-467.

Stejskal, J. (2001). International Certification Study: Accreditation Program in Brazil. *The ATA Chronicle, July*.

Stejskal, J. (2002a). International Certification Study: Argentina. *The ATA Chronicle, June*, 13-17.

Stejskal, J. (2002b). International Certification Study: Finland and Sweden. *The ATA Chronicle, February*, 14-15.

Stejskal, J. (2002c). International Certification Study: Japan. *The ATA Chronicle, September*, 17-19.

Stejskal, J. (2002d). International Certification Study: Norway. *The ATA Chronicle, July*, 13-22.

Stejskal, J. (2002e). International Certification Study: Spain and Portugal. *The ATA Chronicle, October*, 20-30.

Stejskal, J. (2002f). International Certification Study: U.K. and Ireland. *The ATA Chronicle, May*, 12-23.

Stejskal, J. (2002g). International Certification Study: Ukraine. *The ATA Chronicle, November/December*, 12-18.

Stejskal, J. (2005). Survey of the FIT Committee for Information on the Status of the Translation & Interpretation Profession. Retrieved 28 November 2011, from FIT [Online]: http://www.fit-europe.org/vault/admission/FITsurvey2005.pdf

STIBC [Society of Translators and Interpreters, British Columbia],. (2008). Court Interpreting - CTTIC Certification Examination Information for Candidates.   Retrieved 8 Dec, 2011, from http://www.stibc.org/page/court%20interpreter%20-%20criteria%20and%20procedures%20for%20certification%20by%20exam.aspx

Supreme Court of Queensland. (2005). Equal Treatment Benchbook. from http://www.courts.qld.gov.au/The_Equal_Treatment_Bench_Book/S-ETBB.pdf on 18/12/2009

Timarova, S., & Ungoed-Thomas, H. (2008). Admission testing for interpreting courses. *The Interpreter and Translator Trainer, 2*(1), 29-46.

Tiselius, E. (2010). Revisiting Carroll's scales. In C. V. Angelelli & H. Jacobson (Eds.), *Testing and Assessment in Translation and Interpreting Studies : A Call for Dialogue between Research and Practice* (Vol. 14, pp. 95-121). Amsterdam/Philadelphia: John Benjamins Publishing Co. .

Turner, B., Lai, M., & Huang, N. (2010). Error deduction and descriptors–A comparison of two methods of translation test assessment. *Translation & Interpreting, 2*(1), 11-23.

Turner, B., & Ozolins, U. (2007). The standards of linguistic competence in English and LOTE among NAATI accredited interpreters and translators: RMIT University, Melbourne.(Available at www. mit. edu/gsssp/ti.).

University of Oslo. (2001). Measuring Quality in Interpreting: A report on the Norwegian Interpreter Certification Examination (NICE).   Retrieved 12 Dec, 2011, from http://folk.uio.no/dianem/IntQuality-Internet.pdf

Veisbergs, P. (2007). Terminology issues in interpeter training. *Proceedings of the Baltic Sea Region University Network: Quality and Qualifications in Translation and Interpreting II* Retrieved 1 Mar, 2012, from http://www.tlu.ee/files/arts/645/Quali698bb7e395e0b88d73d603e33f5b153f.pdf

Vermeiren, H., Van Gucht, J., & De Bontridder, L. (2009). Standards as critical success factors in assessment. *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue Between Research and Practice, 14*, 297-330.

WA Office of Multicultural Interests. (2008). The Western Australian Language Services Policy 2008. from http://multicultural.wa.gov.au/OMI_language.cfm on 14/01/2010

Waddington, C. (2001). Different methods of evaluating student translations: The question of validity. *Meta: Journal des traducteurs, 46*(2).

Waddington, C. (2004). Should student translations be assessed holistically or through error analysis? *Lebende Sprachen, 49*(1), 28-35.

Wadensjö, C. (1998). *Interpreting as interaction*. London & New York: Longman.

Wakabayashi, J. (1996). *The Translator's voice*. Paper presented at the Jill Blewitt Memorial Lecture. Retrieved from http://ausit.org/national/?page_id=253

Weir, C. J. (2005). *Language testing and validation*: Palgrave Macmillan Houndmills, Basingstoke.

Wilss, W. (1982). *The Science of Translation: Problems and Methods*. Tubingen: Gunter Narr.

Winteringham, S. (2010). The usefulness of ICTs in interpreting practice. *The Interpreter's Newsletter*(15), 87-99.

Appendices

Appendix 1   NAATI Project Specialist Working Group Memberships

1.   Working group on rubrics, descriptors and competency-based assessment

Name	Role	Country
Barry Turner	Co-chair	Australia
Miranda Lai	Co-chair	Australia
Gyde Hansen	Consultant (T&I Educator)	Denmark
David Deck	Research Assistant	Australia
Claudia Angelelli	Consultant (T&I Educator, Researcher)	USA
Dave Gilbert	Participant (Translator)	Australia

2.   Working group on technology and interpreting testing

Name	Role	Country
Jim Hlavac	Chair	Australia
Marc Orlando	Advisor (Educator, Interpreter, Translator)	Australia
Elizabeth Bryer	Research Assistant	Australia

3.   Working group on technology and translation testing

Name	Role	Country
Ignacio Garcia	Chair	Australia
Anthony Pym	Participant (T&I Educator, Researcher)	Spain
Vivian J. Stevenson	Participant (T&I Practitioner)	Australia
Miguel A. Jimenez-Crespo	Participant (T&I Educator, Researcher)	USA
Alan K. Melby	Participant (T&I Educator, Researcher)	USA
Sharon O'Brien	Participant (T&I Educator, Researcher)	Ireland
Anne Lafeber	Participant (T&I Practitioner, Researcher)	Spain

4.   Working group on test validity and reliability

Name	Role	Country
Helen Slatyer	Chair	Australia
Catherine Elder	Consultant (Testing & Assessment Researcher)	Australia
Jim Hlavac	Researcher (T&I Educator)	Australia
Robert Foote	Participant (NAATI)	Australia

5.   Working group on pre-requisites and specialisations

Name	Role	Country
Sandra Hale	Chair	Australia
Louise Hadfield	Research Assistant	Australia
Mira Kim	Researcher	Australia
Jemina Napier	Advisor (Sign Language Educator, Researcher)	Australia
Marc Orlando	Advisor (Educator, Interpreter, Translator)	Australia
Michael Cooke	Advisor (Aboriginal Interpreting)	Australia
Magdalena Rowan	Participant (TAFE Educator)	Australia
Claudia McQuillan	Participant (Translator, Interpreter, Educator)	Australia
Tony Foley	Participant (Lawyer/Academic)	Australia
Anne Kenny	Participant (Health Care Interpreters Service)	Australia

Appendix 2   Survey distribution lists

Interpreter and Translator Agencies

- ABC International Translating and Interpreting Services
- 2M Language Services (Queensland, NSW, South Australia)
- Aboriginal Interpreter Service, Department of Housing, Local Government and Regional Services, NT
- Aboriginal Interpreting Services, NT
- Absolute translations, Perth, WA
- ACT Interpreter Service (AUSLAN)
- Allgraduates (National)
- AllWorldLanguages
- AmigosTranslate
- Anecsys, (VIC, NSW)
- Associated Translators & Linguists Pty Ltd, NSW
- Australian Multi Lingual Services AMLS
- CENTRELINK
- Chin Communications
- Chris Poole Translation
- Connect Language Services (VIC, NSW, ACT)
- CRC Community Relations Commission, NSW
- Deaf Society of NSW, AUSLAN Interpreting
- International Interpreting Agency, VIC
- Interpreters On Site, NSW
- Interpreting & Translating Service, NT
- Kimberley Interpreting Services, NT
- Language Connect, NSW
- Language Professionals, Langpro, NSW
- NSW Multicultural Health Communication Service
- Oncall Interpreters, (National)
- Polaron Language Services, VIC
- SBS Language Translation Services, NSW
- SWAHS Health Care Interpreter Service
- TIS National
- Translation Consultants International
- WA Deaf - Interpreting Booking Service
- Translation House, WA
- Verbatim Language Services, NSW
- VICDEAF, AUSLAN Interpreting
- VITS Language Link, VIC

Examiners and Educators

All NAATI examiners on each language panel and Educators from:

- Abbey College, NSW
- Macquarie University, NSW
- Monash University
- RMIT, VIC
- Sydney Institute of Interpreting and Translation, NSW
- TAFE NSW
- TAFE SA
- University of Queensland, QLD
- UNSW (NSW)
- UWS (NSW)

Practitioners

UNSW Distribution list (comprising hundreds of I&T contacts throughout Australia)

AUSIT e-bulletin

Interpreter and Translator panels of the following agencies to distribute to their panels of interpreters and translators:

- 2M Language Services (Queensland, NSW, South Australia)
- Aboriginal Interpreter Service, Department of Housing, Local Government and Regional Services, NT
- Aboriginal Interpreting Services, NT
- Absolute translations, Perth, WA
- Allgraduates (National)
- Anecsys, (VIC, NSW)
- Associated Translators & Linguists Pty Ltd, NSW
- Australian Multi Lingual Services AMLS
- CENTRELINK
- Connect Language Services (VIC, NSW, ACT)
- Community Relations Commission, CRC, NSW
- International Interpreting Agency, VIC
- Interpreters On Site, NSW
- Interpreting & Translating Service, NT
- Kimberley Interpreting Services, NT
- Language Connect, NSW
- NSW Multicultural Health Communication Service
- Oncall Interpreters, (National)
- Polaron Language Services, VIC
- SBS Language Translation Services, NSW
- SWAHS Health Care Interpreter Service
- TIS National
- Translation House, WA
- Translators International Pty Ltd, WA
- Verbatim Language Services, NSW
- VICDEAF, AUSLAN Interpreting, VIC
- VITS Language Link, VIC
- WAITI

Appendix 3   Questionnaire administered to Translation and Interpreting Agencies

1. Are you:

> Please check all that apply and/or add your own variant.
> a private agency
> a government agency
> Other

2. In what state or territory do you reside?

> Please pick one of the answers below.
> ACT
> NSW
> NT
> QLD
> SA
> TAS
> VIC
> WA

3. How crucial is the level of NAATI accreditation in allocating work to practitioners?

> Please mark the corresponding circle - only one per line.
> 1. Not at all 2. 3. 4. 5. Very much

4. Does your agency record how NAATI accreditation was obtained by each practitioner (by training/testing)?

> Please pick one of the answers below.
> Yes
> No

5. Do you give preference to practitioners with formal training in addition to NAATI accreditation?

> Please mark the corresponding circle - only one per line.
> Never Rarely Sometimes Usually Always

6. Do you pay higher fees to practitioners with higher NAATI accreditation level?

> Please pick one of the answers below or add your own.
> Yes
> No
> Comment

7. Do you pay higher fees to practitioners with Interpreting/Translating tertiary qualifications plus NAATI accreditation?

> Please pick one of the answers below or add your own.
> Yes
> No
> Comment

8. How often do you receive feedback from users about interpreter performance?

> Please pick one of the answers below and add your comments.
> very often
> not at all
> Comments

9. List the top five comments you receive as feedback from clients:

Please use the blank space to write your answers.

1.
....
2.
....
3.
....
4.
....
5.
....

10. Please indicate your level of agreement with the following statements:

Please mark the corresponding circle - only one per line.
1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree

Interpreters should complete compulsory training before being accredited.
Translators should complete compulsory training before being accredited.
A minimum amount of experience should be mandatory before
Translators can be accredited.
A minimum amount of experience should be mandatory before Interpreters can be accredited.
NAATI accreditation should not be necessary if an I&T education program was completed.
NAATI should continue approving training programs that lead to accreditation.
There should be mandatory specialised training for all legal, medical and conference interpreters
There should be mandatory training for all NAATI examiners.
Translators and interpreters should undertake continuous professional development
Employing agencies should pay for interpreter and translator professional development

11. Do you have any other suggestions for the review of NAATI testing and related issues?

Please write your answer in the space below.

....
....
....
....

Appendix 4   Questionnaire administered to Examiners and Educators

1. Are you:

> Please check all that apply and/or add your own variant.
> a NAATI panel interpreting/translation examiner
> a TAFE intepreting/translation educator
> a university interpreting/translation educator
> Other

2. What is your language combination? Please specify:

> Please write your answer in the space below.
> ....
> ....
> ....
> ....

3. How long have you been on the examiners' panel?

> Please pick one of the answers below.
> Not applicable
> 1 year or less
> 5 years or less
> 10 years or less
> more than 10 years

4. How long have you taught I&T?

> Please pick one of the answers below.
> Not applicable
> 1 year or less
> 5 years or less
> 10 years or less
> more than 10 years

5. In what state or territory do you reside?

> Please pick one of the answers below.
> ACT
> NSW
> NT
> QLD
> SA
> TAS
> VIC
> WA

6. What NAATI accreditation do you hold in TRANSLATION?

> Please pick one of the answers below.
> None
> Recognition
> Paraprofessional
> Professional
> Advanced
> Advanced (Senior)

7. How did you obtain your TRANSLATION accreditation?

Please pick one of the answers below or add your own.
I don't have NAATI accreditation or recognition in Translation
By recognition
By sitting a NAATI test
By completing a NAATI approved VET course in Australia
By completing a NAATI approved University course in Australia
In recognition of a University course overseas
Other

8. If you obtained your TRANSLATOR accreditation by course completion, how long was the course?

Please pick one of the answers below and add your comments.
Six months equivalent full time
Twelve months equivalent full time
Eighteen months equivalent full time
Three years equivalent full time
Other
What was the name of the course?

9. What NAATI accreditation do you hold in INTERPRETING?

Please pick one of the answers below.
None
Paraprofessional
Professional
Conference
Conference (Senior)

10. How did you obtain your NAATI accreditation in INTERPRETING?

Please pick one of the answers below or add your own.
I don't have NAATI accreditation in Interpreting
By recognition
By sitting a NAATI test
By completing a NAATI approved VET course in Australia
By completing a NAATI approved University course in Australia
In recognition of a University course overseas
Other

11. If you obtained your INTERPRETER accreditation by course completion, how long was the course?

Please pick one of the answers below.
Six months equivalent full time
Twelve months equivalent full time
Eighteen months equivalent full time
Three years equivalent full time
Other

12. What other formal qualifications do you have? Please specify qualifications and country where they were awarded:

Please write your answer in the space below.
....
....
....
....

13. Do you have any formal qualifications in assessment and evaluation? If YES, please specify which ones:

    Please pick one of the answers below and add your comments.
    Yes
    No

14. Are you a practising interpreter and/or translator? If YES, go to questions 15 and/or 16 , if NO, to question 17.

    Please pick one of the answers below.
    Yes
    No

15. How often do you practise as interpreter?

    Please mark the corresponding circle - only one per line.
    More than once a week
    At least once a week
    More than once a month
    At least once a month
    Less than once a month
    Never

16. As translator

    Please mark the corresponding circle - only one per line.
    more than 1000 words a week
    more than 500 words a week
    less than 1000 words a month
    less than 500 words a month

17. Have you practised interpreting/translation in the past?

    Please pick one of the answers below and add your comments.
    Yes
    No
    Comments

18. Please indicate your level of agreement with the following statements:

    Please mark the corresponding circle - only one per line.
    1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree

    Interpreters should complete compulsory training before being accredited.
    Translators should complete compulsory training before being accredited.
    A minimum amount of experience should be mandatory before Translators can be accredited.
    A minimum amount of experience should be mandatory before Interpreters can be accredited.
    NAATI accreditation should not be necessary if an I&T education program was completed.
    NAATI should continue approving training programs that lead to accreditation.
    There should be mandatory specialised training for all legal, medical and conference interpreters.
    There should be mandatory training for all NAATI examiners.
    There should be different types of accreditation according to training and specialisation
    Translators and interpreters should undertake continuous professional development
    New interpreters should be mentored by an experienced interpreter for an initial period of time
    New translators should be mentored by an experienced translator for a number of assignments
    The current NAATI translator exam adequately assesses the skills and knowledge needed to practice as a professional translator
    The current NAATI intepreter exam adequately assesses the skills and knowledge needed to practice as a professional interpreter in all settings
    The current NAATI Paraprofessional interpreter exam should continue to exist
    The current NAATI marking guidelines for Translation are adequate
    The current NAATI marking guidelines for Intepreting are adequate
    The current NAATI pass rate is adequate to ensure quality of intepreting and/or translation

19. List the top five skills you think a translator test should test

Please write your answer in the space below.

....
....
....
....

20. List the top five skills you think an Interpreter test should test

Please write your answer in the space below.

....
....
....
....

21. Do you have any other suggestions for the review of NAATI testing and related issues?

Please write your answer in the space below.

....
....
....
....

Appendix 5   Questionnaire administered to Practitioners

1. Are you:

> Please check all that apply and/or add your own variant.
> an interpreter
> a translator
> both an interpreter and translator
> Other

2. In what state or territory do you reside?

> Please pick one of the answers below.
> ACT
> NSW
> NT
> QLD
> SA
> TAS
> VIC
> WA

3. What NAATI accreditation do you hold in TRANSLATION?

> Please pick one of the answers below.
> None
> Recognition
> Paraprofessional
> Professional
> Advanced
> Advanced (Senior)

4. How did you obtain your TRANSLATION accreditation?

> Please pick one of the answers below or add your own.
> I don't have NAATI accreditation or recognition in Translation
> By recognition
> By sitting a NAATI test
> By completing a NAATI approved VET course in Australia
> By completing a NAATI approved University course in Australia
> In recognition of a University course overseas
> Other

5. If you obtained your TRANSLATOR accreditation by course completion, how long was the course?

> Please pick one of the answers below and add your comments.
> Six months equivalent full time
> Twelve months equivalent full time
> Eighteen months equivalent full time
> Three years equivalent full time
> Other
> What was the name of the course?

6. What NAATI accreditation do you hold in INTERPRETING?

> Please pick one of the answers below.
> None
> Paraprofessional
> Professional
> Conference
> Conference (Senior)

7. How did you obtain your NAATI accreditation in INTERPRETING?

> Please pick one of the answers below or add your own.
> I don't have NAATI accreditation in Interpreting
> By recognition
> By sitting a NAATI test
> By completing a NAATI approved VET course in Australia
> By completing a NAATI approved University course in Australia
> In recognition of a University course overseas
> Other

8. If you obtained INTERPRETING accreditation by course completion, how long was the course?

> Please pick one of the answers below and add your comments.
> Six months full time equivalent?
> Twelve months full time equivalent?
> Eighteen months full time equivalent?
> Three years full time equivalent?
> What was the name of the course?

9. Have you ever failed a NAATI test?

> Please pick one of the answers below and add your comments.
> Yes
> No
> If yes, do you know why?

10. When did you gain your accreditation in TRANSLATION?

> Please pick one of the answers below.
> Between 1977 and 1980
> Between 1981 and 1990
> Between 1991 and 2000
> Between 2001 and 2011
> Not applicable

11. When did you gain your accreditation in INTERPRETING?

> Please pick one of the answers below.
> Between 1977 and 1980
> Between 1981 and 1990
> Between 1991 and 2000
> Between 2001 and 2011
> Not applicable

12. How long have you been practising as a TRANSLATOR?

> Please pick one of the answers below and add your comments.
> For over 20 years
> Between 10 and 20 years
> Between 5 and 10 years
> Less than 5 years
> I don't practise as a translator
> If you don't practice, comment why

13. How long have you been practising as an INTERPRETER?

> Please pick one of the answers below and add your comments.
> Over 20 years
> Between 10 and 20 years
> Between 5 and 10 years
> Less than 5 years
> I don't practise as an interpreter
> If you don't practise, comment why

14. If you gained your accreditation by sitting a test, please indicate your level of agreement with the following statements:

> Please mark the corresponding circle - only one per line.
> 1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree
>
> I was well prepared to practise as translator after passing the test
> I was well prepared to practise as an interpreter after passing the test
> I was well prepared to translate simple texts after passing the test
> I was well prepared to translate complex texts after passing the test
> I was well prepared to interpret for simple interactions after passing the test
> I was well prepared to interpret in complex settings such as the courtroom, after passing the test

15. If you gained your accreditation by completing a formal course of study, indicate your level of agreement with the following statements:

> Please mark the corresponding circle - only one per line.
> 1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree
>
> I was well prepared to practise as a translator after completing the course
> I was well prepared to practise as an interpreter after completing the course
> I was well prepared to translate simple texts after completing the course
> I was well prepared to translate complex texts after completing the course
> I was well prepared to interpret at complex settings such as the courtroom, after completing the course
> I was well prepared to interpret for simple interactions after completing the course

16. Please indicate your level of agreement with the following statements:

> Please mark the corresponding circle - only one per line.
> 1. Strongly Disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly Agree
>
> Interpreters should complete compulsory training before being accredited.
> Translators should complete compulsory training before being accredited.
> A minimum amount of experience should be mandatory before Translators can be accredited.
> A minimum amount of experience should be mandatory before Interpreters can be accredited.
> NAATI accreditation should not be necessary if an I&T education program was completed.
> NAATI should continue approving training programs that lead to accreditation.
> There should be mandatory specialised training for all legal interpreters
> There should be mandatory specialised training for all medical interpreters
> There should be mandatory specialised training for all conference interpreters
> There should be mandatory training for all NAATI examiners.
> There should be different types of accreditation according to training and specialisation
> Translators and interpreters should undertake continuous professional development
> New interpreters should be mentored by an experienced interpreter for an initial period of time
> New translators should be mentored by an experienced translator for a number of assignments

17. Would you be willing to mentor new interpreters and/or translators?

> Please pick one of the answers below and add your comments.
> Yes
> No
> Comments

18. If you are new in the profession, would you be willing to be mentored by experienced interpreters and/or translators?

> Please pick one of the answers below and add your comments.
> Yes
> No
> Comments

19. List the top five skills you think a translator test should assess

Please write your answer in the space below.

………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………

20. List the top five skills you think an Interpreter test should assess

Please write your answer in the space below.

………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………

21. Do you have any other suggestions for the review of NAATI testing and related issues?

Please write your answer in the space below.

………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………
………………………………………… ………………………………………… …………………………………………

Appendix 6   CEFR B2 Descriptors for the macro-skills of Speaking and Listening (Association of Language Teachers of Europe, n.d.)

Level Three may be referred to as an intermediate stage of proficiency. Users at this level are expected to be able to handle the main structures of the language with some confidence, demonstrate knowledge of a wide range of vocabulary and use appropriate communicative strategies in a variety of social situations. Their understanding of spoken language and written texts should go beyond being able to pick out items of factual information, and they should be able to distinguish between main and subsidiary points and between the general topic of a text and specific detail. They should be able to produce written texts of various types, showing the ability to develop an argument as well as describe or recount events. This level of ability allows the user a certain degree of independence when called upon to use the language in a variety of contexts. At this level the user has developed a greater flexibility and an ability to deal with the unexpected and to rely less on fixed patterns of language and short utterances. There is also a developing awareness of register and the conventions of politeness and degrees of formality as they are expressed through language.

Examinations at Level B2 are frequently used as proof that the learner can do office work or take a non-academic course of study in the language being learned, e.g. in the country where the language is spoken. Learners at this level can be assumed to have sufficient expertise in the language for it to be of use in clerical, secretarial and managerial posts, and in some industries, in particular tourism.

Productive Skills

Speaking
In social and travel contexts, users at this level can deal with most situations that may arise in shops, restaurants, and hotels; for example, they can ask for a refund or for faulty goods to be replaced, and express pleasure or displeasure at the service given. Similarly, routine situations at the doctors, in a bank or post office or at an airport or station can all be handled. In social conversation they can talk about a range of topics and express opinions to a limited extent. As tourists they can ask for further explanations about information given on a guided tour. They themselves can show visitors around, describe a place and answer questions about it.

In the workplace, users at this level can give detailed information and state detailed requirements within a familiar topic area, and can take some limited part in a meeting. They can take and pass on messages, although there may be difficulties if these are complex, and can carry out simple negotiations, for example on prices and conditions of delivery.

If studying, users at this level can ask questions during a lecture or presentation on a familiar or predictable topic, although this may be done with some difficulty. They can also give a short, simple presentation on a familiar topic. They can take part in a seminar or tutorial, again with some difficulty.

Receptive Skills

Listening
In social and travel contexts, users at this level can cope with casual conversation on a fairly wide range of familiar, predictable topics, such as personal experiences, work and current events. They can understand routine medical advice. They can understand most of a TV programme because of the visual support provided, and grasp the main points of a radio programme. On a guided tour they have the understanding required in order to ask and answer questions.

In the workplace, they can follow presentations or demonstrations of a factual nature if they relate to a visible, physical object such as a product.

If studying, they can understand the general meaning of a lecture, as long as the topic is predictable.

Appendix 7   The National Framework of Adult English Language, Literacy and Numeracy Competence

Excerpts of the first-mentioned capabilities for each group only (Cope, et al., 1995, pp. 18-19).

Stage 3: Collaborative Competence

Task:	Can use diverse cultural/gender and linguistic/mathematic experiences and resources to establish alternative approaches to tasks.
Technology:	Can co-ordinate and synthesise discussions on ways of utilising the range of technologies in relation to diverse cultural/gender and linguistic/mathematic resources.
Identity:	Can analyse and assess the degree of responsibility taken by individuals and groups in relation to the organisation.
Group:	Can collaboratively use diverse resources (experience and knowledge) or the group to formulate team/group goals and objectives.
Organisation:	Can review organisational goals in terms of cultural/gender/linguistic/mathematic resources of the organisation.
Community:	Can collaboratively initiate and manage group discussions which relate community functions and objectives to the needs and the expectations of diverse social groups.

Appendix 8   Samples of Jacobson's Rubrics

Rubrics for competence in the use and transfer of contextualisation cues (paralinguistic features of speech) (Jacobson, 2009, p. 63)

Contextualisation Cues	
Superior	Demonstrates superior ability in understanding meaning of contextualisation cues (voice volume, intonation, prosody, and paralinguistic features) accompanying the utterances of primary interlocutors; produces effective and natural renditions of such cues in the target language; demonstrated balanced focus on both accuracy of information and interactional features; produces consistently dynamic renditions with appropriate intonation contour in the target language.
Advanced	Demonstrates advanced ability in understanding meaning of contextualisation cues (voice volume, intonation, prosody, and paralinguistic features) accompanying the utterances of primary interlocutors; is usually able to interpret such cues into target language, with some difficulty at times due to inability to consistently focus on both accuracy of information and interactional features; renditions are usually dynamic and appropriate, with occasional lapses into monotone renditions.
Fair	Consistently demonstrates difficulty in understanding meaning or contextualisation cues (voice volume, intonation, prosody, and paralinguistic features) accompanying utterances of primary interlocutors; is often unable to focus on both accuracy of information and interactional features; renditions tend to be monotone and dull, characterised by frequent backtracking.
Poor	Demonstrates clear inability to understand meaning of contextualisation cues (voice, intonation, prosody, and paralinguistic features) in the utterances of primary interlocutors; is unable to interpret such cues into target language due to inability to focus on both accuracy of information and interactional features; renditions are consistently monotone, characterised by excessive back-tracking and stuttering.

Rubrics for competence in the use and transfer of discourse management (Jacobson, 2009, p. 65)

Discourse management	
Superior	Provides a clear, concise pre-session to primary interlocutors on interpreter's role when possible; consistently uses the first person while interpreting, switching to third person for clarifications; encourages interaction, including eye contact, between interlocutors, both verbally and through other paralinguistic cues; allows interlocutors to complete turns due to strong memory and note-taking skills; demonstrates strategies for dealing with overlap.
Advanced	Provides a clear, concise pre-session to primary interlocutors on interpreter's role when possible; consistently uses the first person while interpreting, switching to third person for clarifications; usually encourages interaction between interlocutors, both verbally and through other paralinguistic cues; usually demonstrates skill in allowing interlocutors to complete turns without interrupting for clarifications, with some difficulty due to need to further develop memory and note-taking skills; generally deals calmly and effectively with overlaps, with demonstrated need for further practice.
Fair	In most cases, provides a clear, concise pre-session to primary interlocutors on interpreter's role, although at least one or two of the principal points are usually left out; is inconsistent in using the first person while interpreting, and exhibits excessive use of the third person, leading to awkward renditions; does not often encourage interaction between interlocutors, either verbally or through other paralinguistic cues; often interrupts interlocutors mid-turn for clarifications due to need to develop memory and note-taking skills and to build vocabulary; becomes nervous when challenged by overlaps, demonstrating clear need for further practice.
Poor	Does not always provide a clear, concise pre-session to primary interlocutors on interpreter's role, leaving out principal points; is inconsistent in using the first person while interpreting and almost always uses the third person, leading to awkward renditions; does not encourage interaction between interlocutors, either verbally or through other paralinguistic cues; does not allow interlocutors to complete turns, and interrupts frequently to request clarification, resulting in choppy discourse; note-taking and memory skills are poor; does not deal effectively with overlaps, leading to interruptions in the dialogue and excessive omissions.

Appendix 9   Interpreter Performance Evaluation Rubric (Bontempo, 2009b)

Interpreter's Name: _____ Setting / context: _____

ELEMENTS TO CONSIDER	COMMENTS
1. Interpreting Aspect	
1.1 Equivalence of message (appropriate for context? Contains textual integrity and fidelity? Info exchange is successful overall?)	
1.2 Avoids distracting mannerisms that impact on performance (whispering, vocalisations, upper body shifts, inappropriate eye gaze etc)	
1.3 Uses appropriate time lag to allow concepts to be conveyed accurately	
1.4 Miscues (omissions, additions, substitutions, intrusions, anomalies) – any strategic?	
TRANSMISSION ACCURACY: 1 2 3 4 5 5 = VERY ACCURATE	
2. Language Aspect	
2.1 Comprehends source message (English vocabulary, denotative/connotative meaning, Auslan signs/fingerspelling/numbers.)	
2.2 Paralinguistic elements (facial expression, pace, size of signing space, mouth movements, etc. English prosody / inflection)	
2.3 Articulation (clear production of signs, fingerspelling, numbers, etc in Auslan. Clear production in English at correct volume.)	
2.4 Uses correct grammar & structure in target message (complete thoughts in English & Auslan; use of space, classifiers, tenses, indexing, etc in Auslan)	
2.5 Fluency ('smoothness', control and flow of language; comprehensibility/ease of viewing or listening to target text – care taken not to overly smooth out rough source text)	
2.6 Vocabulary and register (using correct signs, right style of language, appropriate vocabulary, idioms, strategies for unknown / key vocab etc)	
OVERALL LANGUAGE SKILLS: 1 2 3 4 5 5 = EXCELLENT	
3. Interaction / Role aspects	
3.1 Roles specific to setting (contextual adjustments; preparation/prior knowledge; checking comprehension; challenges of setting)	
3.2 Managing overlap, turn-taking (& indicates speakers), questions, interruptions, clarifications & introductions	
3.3 Handling ethical dilemmas & demonstrating ethical behavior.	
3.4 Social / cultural / professional sensitivity (use of appropriate strategies to gain attention; facilitation of social interactions; interpreter interaction with parties present)	
MANAGING INTERACTION / ROLE : 1 2 3 4 5 5 = EXCELLENT	
4. Professional Conduct	
4.1 Environmental management (to extent possible), appropriate positioning, accessibility in general	
4.2 Appearance / presentation, demeanor, punctuality, posture etc.	
4.3 Response to errors / overall confidence	
4.4 General attitude, conduct and body language	
OVERALL PROFESSIONALISM: 1 2 3 4 5 5 = EXCELLENT	
5. General Comments	

Appendix 10 Description of Angelelli's four remaining Rubrics

Draft rubrics for translation evaluation: style and cohesion (addressing textual sub-component); situational appropriateness (addressing pragmatic sub-component); grammar and mechanics (addressing micro-linguistic sub-component); and translation skill (strategic sub-component) (Angelelli, 2009, pp. 40-41)

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 forma 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 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 skilful 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 translation's intended purpose in the TL. Cultural references, discourse, and register are mostly appropriate for the TL domain but some phrasing or work 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 achive the translation's intended purpose in the TL. Cultural references, discourse and register are at times inappropriate for the TL domain. Numerous phrasings 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 translation's 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. Skilful 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.

Appendix 11 'Rubrics' Questionnaire

At present, NAATI translating & interpreting candidates are awarded pass/fail for accreditation based on a 'subtractive' system that awards an overall numerical score by deducting varying numbers of marks according to the type and seriousness of errors detected.

An alternative or complementary marking system uses 'rubrics': candidate performance is assessed over a range of areas (but still on the basis of errors detected), with testees being allocated a level in each area on the basis of descriptions of performance at each level.

The purposes of this focus group are:

- to explain to you, with some examples, what a 'rubrics-based' marking system might look like if NAATI decided to adopt one; and
- to allow you to try using a set of rubrics to briefly assess some translations, then seek your reactions and impressions.

However, please note that the example rubrics used in this focus group (from Prof Claudia Angelelli's work) do **not necessarily represent** the system that NAATI **might** eventually adopt. Both the decision to adopt a rubrics-based system, and exactly what system to adopt, are questions that still require much further evaluation (of which this is a part).

We would appreciate it if you could respond to the following questionnaire.

1. Examiners / Educators

a. Compared with the present marking system(s), do you feel that the use of the rubrics-based system provided you with **clearer** guidance on assessing? **Yes / Unsure / No**

Comments: _____

_____

_____

b. Compared with the present marking system(s), do you feel that a rubrics-based system would be **easier** to use / apply?   **Yes / Unsure / No**

Comments: _____

_____

_____

c. Compared with the present marking system(s), do you feel that the rubrics-based system encouraged you to take a **wider range of factors** into account when marking?
Yes / Unsure / No

Comments: _____

_____

_____

d. The example set of rubrics describes five levels for each assessment area. On the basis of this trial of using these rubrics, what level do you think should be the 'adequate' / 'passing' level?
5 4 3 2 1

Comments: _____

_____

_____

e. Some rubrics-based assessment systems also determine 'hurdle' levels; that is, in any or all assessment areas, if candidates are awarded a **certain level or lower**, no matter how well they have performed in other assessment areas, a pass is automatically **precluded**. On the basis of this trial of using these rubrics, if such a 'hurdle' level were to be applied, what level do you feel should automatically preclude a pass?

5 4 3 2 1

Comments: _____

_____

_____

f.   The example set of rubrics makes assessments in five areas. Are there any areas of assessment that you would suggest to be added or removed?

Comments: _____

_____

_____

g.   Do you have any other general comments about the idea of NAATI looking into using a system of assessment that was partly or wholly rubrics-based?

_____

_____

_____

Thank you for your responses and your time.

Appendix 12 Marking System for Tests in IoL Diplomas

The Diploma of Public Service Interpreting (UK), Assessment Criterion Statements as reproduced in Turner & Ozolins (2007: 136-137)

NATIONAL QUALIFICATIONS FRAMEWORK LEVEL 4

DPSI ASSESSMENT CRITERION STATEMENTS FOR TASK 1-INTERPRETING

	Accuracy	Delivery	Language Use
Band A	**Mark Range 10-12**	**Mark Range 10-12**	**Mark Range 10-12**
	The candidate: -conveys sense of original message with complete accuracy -transfers all information without omissions, additions, distortions -demonstrates complete competence in conveying verbal content and familiarity with subject matter	The candidate: -demonstrates complete competence in languages -switches effortlessly between languages -interprets clearly and smoothly -reflects tone, emotion and non-verbal signs appropriate to situation -displays a courteous and confident manner -remains unobtrusive and impartial -handles intercultural references correctly -displays good management strategies intervening appropriately and only when necessary to clarify or ask for repetition or prevent breakdown of communication	The candidate: -demonstrates excellent command of grammar, syntax, vocabulary, specialist terminology, with minimum paraphrasing -chooses language and register entirely appropriate to situation -has clear, distinct pronunciation -has an accent which in no way affects ease of comprehension
Band B	**Mark Range 7-9**	**Mark Range 7-9**	**Mark Range 7-9**
	The candidate: -accurately conveys sense of original message -makes only one or two minor omissions/distortions not affecting correct transfer of information or complete comprehension	The candidate: -demonstrates good competence in language -switches easily between languages -interprets for most part clearly and smoothly -reflects tone, emotion and non-verbal signals of interlocutors -displays a courteous and confident manner -remains unobtrusive and impartial -handles intercultural references correctly -intervenes justifiably and appropriately -makes occasional slip or sign of nervousness but not leading to communication problem	The candidate: -demonstrates good command of grammar, syntax, vocabulary, and specialist terminology -paraphrases in clear, concise way, where appropriate -chooses language, register largely appropriate to situation -has clear, distinct pronunciation -has an accent which in no way or only occasionally affects ease of comprehension
Band C	**Mark Range 4-6**	**Mark Range 4-6**	**Mark Range 4-6**
	The candidate: -adequately conveys sense of original message -makes no serious inaccuracies, omissions or distortions affecting comprehension or transfer of information	The candidate: -demonstrates adequate competence in language -switches between languages without major problem -shows some confidence while interpreting -makes reasonable attempt to reflect suitable tone, emotion and demeanour -displays a manner, delivery and interventions, occasionally not completely appropriate, but not leading to irretrievable breakdown of communication	The candidate: -demonstrates adequate command of grammar, syntax, vocabulary and specialist terminology -keeps paraphrasing to acceptable level -may choose inappropriate language/register at times but not impairing overall transfer of information -may occasionally display faulty pronunciation or a pronounced accent but without impairing message
Band D	**Mark Range 1-3**	**Mark Range 1-3**	**Mark Range 1-3**
	The candidate: -does not, or only partially, convey sense of original message -makes serious inaccuracies, omissions, distortions affecting comprehension and transfer of information -demonstrates inadequate grasp of language and/or subject matter	The candidate: -demonstrates inadequate competence in language -has problems switching between languages -lacks confidence and clarity -does not attempt to reflect tone, emotion relevant to situation -sounds flat and mechanical or too loud and overbearing -fails to apply suitable management strategies, where appropriate, e.g. asking for repetition/clarification -makes excessive requests for repetition/clarification	The candidate: -has inadequate command of grammar, syntax, vocabulary and specialist terminology -uses excessive and inaccurate paraphrasing which distorts meaning -uses register which prevents successful transfer of message -has a strong accent, intonation or stress patterns, making it difficult to understand meaning of message

DPSI ASSESSMENT CRITERION STATEMENTS FOR TASK 3 – WRITTEN TRANSLATION

NATIONAL QUALIFICATIONS FRAMEWORK LEVEL 4

	ACCURACY/APPROPRIACY OF TRANSLATED TEXT	COHESION, COHERENCE & GENRE CONVENTIONS	EFFECTIVENESS OF COMMUNICATION
Band A	**Mark range 10-12** The candidate: • conveys information with complete accuracy • conveys all information without omissions, additions or distortions • shows excellent use of vocabulary throughout • uses excellent grammatical/syntactical constructions • displays faultless spelling, accentuation/use of diacritics, faultless punctuation • has excellent knowledge of specialist terminology with minimum paraphrasing	**Mark Range 10-12** The candidate: • provides text which conforms fully to the conventions of the relevant genre • uses cohesive devices which are completely appropriate • provides text which facilitates a completely coherent reading	**Mark Range 10-12** The candidate: • provides text which enables reader to act entirely appropriately on the basis of the translation • displays excellent presentation and legibility • displays excellent use of style and register • shows excellent awareness of intercultural differences and handles these consistently well
Band B	**Mark Range 7-9** The candidate: • conveys information accurately with only few minor errors • transfers information without serious omissions or distortions • shows good use of vocabulary • uses sound grammatical/syntactical constructions • displays good spelling, careful accentuation/use of diacritics/punctuation • has good knowledge of specialist terminology, paraphrasing in clear, concise way, where necessary	**Mark Range 7-9** The candidate: • provides text which generally conforms well to the conventions of the relevant genre • makes generally good use of cohesive devices • provides text which facilitates generally a good coherent reading	**Mark Range 7-9** The candidate: • provides text which enables reader to act appropriately on the basis of the translation • displays good presentation and legibility • displays good style and register with only minor errors • shows good awareness of intercultural differences and handles these well
Band C	**Mark Range 4-6** The candidate: • makes some inaccuracies but no misunderstandings adequately conveys content • makes minor omissions/distortions but not leading to misunderstandings • makes adequate use of vocabulary • uses adequate grammatical/syntactical constructions • displays no major faults in spelling, accentuation, diacritics, punctuation • has adequate knowledge of specialist terminology with paraphrasing kept to an acceptable level	**Mark Range 4-6** The candidate: • provides text which conforms adequately to the conventions of the relevant genre • makes adequate use of cohesive devices • provides text which adequately facilitates a coherent reading	**Mark Range 4-6** The candidate: • provides text which adequately enables reader to act appropriately on the basis of the translation • displays adequate presentation and legibility • displays adequate style and register • shows adequate awareness of intercultural differences, these in an acceptable manner
Band D	**Mark Range 1-3** The candidate: • makes inaccuracies leading to misunderstandings inadequately conveys content • makes serious omissions/distortions • has inadequate command of vocabulary • has inadequate knowledge of grammar and syntax • displays considerable number of technical faults: spelling, punctuation, accentuation, diacritics • has inadequate knowledge of specialist terminology and/or excessive and inaccurate paraphrasing	**Mark Range 1-3** The candidate: • provides text which does not conform to the conventions of the relevant genre • uses inadequate cohesive devices • provides text which does not adequately facilitate a coherent reading	**Mark Range 1-3** The candidate: • provides text unfit for the purpose of enabling reader to act on the basis of the translation • displays poor presentation and legibility • displays inappropriate style/register • shows lack of awareness of intercultural differences and handles these badly

Appendix 13   ATA (2011b) Framework for Standardized Error Marking Explanation of Error Categories

ATA Certification Exam – Type and Frequency of Errors

Addition: (A): An **addition** error occurs when the translator introduces superfluous information or stylistic effects. Candidates should generally resist the tendency to insert "clarifying" material.
Explicitation is permissible. Explicitation is defined as "A translation procedure where the translator introduces precise semantic details into the target text for clarification or due to constraints imposed by the target language that were not expressed in the source text, but which are available from contextual knowledge or the situation described in the source text." (Translation Terminology, p. 139)

Ambiguity: (AMB): An **ambiguity** error occurs when either the source or target text segment allows for more than one semantic interpretation, where its counterpart in the other language does not.

Capitalization: (C): A **capitalization** error occurs when the conventions of the target language concerning upper and lower case usage are not followed.

Cohesion: (COH): A **cohesion** error occurs when a text is hard to follow because of inconsistent use of terminology, misuse of pronouns, inappropriate conjunctions, or other structural errors. Cohesion is the network of lexical, grammatical, and other relations which provide formal links between various parts of a text. These links assist the reader in navigating within the text. Although cohesion is a feature of the text as a whole, graders will mark an error for the individual element that disrupts the cohesion.

Diacritical marks / Accents: (D): A **diacritical marks** error occurs when the target-language conventions of accents and diacritical marks are not followed. If incorrect or missing diacritical marks obscure meaning (sense), the error is more serious.

Faithfulness: (F): A **faithfulness** error occurs when the target text does not respect the meaning of the source text as much as possible. Candidates are asked to translate the meaning and intent of the source text, not to rewrite it or improve upon it. The grader will carefully compare the translation to the source text. If a "creative" rendition changes the meaning, an error will be marked. If recasting a sentence or paragraph—i.e., altering the order of its major elements—destroys the flow, changes the emphasis, or obscures the author's intent, an error may be marked.

Faux ami: (FA): A ***faux ami*** error occurs when words of similar form but dissimilar meaning across the language pair are confused. *Faux amis*, also known as false friends, are words in two or more languages that probably are derived from similar roots and that have very similar or identical forms, but that have different meanings, at least in some contexts.

Grammar: (G): A **grammar** error occurs when a sentence in the translation violates the grammatical rules of the target language. Grammar errors include lack of agreement between subject and verb, incorrect verb tenses or verb forms, and incorrect declension of nouns, pronouns, or adjectives.

Illegibility: (ILL): An **illegibility** error occurs when graders cannot read what the candidate has written. It is the candidate's responsibility to ensure that the graders can clearly discern what is written. Candidates are instructed to use pen or dark pencil and to write firmly enough to produce legible photocopies. Deletions, insertions, and revisions are acceptable if they do not make the intent unclear.

Indecision: (IND): An **indecision** error occurs when the candidate gives more than one option for a given translation unit. Graders will not choose the right word for the candidate. Even if both options are correct, an error will be marked. More points will be deducted if one or both options are incorrect.

Literalness: (L): A **literalness** error occurs when a translation that follows the source text word for word results in awkward, unidiomatic, or incorrect renditions.

Mistranslation: (MT): A **mistranslation** error occurs when the meaning of a segment of the original text is not conveyed properly in the target language. "Mistranslation" includes the more specific error categories described in separate entries. Mistranslations can also involve choice of prepositions, use of definite and indefinite articles, and choice of verb tense and mood.

Misunderstanding: (MU): A **misunderstanding** error occurs when the grader can see that the error arises from misreading a word, for example, or misinterpreting the syntax of a sentence.

Omission: (O): An **omission** error occurs when an element of information in the source text is left out of the target text. This covers not only textual information but also the author's intention (irony, outrage). Missing titles, headings, or sentences within a passage may be marked as one or more errors of omission, depending on how much is omitted.
Implicitation is permissible. Implicitation is defined as "A translation procedure intended to increase the

economy of the target text and achieved by not explicitly rendering elements of information from the source text in the target text when they are evident from the context or the described situation and can be readily inferred by the speakers of the target language." (Translation Terminology, p. 145)

Punctuation: (P): A **punctuation** error occurs when the conventions of the target language regarding punctuation are not followed, including those governing the use of quotation marks, commas, semicolons, and colons. Incorrect or unclear paragraphing is also counted as a punctuation error.

Register: (R): A **register** error occurs when the language level or degree of formality produced in the target text is not appropriate for the target audience or medium specified in the Translation Instructions. Examples of register errors include using everyday words instead of medical terms in a text intended for a medical journal, translating a text intended to run as a newspaper editorial in legalese, using the familiar rather than the polite form of address, and using anachronistic or culturally inappropriate expressions. Register is defined as "A property of discourse that takes into account the nature of relationships among speakers, their socio-cultural level, the subjects treated and the degree of formality and familiarity selected for a given utterance or text." (Translation Terminology, p. 172)

Spelling: (SP) (**Character** (CH) for non-alphabetic languages): A **spelling/character** error occurs when a word or character in the translation is spelled/used incorrectly according to target-language conventions. A spelling/character error that causes confusion about the intended meaning is more serious and may be classified as a different type of error using the Flowchart and Framework. If a word has alternate acceptable spellings, the candidate should be consistent throughout the passage.

Style: (ST): A **style** error occurs when the style of the translation is inappropriate for publication or professional use as specified by the Translation Instructions. For example, the style of an instructional text should correspond to the style typical of instructions in the target culture and language, or the temper of a persuasive essay may need to be toned down or amplified in order to achieve the desired effect in the target language.

Syntax: (SYN): A **syntax** error occurs when the arrangement of words or other elements of a sentence does not conform to the syntactic rules of the target language. Errors in this category include improper modification, lack of parallelism, and unnatural word order. If incorrect syntax changes or obscures the meaning, the error is more serious and may be classified as a different type of error using the Flowchart and Framework.

Terminology: (T): A **terminology** error occurs when a term specific to a special subject field is not used when the corresponding term is used in the source text. This type of error often involves terms used in various technical contexts. This also applies to legal and financial contexts where words often have very specific meanings. In more general texts, a terminology error can occur when the candidate has not selected the most appropriate word among several that have similar (but not identical) meanings.

Unfinished: (UNF): A substantially **unfinished** passage is not graded. Missing titles, headings, or sentences within a passage may be marked as one or more errors of omission, depending on how much is omitted.

Usage: (U): A **usage** error occurs when conventions of wording in the target language are not followed. Correct and idiomatic usage of the target language is expected.

Word form / **Part of speech:** (WF / PS): A **word form** error occurs when the root of the word is correct, but the form of the word is incorrect or nonexistent in the target language (e.g. "conspiration" instead of "conspiracy"). A **part of speech** error occurs when the grammatical form (adjective, adverb, verb, etc.) is incorrect (e.g. "conspire" instead of "conspiracy").

ATA (2011b) Certification Programme – Rubric for Grading

Exam number:

Version 2011 Exam passage: **Evaluation by Dimensions**

Instructions: In each column, the grader marks the box that best reflects performance in that dimension, measured against the ideal performance defined for that dimension in the "Standard" row. The grader may also insert, circle, and/or cross out words in a description to make the evaluation more specific.

Note: A passage may show uneven performance across the dimensions. For example, a candidate with excellent command of the target language but limited knowledge of the source language might show *Strong* performance for *Target mechanics* but *Minimal* performance for *Usefulness / transfer*.

See also the Explanation on the reverse.

STANDARD	Usefulness / transfer	Terminology / style	Idiomatic writing	Target mechanics
	The translated text is fully usable for the purpose specified in the Translation Instructions. The meaning and sense of the source text have been fully and appropriately transferred to the translated text.	Terminology is appropriate in context. Style and register are appropriate for the topic in the target language and for the specified audience.	Translated text reads smoothly. Wording is idiomatic and appropriate for the topic in the target language and for the specified audience.	Translated text fully follows the rules and conventions of target language mechanics (spelling, grammar, punctuation, etc.).
Strong	Translated text transfers meaning in a manner fully consistent with the Translation Instructions. Translation contains few or no transfer errors, and those present have a minor effect on meaning.	Translated text contains few or no inappropriate term or style/register choices. Any errors have a minor effect on meaning.	Translated text is almost entirely idiomatic and appropriate in context. Any errors have a minor effect on meaning.	Translated text contains few or no errors in target language mechanics.
Acceptable	Translated text transfers meaning in a manner sufficiently consistent with the Translation Instructions. Translation contains occasional and/or minor transfer errors that slightly obscure or change meaning.	Translated text contains occasional and/or minor inappropriate term or style/register choices. Such errors may slightly obscure meaning.	Translated text contains occasional unidiomatic or inappropriate wording. Such errors may slightly obscure meaning.	Translated text contains occasional errors in target language mechanics.
Deficient	Translated text transfers meaning in a manner somewhat consistent with the Translation Instructions. Translation contains more than occasional transfer errors that obscure or change meaning.	Translated text contains frequent inappropriate and/or incorrect terms or style/register choices. Such errors may obscure or change meaning.	Translated text contains frequent and/or obvious unidiomatic or inappropriate wording. Such errors may obscure or change meaning.	Translated text contains frequent and/or obvious errors in target language mechanics.
Minimal	Translated text transfers meaning in a manner inconsistent with the Translation Instructions. Translation contains frequent and/or serious transfer errors	Translated text contains excessive inappropriate and/or incorrect terms or style/register choices. Such errors obscure or	Translated text contains excessive and/or disruptive unidiomatic or inappropriate wording. Such errors obscure or change meaning.	Translated text contains excessive and/or disruptive errors in target language mechanics.

that obscure or
change meaning.

change meaning.

Appendix 14 The Federal Court Interpreter Certification Examination (FCICE) (USA)

Source: Federal Court Interpreter Certification Examination For Spanish/English Examinee Handbook (2011)

The FCICE is a two-phase process, involving a Spanish-English Written Examination (Phase One) and an Oral Examination (Phase Two) administered on a biennial basis with Phase One and Phase Two occurring in alternating years. Interpreters must pass the Written Examination with a score of 75 percent or higher in order to be eligible to sit for the Oral Examination.

Written Examination

The Phase One Written Examination serves primarily as a screening test for linguistic competence in English and Spanish and is a prerequisite for the Phase Two Oral Examination. The Written Examination is a four-option, multiple choice examination of job-relevant language ability in English and Spanish. In 2008, and possibly in the future, there will be 100 items in the English section and 100 items in the Spanish section of the test. When that happens, additional time will be provided for the candidates to take this longer, 200 item test. Each section consist of five parts, and each part involves a task that is considered to be relevant for a court interpreter. The Written Examination tests comprehension of written text, knowledge of vocabulary, idioms, and grammatically correct expression, and the ability to select an appropriate target language rendering of source language text.

The English and Spanish sections of the examination are scored separately and the criterion to pass is 75 percent correct answers on <u>each</u> section of the test.

Oral Examination

The Phase Two Oral Examination directly measures interpreting skills. Because it fulfils the legal mandate for a "criterion-referenced performance examination" the Oral Examination is the basis for certification to interpret in the federal courts. The Oral Exam assesses the ability of the interpreter to adequately perform the kinds of interpretation discourse that reflects both form and content pertinent to authentic interpreter functions encountered in the federal courts. It consists of five parts: Interpreting in the consecutive mode; interpreting a monologue in the simultaneous mode; interpreting a witness examination in the simultaneous mode; sight translation of a document from English into Spanish; and sight translation of a document from Spanish into English. All five parts are simulations of what interpreters do in court.

The language used in the examinations varies widely across speech registers and vocabulary range. Test items include both formal and informal/colloquial language, technical and legal terminology, and other specialized language that is part of the active vocabulary of a highly articulate speaker, both in English and in Spanish. Overall, there are 220 scored items in the test and the examinee must render 80 percent of them correctly to pass the test. In addition, the examinee's performance is scored holistically on three skills, including interpreting skills, English skills, and Spanish skills.

Practice Oral Examination

Overview and instructions for the Oral Examination Practice Test

This section contains all of the Oral Practice Examination material relating to the Oral Examination for the Federal Court Interpreter Certification Examination (FCICE). These instructions are for the web-based Practice Examination. You will need to follow the online instructions and use your computer to play .mp3 files. If you ordered a hard copy of the Handbook, these materials are included as Part 8 of the book and a CD is included, containing all of the audio files.

 These materials contain everything you need to self-administer the Practice Test in a way that closely simulates the actual test experience. When you are finished administering the practice examination, you can then score your examination. You can also listen to and score an example of a strong passing performance.

The practice test materials are presented in the same sequence that they are given during the examination itself, as follows:

- Part 1-English to Spanish Sight Translation
- Part 1-Spanish to English Sight Translation
- Part 2 Simultaneous Monologue
- Part 3 Consecutive
- Part 4 Simultaneous Witness Examination

Appendix 15   Marker's Guide for the CTTIC (Canadian Translators, Terminologists and Interpreters Council) Translation Test (CTTIC [Canadian Translators, n.d., pp. 3-5)

Marking Scale

Errors are considered to fall into **two main categories**: **Translation** (Comprehension – failure to render the meaning of the original text) and **Language** (Expression – violation of grammatical and other rules of usage in the target language).

Major and minor errors must be identified within each category.

Note: Errors in the text must be highlighted (underlined, etc.).

Translation (Comprehension)

Major mistakes--shown in margin as **(T)**   -10
e.g. serious misinterpretation denoting
a definite lack of comprehension of the source language,
nonsense, omission of a phrase or more

Minor mistakes--shown in margin as **T**   -5
e.g. mistranslation of a single word,
omission/addition affecting meaning,
lack of precision, wrong shade of meaning.

Language (Expression)

Major mistakes--shown in margin as **(L)**   -10
e.g. gibberish, unacceptable structure

Minor mistakes--shown in margin as **L**   -5
e.g. syntax, grammar, ambiguity, unnecessary repetition,
convoluted structure, non-idiomatic structure,
unacceptable loan translation

Minor mistakes--shown in margin by **I**   -3
e.g. breach of spelling, punctuation,
typographical conventions

Application

If an error recurs consistently throughout the text (punctuation, pronouns, etc.), it should only be penalized the first time it appears, except in cases where it affects the meaning.

Errors must be indicated in the margin of the paper **using the appropriate letter**. When a paper has been corrected, the various types of errors must be entered at the end of each text, together with the total points deducted-- e.g.

    (T) 1 x 10 =10
    T 1 x 5 = 5
    (L) 1 x 10 =10
    L 1 x 5 = 5
    I 1 x 3 = 3
    -33

In this example, the candidate obtains 67%

[(100 - 33) = 67]

N.B. Please consider this example for a moment. As can be seen, the candidate is only three points short of the pass mark. If we were to take the combined mark given by the two markers for the two texts, we would then be faced with the worst-case scenario. So as to avoid CTTIC having to deal with complaints from unhappy candidates (which can be a costly and time-consuming process), we must try to distance ourselves as far as possible from the 70% pass mark.

In a case such as this one, the two markers must try to confirm the failure or success of the candidate, leaning in so far as possible towards success. Would it not be possible here to overlook the three points taken off for misuse of punctuation?

Let's now presume that the two translations, despite a major error of transfer and one of language, are generally well done, that the style used makes for pleasant reading. In such a case, we would try to slightly offset the two or three major mistakes by giving a positive overall mark (maximum of 10 points) in order to recognize the quality of each of the translations.

Experience shows us, however, that such a case seldom occurs and that the style of borderline candidates usually leaves something to be desired. In such a situation, read the translations again and see if you have been too generous towards the candidate. Might you not have failed to note one or two spelling or punctuation mistakes, which would in fact push the mark further down in the 60% range.

Appendix 16 Community interpreting services of Ottawa-Carleton test (Roberts, 2000)

The marking system requires two markers, each one having as their L1 one of the languages of the pair. The dialogue interpreting component of the test has five areas of assessment which are awarded up to 2 marks each:

General vocabulary	2
Technical terms	2
Grammar	2
Appropriate register, level of language and tone	2
Pronunciation and audibility	2

Gradings:

- very good (native fluency) = 2;
- good for a non-native = 1.5;
- understandable but numerous errors = 1;
- definitely too poor to be an interpreter = 0

The testee is given a mark out of 10 for each respective language. These are added together to give a mark out of 20. A marking key is provided for examiners. For the consecutive interpreting exercise, 'global' criteria are presented with recommended mark allocations:

1. How well is the source text understood? Total of 6 points. For each mistranslation – deduct 1 point.
2. How accurately does the candidate present the ideas in the target language (excluding names and numbers)? Total of 6 points. For each omission, addition or distortion of ideas, deduct 1 point.
3. How well does the candidate handle names and numbers? Total of 2 points.
4. Does the candidate use appropriate target language grammar and syntax? Total of 6 points based on overall impression (3 points for English, 3 points for the foreign language).

Pass mark: 14/ 20.

Appendix 17 Court Interpreting - CTTIC Certification Examination

Information for candidates - Society of Interpreters and Translators or British Columbia (STIBC [Society of Translators and Interpreters, 2008)

Examination format
The examination is a two-step process consisting of a written and an oral component. Candidates who pass the written will be eligible to sit the oral at a later date. Candidates register and pay for the two examinations separately.

(i)   *Written examination*

a)   Written translation on a general legal or court related subject from a language of specialization into English (or French) AND
Written translation of legal terms & technical vocabulary commonly used in court proceedings from English (or French) into the language of specialization. (Note: no dictionaries will be allowed.)

b)   Legal knowledge test (multiple choice and short answers)

c)   Ethics component (case scenario followed by questions)

(ii)   *Oral examination*

a)   Sight translation from English (or French) to language of specialization and from language of specialization into English (or French).

b)   Simultaneous interpretation from English (or French) into language of specialization.

c)   Consecutive interpretation from English (or French) into language of specialization and from language of specialization into English (or French).

Marking
Examinations will be marked independently by two markers. Candidates must pass all three parts of the written component with a minimum of 70% on each. Similarly, all three parts of the oral component must be passed with a minimum of 70% on each.

How are the exams marked, and what do the comments mean?
(Answers provided by Creighton Douglas, Chair, CTTIC Board of Certification; Oct. 1998)

Let me assure candidates that every paper is carefully read and corrected by two markers, who must agree on the final mark. If they do not agree, the paper is referred to a third marker, whose decision is final.

The pass mark is 70% and any paper that falls between 65% and 70% is reviewed very carefully to ensure that a pass or failure is clearly justified.

Re "General Comments": they will usually seem repetitive, since the pattern of errors from candidate to candidate and from year to year is very similar. These comments in no way *disqualify a candidate*, but are written after the paper is marked to indicate the nature of the problems in a general way.

The term "transfer error" refers to a shift in meaning, sometimes quite subtle, between the original meaning in the source text and the meaning as translated into the target language. Such errors can be very important, but at the same time difficult for the candidate to recognize – if the candidate had perceived the error, they likely would not have made it!

Appendix 18 Association of Visual Language Interpreters of Canada test of interpretation. American Sign Language Criteria (linguistic criteria) AVLIC Evaluation Committee (2007, as cited in Russell & Malcolm, 2009, p. 370)

I	DISCOURSE STRATEGIES
1.	Standard: Overall discourse strategies used result in coherent text.
a.	Appropriate use of opening/closing comments
b.	Essential elements of meaning with adequate supporting detail
c.	Appropriate use of topic transition and topic maintenance strategies - Exhibits strategies for comparing and contrasting ideas - References within the text to previously introduced information
d.	Avoids restatement of ideas that do not add meaning to the text.
II	FORM
1.	Standard: Overall sign production is clear and intelligible
a.	Sign production is clear and accurate
b.	Fingerspelling is clear and appropriate for the context
c.	Pausing is appropriate
2.	Standard: overall use of grammatical markers is accurate
a.	Cohesive use of markers (e.g. tense/time indicators, plurals etc.
b.	Use of space appropriate
c.	Effective use of classifiers
d.	Accurate use of pronouns
e.	Accurate use of non-manual sign modifications (e.g. mouth movement, eyebrows, sign movement/intensity, etc.)
3.	Standard: Overall use of sentence structures is appropriate
a.	Use of complete sentences
b.	Sentence structures are appropriately marked (e.g. eyebrows, eye gaze, mouth movements, used to indicate negation, questions etc.

I	**MESSAGE PROCESSING**
1.	Standard: overall message processing results in coherent and accurate interpretation
a.	Understands and conveys speaker/signer (source) goals
b.	Essential elements of meaning/main points conveyed
c.	Appropriate detail conveyed to support main points
d.	Appropriate use of expansions and reductions
e.	Overall discourse strategies use result in coherent target text
f.	Successful management of processing levels
g.	Interpretation is not marked with numerous false starts
h.	Interpretation is generally successful. If not, is there a pattern (deceptive, intrusive or dysfunctional)?
II	**INTERPRETING SUB-TASKS**
1.	Standard: Overall, interpreter comprehends the source message
a.	Analysis of source message, syntax and grammatical features
b.	Monitoring of own work demonstrated/makes corrections appropriately
c.	Effectively mediates culturally-laden elements of the message
d.	Conveys cultural (and other) gestures; verbal and non-verbal cues
e.	Demonstrates awareness of the register for that given situation
2.	Standard: overall target message is accurate
a.	Target language output: overall interpretation grammatically correct
b.	Target language output: overall interpretation semantically accurate
c.	Target language output: appropriate use of discourse markers
III	**MISCUE PATTERNS**
1.	Standard: Overall impact of miscues on interpretation is minimal
a.	If miscues are excessive, is there a pattern (omissions, additions, substitutions, anomalies)?
IV	**ADDITIONAL OBSERVATIONS**
1.	Delivery and flow look natural
2.	Interpreter looks confident
3.	Interpreter demonstrates few/no personal distracting mannerisms

Appendix 19   Official Journal of the European Union / Amtsblatt der Europäischen Union (2006)

Requirements of Interpreting Test

Two admission tests:

1.   10 questions, multiple-choice answers about the EU, its organs of power and its areas of political influence

2.   20 questions, multiple-choice answers about linguistic proficiency and logical thought.
     A pass mark of 15/30 is required to progress to the interpreting examinations:

Interpreting examinations

1.   Consecutive interpreting – maximum length of speech: 6 minutes

2.   Simultaneous interpreting – maximum length of speech: 12 minutes.

A pass mark of 10/20 is required in all interpreting tests

Part I: (A + B + C)

     Consecutive B > A
     Simultaneous B > A
     Consecutive C > A
     Simultaneous C > A

Part II

Option 1: (A + CCC)

     Consecutive C1 > A
     Simultaneous C1 > A
     Consecutive C2 > A
     Simultaneous C2 > A
     Consecutive C3 > A
     Simultaneous C3 > A

Or

Option 2: (AA + C)

     Consecutive A2 > A1
     Simultaneous A2 > A1
     Consecutive A1 > A2
     Simultaneous A1 > A2
     Consecutive C > A1
     Simultaneous C > A1

Formally assessed interview. Candidate's knowledge of the EU, about the workings of the EU, its areas of responsibility. Candidate's ability to work in the multi-cultural and multi-lingual environment of an EU department is also assessed. Length of interview: 30 minutes in candidate's A language.

Pass   mark: 20/40.

Improvements to NAATI testing

Development of a conceptual overview for a new model
for NAATI standards, testing and assessment

Never Stand Still

For:

**The National Accreditation Authority for
Translators and Interpreters (NAATI)**

30 November 2012

Authorised Contact:

Louise Milazzo
Consulting and Contracts Officer
Grants Management Office
Level 3 South Wing
Rupert Myers Building
The University of New South Wales
USNW SYDNEY | NSW | 2052
T: +61 2 9385 4465
F: +61 2 9385 7238
E: l.milazzo@unsw.edu.au
W: www.unsw.edu.au