



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## ARCHIVIO ISTITUZIONALE DELLA RICERCA

### Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Aptitude for conference interpreting

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Russo, M. (2022). Aptitude for conference interpreting. London/New York : Routledge.

*Availability:*

This version is available at: <https://hdl.handle.net/11585/882285> since: 2024-09-28

*Published:*

DOI: <http://doi.org/>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

## **Aptitude for conference interpreting**

*Mariachiara Russo*

### **The concept of aptitude and the construct of aptitude for interpreting**

The word aptitude, which is derived from the Latin term *aptus*, has three different meanings in the online Collins Dictionary: (i) inherent or acquired ability, (ii) ease in learning or understanding; intelligence, and (iii) the condition or quality of being apt. Over the years, various studies and models have been published focusing on general aptitude (i.e. not for a specific competence, such as interpreting), with no universally accepted definition of the term. Indeed, there have been considerable differences of opinion. Nowadays, biological and genetic or environmental factors per se are no longer used to explain a person's predisposition and development. Busch and Reinhart (2005) argue that interpersonal differences are the result of person-environment interaction, where various personality factors come into play, such as: i) above-average intellectual capacity; ii) motivation and readiness to engage; iii) creativity, also in cooperation with primary social environments (family, school and friends); and iv) environmental stimuli. An important criterion of aptitude models is that (...) “they need to describe, explain and predict gifted behavior over time and across situations” (Davidson 2009: 85). A possible division into two separate schools of thought emerges as regards the models and theories relating to aptitude (Chabasse 2009): one sees general aptitude as a product of three equally important and inseparable factors, namely intelligence, creativity and motivation (Mönks 1992; Renzulli 1978); the other supports the existence of natural abilities, distinguishing between giftedness (at least in one domain) and talent, i.e. exceptional performances developed through commitment and practice (Gagné 2000).

Aptitude soon became an issue in the training of aspiring interpreters as a result of the complex psycholinguistic tasks performed by conference interpreters (especially simultaneous interpreters). Indeed, the debate over whether an interpreter is ‘born or made’ (to paraphrase Mackintosh’s seminal work published in 1999) dates back to the very beginnings of the profession. Investigations into the skills of professional interpreters started as early as the 1930s with Sanz (as quoted in Shlesinger & Pöchhacker 2011). The impression that interpreters were extraordinary creatures created by the early consecutive and, above all, simultaneous interpreters at the United Nations,

described by the press as “mythical word wizards” (*The New York Times*, 3 Oct. 1948, as quoted in Baigorri-Jalón 2004: 79), contributed to reinforce the myth of the gifted interpreter with innate talent.

Today, more sophisticated training methods and more systematic training programmes consider interpreting a complex skill that can be acquired, but only after a substantial amount of practice (see Moser-Mercer, Chapter 28, this volume). The difficult task of identifying exactly what it takes to become a conference interpreter has sparked a lively debate on aptitude since a variety of factors contribute to the construct of aptitude for interpreting. In trying to circumscribe the elusive concept of aptitude for interpreting, two insightful definitions from educational sciences appear particularly relevant. The first of these defines the aptitude of learners as “an inverse function of the amount of time, other things being optimal, which will be required for the learner to attain a criterion mastery in the task to be learned” (Carroll 1962, as cited in Bowen & Bowen 1989: 110). The second refers to individual characteristics in relation to learning: “Any aspect of the individual, including some matters untouched by conventional ability and personality measures, can predict response to instruction and hence can be a source of ‘aptitude’” (Cronbach & Snow 1977: 6).

The early detection of personal features conducive to the successful attainment of a training goal leads to ‘aptitude’ predictors. For this very reason, selecting interpreting candidates has become not only a practical necessity for Conference Interpreter Training Programs (CITPs) faced with human and financial restrictions, but also an ethical requirement (especially for those CITPs with a high attrition rate, that is where only a few of the trainees succeed in obtaining conference interpreter certification), as discussed at the *Colloque sur l’enseignement de l’interprétation* (Keiser 1965), where the issue of aptitude was first debated. Keiser stated that anybody could be trained, adding, however:

Mais il serait anormal de permettre à des candidats dépourvus de dons de se lancer dans des études d’interprétation alors qu’il serait patent qu’ils ne deviendront jamais interprètes, ou qu’ils ne le deviendront qu’après quinze ans d’études, ce qui entraînerait une disproportion des efforts injustifiable<sup>1</sup>.  
(Keiser 1965: 3)

The next event in which the issue of aptitude in conference interpreting was discussed was the School Seminar organised by the International Association of Conference Interpreters (AIIC) in 1974 (Bowen & Bowen 1989). Subsequent conferences continued the debate on aptitude: the NATO Symposium (Gerver & Sinaiko 1978), the European Parliament’s Colloquium on Interpretation in a Multilingual Institution (Bowen & Bowen 1989) and the conference organised by the University of Trieste (Gran & Dodds 1989). While aptitude testing was on the agenda at all of

these events and was clearly seen as relevant to training, it was not until the 2009 Antwerp International Symposium, *Aptitude for Interpreting: Towards Reliable Admission Testing*, that an entire conference was devoted to this topic. A selection of the papers presented there can be found in Shlesinger and Pöchhacker (2011).

### **Profile of the ‘ideal candidate’: hard and soft skills**

The moot points of aptitude testing for interpreting revolve around two key issues: what to look for in a prospective candidate and how to measure this in a valid and reliable way. Interpreter trainers (most of whom tend to be professional interpreters) rely on their intuition and personal experience to determine the features and abilities that define ‘aptitude’ for interpreting. The core characteristics of the ‘ideal profile’ for an interpreter do not seem to have changed much over time (see Russo 2011: 9–10). The same linguistic, cognitive and personal traits identified forty years ago are still mentioned in the AIIC website in the section “Linguistic ability alone is not enough to make a conference interpreter. Have you got what it takes?”, which also includes ethical characteristics of professionalism (AIIC 2019). The characteristics listed are the ability to listen and concentrate, a good memory, mental dexterity, speed of thought, a broad general knowledge, analytical skills, a pleasant voice, the ability to keep cool under pressure, the poise required to perform before a large audience and respect for professional confidentiality.

In the current debate on aptitude, only some of these features are still considered core prerequisites for interpreter training, namely command of native and working languages, good memory, broad general culture, verbal intelligence and word fluency (Carroll 1978), language transfer skills, mental quickness, ability to concentrate, ability to deliver the message in a confident and pleasant voice. These can be considered hard skills, or knowledge, i.e., skills and abilities (KSAs), which can be tested empirically or through validated tests. Memory capacity, for example, can be tested through a recall exercise or through standardised tests, such as the Wechsler memory scale; knowledge of L1 and other languages can be tested through an interview or a standard language test, such as a cloze test. Other features could be more difficult to measure, such as the “ability to grasp rapidly and to convey the meaning of discourse” (Gerver et al. 1984: 19; Lambert 1992: 25).

Other features now deemed equally important as these hard skills are what are known as soft skills as they constitute more elusive personal traits and skills that are not amenable to straightforward standard testing procedures. They include love for languages, motivation, open-mindedness, willingness to acquire new information, tact and diplomacy or physical and psychological stamina,

ability to learn quickly and ability to work as a team member. Another set of the personal qualities of an ‘ideal candidate’ based on the experience of leading CITPs are listed by Setton and Dawrant (2016b: 69):

(i) intellectually honest: can faithfully represent another’s point of view despite personal beliefs, knows what s/he doesn’t know, and is not afraid to admit ignorance, has personal integrity and is willing to abide by professional ethics and standards, has calm nerves, composure and the ability to perform under pressure; (ii) is highly motivated: sincerely wants to help people from different cultures communicate, is ready to accept instruction and advice (‘coachability’), is willing to devote long hours outside class to group and individual practice as well as to language and knowledge enhancement, enjoys public speaking and has a pleasant voice, appropriate demeanour and some stage presence, has a sense of humour and can work well as part of a team; shows maturity and developed social skills (‘Emotional Quotient’), is capable of empathy.

Both hard and soft skills can be considered key features to be looked for in an ‘ideal’ interpreter candidate according to both the literature and the experience of CITPs. It remains to be seen, however, which of these hard and soft skills are most predictive. This explains the variety of approaches in the studies on interpreting aptitude (for an overview, see Russo 2011).

### **Key issues in the debate on aptitude testing**

A preliminary issue in the debate on aptitude testing concerns terminology. Some scholars make a distinction between aptitude testing and entrance examinations. During the previously mentioned AICC School Seminar, it was suggested that an ‘entrance examination’ would be preferable to aptitude testing because candidates are not being asked to prove their ability to interpret (Keiser 1978: 18). Furthermore, Bontempo and Napier (2009) distinguished between: (i) true aptitude, i.e. abilities that the candidate must have and that cannot be taught, (ii) abilities expected to have been acquired before interpreter training, and (iii) abilities that will be acquired during training.

It is generally accepted that both hard and soft skills are relevant in a candidate, yet, as Timarová and Ungoed-Thomas (2008) showed, only hard skills have normally been tested, either directly or indirectly (such as memory and language knowledge), whereas soft skills (such as learning ability and affectivity) have mostly been neglected. Only recently have soft skills been attracting more attention (Pöchhacker & Liu 2014; Shlesinger & Pöchhacker 2011; Zannirato 2013).

Another important issue emerging from the debate on interpreting aptitude—and, consequently, also selection procedures—is whether or not greater emphasis should be placed on a skill that is typically acquired during an interpreter programme, such as sight translation (Sunnari 2002) or

short consecutive, which was tested in 14 out of the 18 schools surveyed by Timarová and Ungoed-Thomas (2008), or on the ability to learn interpreting-related skills (Angelelli 2007; Donovan 2003; and Lederer 1975, among others, talked about ‘interpreter readiness’). Seleskovitch and Lederer, in particular, focused on the concept of ‘teachability’ (Keiser 1965) since they considered most training goals as achievable through practice, but others as unteachable, despite their vital importance in interpreter competence. Screening procedures should, therefore, identify ‘teachable’ candidates who are able to think logically, for example, as it would be futile to try to teach interpreting to candidates unable to think ‘logically’.

Another relevant issue concerns the generalisation of test results. Results from the literature over the last forty years indicate that certain tests are more predictive for simultaneous and others for consecutive interpreting (Chabasse 2009; Gerver et al. 1984, 1989; Shang & Xie 2020), indicating that what is tested cannot be generalised for both modes. Furthermore, certain studies indicate that administering several test items at the same time may be redundant and even counter-productive, or unfeasible due to a lack of human resources (Chabasse & Kader 2014; Gerver et al. 1984; Timarová & Ungoed-Thomas 2008;).

Another aspect of aptitude testing concerns course design and content, and also when the test is administered (Bowen & Bowen 1989; Moser-Mercer 1994). In other words, aptitude testing should be considered a relative rather than an absolute construct. Its efficacy in detecting aptitudes depends on the length of the course and whether the test is taken before the course starts, as is the case for almost all CITPs, or after several weeks of interpreting-related exercises.

### **Aptitude testing practice and design**

“Screening potential interpreters”, a seminal paper by Moser-Mercer (1985), is a key issue for CITPs. Nevertheless, most research on aptitude has been carried out by individual scholars and has not necessarily led to the implementation of evidence-based entrance examinations in institutions where they are used (for a similar situation as regards assessment and testing, see Dawrant & Han, Chapter 24). The tests in the first entrance examinations were based on intuition and common sense, presenting what Clifford (2005) called ‘face validity’, that is, they seemed to measure skills a potential interpreting student would need to have (such as the ability to listen and speak at the same time or broad general knowledge). Such tests revealed excessive subjectivity in the assessing of candidates (Dodds 1990; Gerver et al. 1984; Gringiani 1990; Pippa & Russo 2002; Sunnari 2002; Timarová & Ungoed-Thomas 2008). Moreover, they lacked predictive value as “the almost total

lack of data relating pass/fail results in final exams, leaves the evaluation of aptitude for interpreting wide open to criticism” (Moser-Mercer 1994: 65). The issue of reliability led the Advanced School of Modern Languages for Interpreters and Translators of the University of Trieste to revise its non-binding admission procedure. Gringiani (1990) found that the test’s predictive reliability was very poor, with the error percentage for the group that passed the entrance exam, but later dropped out, higher than that for the group that failed, but eventually completed the course.

In his provocative paper “Aptitude of aptitude testing”, Dodds (1990) highlighted the fuzziness of certain expected features (for example, the world knowledge to be reasonably expected of a 20-year-old candidate), the lack of proven scientific soundness of certain procedures (such as sight translation), and the role of psychological factors affecting the result (test-taker’s attitude and motivation or tester’s attitude). Thirty years on, the scientific rigour of some aptitude tests has still not been assessed, while others have provided some evidence-based data, as will be seen.

#### *Skills usually tested in entrance examinations*

The few published reports on the skills tested in entrance examinations provide similar findings. Timarová and Ungoed-Thomas (2008) looked at the entrance tests for 18 interpreting programmes, observing that the screening procedures covered five abilities: language (short consecutive, short speech by candidate, interview, summary and translation), communication (short speech by candidate, short consecutive, summary and interview), comprehension (summary, translation and short consecutive), analysis (summary, short consecutive and translation) and general knowledge (interview, written test and short speech by candidate). The authors concluded that despite the fact that soft skills were considered highly relevant by interviewees, entrance exams only tested hard skills. Furthermore, despite careful selection, on average only 56% of students admitted successfully completed the course. Similar results were obtained by Zha (2016), who provides a comparative picture of entrance examinations in CITPs in Europe (ESIT, Leeds University and Manchester University), Canada (Glendon College, York University), Taiwan (Fu Jen Catholic University and National Taiwan University) and the ten most prestigious CITPs in China (incidentally, entrance examinations are not envisaged in the overwhelming majority of Chinese CITPs, see Dawrant, Wang & Jiang, Chapter 15, this volume). Zha concludes that the entrance examinations in all these programmes focused solely on hard skills—language, communication skills, comprehension skills and encyclopaedic knowledge—which are assessed using procedures similar to those identified by Timarová and Ungoed-Thomas (2008), such as cloze tests, consecutive, translation, oral summary and oral interview.

### *Test designs to be put to the test*

Over the last two decades a number of proposals for aptitude tests have been put forward. Moser-Mercer (2000) described the Geneva Aptitude Project, based on the expert-novice paradigm; Bernstein and Barbier (2000) suggested a rapid automatic screening test based on the PhonePass-SET-10 test methodology, which produces a fully automatic English spoken test; Clifford (2005) introduced a psychometric test for interpreter certification targeted at single operations (such as comprehension), which could be taken in consideration when developing an aptitude test; and Schweda Nicholson (2005) suggested including a personality test—the Myers-Briggs Type Indicator (MBTI)—in the aptitude testing battery. More recently, Zannirato (2013) suggested testing soft skills, in particular affectivity and self-perception, in addition to psychometric testing for hard skills (memory and language skills). He also stated that tests should be accompanied by a qualitative approach based on triangulation, i.e. the use of more than one method or source of information to gain a better insight into a candidate's suitability for the training programme. "Examples may include the analysis of personal statements, group interviews, and one-on-one interviews" (Zannirato 2013: 120). Most of the above-mentioned contributions provide conclusions based on studies carried out by the authors themselves. Setton and Dawrant (2016a), on the other hand, use a different approach in their textbook for trainers adding their own reflections to the debate on aptitude. They suggest an entrance examination organised in five steps, combining both written and oral tests: (1) a selected-response test (such as a multiple choice test), which is written and can be objectively and rapidly scored, and is able to test knowledge of working languages, world knowledge, comprehension and analysis; (2) a voice and delivery test through a reading aloud task; (3) translation (L1<>L2 or L2/L3>L1); (4) a written paraphrase in L1 and L2 and/or a cloze test in L1 or L2; (5) and an oral interview. Test types 3 and 4 are constructed-response tests, which can only be judged subjectively, so it is essential to ensure inter-rater agreement in the scoring of candidates' linguistic output.

### **Empirical studies**

As previously mentioned, a number of trainers and scholars have recognised the need to adopt methodologies that can be checked and replicated, breaking down 'interpreting aptitude' and interpreting performance into constructs that can be defined and measured. They have also advocated the need to administer more valid and reliable testing procedures (see Russo 2011). The methods and main results of the most cited empirical studies published so far will be described in



chronological order and subdivided according to their main focus (hard or soft skills), even though some test batteries address both skills.

### *Testing hard skills*

The first CITP to carry out a pilot test battery based on a scientific approach was the Polytechnic of Central London, where a multidisciplinary team, which also included psychologists, was created (Gerver et al. 1984, 1989). Gerver et al. (1984: 19) concentrated on the “ability to grasp rapidly and to convey the essential meaning of what is being said” from among the several suggested characteristics listed in the ‘ideal interpreter profile’ (see Russo 2011) and used twelve psychometric tests, that they considered to be a reflection of the task of an interpreter. There were three different types of tests, which were administered aurally to 29 subjects (several of whom dropped out during testing): 1) text-based tests aimed at assessing text processing skills: recall of the information presented or the completion of individual target words in a text; 2) subskill-based tests aimed at assessing verbal ability: synonym generation, rewriting and vocabulary selection; 3) stress-based tests aimed at assessing performance under a time constraint: the solving of problems using letter series. The test scores were correlated with final exam outcomes, providing the following main results: i) differences between students who passed and those who failed exams were significant for text memory, logical memory, error detection, cloze and synonym tests; ii) differences in exam outcomes between simultaneous and consecutive interpreting—completion and detection tests correlated more with the marks obtained in simultaneous exams, while recall tests correlated more with those obtained in consecutive exams; iii) for subskills and speed stress tests, only the synonyms test significantly reflected differences between candidates who passed and those who failed exams.

This pilot test had several merits. Firstly, a generic feature of the aptitude construct, such as the “ability to grasp rapidly and to convey the essential meaning of what is being said”, was broken down into scorable tasks, based on psycholinguistic literature and procedures. This showed that some tasks were more accurate than others in predicting successful performance in consecutive or simultaneous interpreting. Given the different nature of the tasks administered to candidates, Gerver et al. (1989) also suggested that since the processing of connected discourse constitutes a crucial feature of the interpreter’s task, it should be included in selection tests. The limitations of this pilot test lie in the sample size and the fact that it was not possible to fully assess its validity and reliability because this would have required the independent assessment of a large enough number of candidates by at least two raters (Gerver et al. 1984: 20).

A reformulation test to detect a candidate's mental flexibility and expressive ability, two key interpreting-related skills, was suggested by Russo (1989) in her oral paraphrasing test. The validity of paraphrase as a diagnostic tool had already been proposed by Moser-Mercer (1978, 1985). She observed that students who performed poorly in paraphrasing obtained the lowest final grades in the introductory course to simultaneous interpreting. Since some authors question the existence of similarities between intralinguistic and interlinguistic translation (see Pippa and Russo 2002; Russo 2014), the use of monolingual paraphrasing as a predictive indicator of interpreting competence could also be challenged. Inspired by Moser-Mercer's work, Russo started to collect data on her time-constrained, synchronous oral Italian-Italian paraphrasing task (Pippa & Russo 2002). The candidates' linguistic reformulations and consequent semantic and pragmatic shifts were classified according to a specific set of categories and, for statistical analysis, the candidates' paraphrases were operationalized into 75 triplets of shift items (Russo & Pippa 2004).

The test was administered to 46 subjects and the preliminary results revealed that only two of the candidates' text-based operations (assessed by two independent raters with an inter-judge agreement of 0.75 Cohen's K-coefficient) were statistically significant predictors: loss of coherence (especially at sentence level) and synonym substitution (with or without pragmatic loss). These correlated with both the students' average mark in the eight interpreting exams in every language combination and the time taken to pass all the interpreting exams. The follow-up study involving 64 students (Russo 2014) confirmed that the ability to use synonyms showed the highest validity in relation to the time taken to pass all the interpreting exams and correctly classified 48 out of the 64 students (75%) as 'slow' (test sensitivity = 80%) or 'fast' (test specificity = 71%). These results were statistically significant. The merit of this empirical investigation is that it reveals that synonym production during a time-constrained synchronous paraphrasing task in the candidate's L1 appears to be a promising procedure to screen the cognitive and linguistic abilities of best performer, and could, therefore, be usefully included in a multi-component test battery (as confirmed by Setton & Dawrant 2016a: 134). The limitation of the second empirical study by Russo (2014) lies in the absence of a second rater and, in general, in the difficulty in scoring meaning equivalents manually, despite the availability of a simplified grid classifying only three types of synonyms.

Another empirical study investigating synonym production as a means of aptitude testing was published by Pöchhacker (2011). His SynCloze test, which combined an oral cloze and a task requiring high expressional fluency, required candidates to rapidly find suitable synonymic

sentence completions. The purpose of the test, administered to 116 students attending an introductory course on interpreting, was to investigate interpreting aptitude by assessing the accuracy and speed of response in producing synonyms to suitably end a sentence (in German the element which is essential to understand the meaning of a sentence, namely the verb, is placed in the final position). For institutional reasons, Pöchhacker was unable to correlate his test scores with the marks obtained in consecutive and simultaneous interpreting exams proper. He was, however, able to correlate them with the students' performance in the introductory course exam (a German-into-German short consecutive), observing a significant correlation with test scores in 24 students and a moderate correlation in 56 students. This study offers further evidence of the relevance of testing expressional fluency in the candidate's L1. The main limitation of the test is the lack of correlations with tests that demonstrate the successful completion of an interpreter training programme, such as final examinations in consecutive and simultaneous interpreting.

Chabasse (2009) developed her cognitive shadowing test to evaluate language proficiency, fluency and the ability to concentrate. The test consisted in six sections, each comprising 30 questions that candidates (without prior experience in simultaneous interpreting) had to answer after a certain time lag. The aim of the test was to assess a candidate's ability to listen, speak and think at the same time. In 2009, the innovative "MA KD Germersheim Project" was launched to find the most predictive screening procedure for simultaneous interpreting (Chabasse & Kader 2014). The authors compared the predictive validity of three aptitude tests for simultaneous interpreting administered in the booth to 24 candidates: Pöchhacker's SynCloze test (2011), Chabasse's cognitive shadowing test (2009) and Timarová's personalised cloze test, which was presented at the 2009 Antwerp Symposium and consisted in listening to a short text and shadowing it in L2, substituting the original story with personal details. The performances in the three tests were scored by adding points for positive items (such as correct answers, substituted items etc.) and deducting points for negative items (such as wrong answers, grammatical errors etc.). The authors do not, however, provide any information on inter-judge agreement. The test scores were correlated with exam grades at the end of the second semester (not at the end of the MA course, which lasts four semesters). Cognitive shadowing was identified as the most predictive test for SI from L2 to L1 and from L1 to L2. The limitation of this empirical study is that it only correlated test scores with interim exam results, with the authors themselves stating that they would only be able to provide more conclusive data after the 24 students had successfully completed their MA course. This study has the merit of showing the suitability, feasibility and shortcomings of different tests when

compared with one another. It is to be hoped that comparative studies of this kind will be replicated, with a wider sample and across different CITPs.

Recall, both intralingual and interlingual, is one of the most widely used tests (Setton & Dawrant 2016a; Timarová & Ungoed-Thomas 2008; Timarová & Ungoed-Thomas 2009;) as it is considered a useful way of assessing linguistic and communicative suitability for conference interpreter training. Its validity has been tested in several empirical studies by correlating the scores obtained with results in the final exams (Arjona-Tseng 1994; Donovan 2003; Gerver et al. 1984, 1989; Timarová & Ungoed-Thomas 2009). A recent study also tested its validity for non-European languages. Shang and Xie (2020) tested recall across English and Chinese using four parameters: fidelity, language, coherence and delivery. The performance of 52 subjects with Chinese as L1 were jointly scored by 4 trained raters (Cronbach's alpha inter-rater reliability over 0.7) and correlated with consecutive interpreting exams. The data show that recall can predict interpreting performance in both directions. Language, however, is a more accurate predictor than fidelity in the Chinese-English recall, while fidelity is a more accurate predictor than language in the English-Chinese recall. The study shows the importance of having a good command of the foreign language from the very start of training, in terms of both production and comprehension, to be able to successfully complete training. The merit of this study lies in its rigorous approach, focusing on scoring reliability, a factor that greatly affects the validity of the results. As is the case in most empirical studies in interpreting, its limitation lies in the sample size and, as also acknowledged by the authors, a still 'crude' scoring rubric.

### *Testing soft skills*

So far research into aptitude testing has mainly focused on hard skills, namely cognitive and linguistic skills, though, more recently, scholars have also begun investigating other components of an 'ideal' entrance examination, namely soft skills. The first results were presented at the 2009 Antwerp Symposium. Timarová and Salaets (2011) investigated learning styles, motivation and cognitive flexibility by means of specific tests. The authors carried out a two-stage study. In the first stage, they compared three groups of students: (1) 9 applicants for the postgraduate conference interpreter programme; (2) 23 applicants for the MA liaison interpreting programme; and (3) 104 undergraduate students in applied language studies. Students in groups (1) and (2) were considered a single group of "self-selected interpreting students", while students in group (3) were considered the control group. The test scores showed that self-selected interpreting students were cognitively more flexible, more achievement-driven and less negatively affected by anxiety. The differences

between the two groups were used in the second stage of the study to verify whether these differences could highlight an aptitude for interpreting. 14 of the 32 self-selected interpreting students started the conference interpreter programme, so the authors were able to correlate their test scores with their pass/fail grade in their final exams. The study showed that, compared with the control group of undergraduate students, conference interpreting students who passed their final exams (but not unsuccessful ones) were cognitively more flexible and benefited more from some level of anxiety. Furthermore, all conference interpreting students were less affected by stress, with apparently clearly developed learning styles. The authors, therefore, conclude that these tests could be considered promising aptitude predictors.

Shaw (2011) investigated cognitive and motivational contributors to aptitude measured by standardised performance and motivation tests. She studied spoken (SP) and sign language (SL) interpreting students (29 and 18 respectively) at their respective entry and advanced levels, and found that SL students have more highly developed visual memory skills in the presence of distracting information and retain visual information longer than SP students, and SP students were more vulnerable to distractions and more likely than SL students to believe that their successes were due to their own efforts. She also found differences between advanced and entry-level students in the areas of concentration and eagerness to learn new concepts in the absence of external reward. The finding that students required less time to perform certain tasks as they progress through the programme demonstrate that certain skills, such as faster reaction times or the ability to concentrate for longer period of time, are teachable throughout the course of study. The author believes that this can be a helpful indicator in the selection of items used in screening procedures.

The role of motivation, linguistic self-confidence and language anxiety as potential aptitude predictors is also confirmed by the study on individual difference (ID) variables by Rosiers et al. (2011). They compared the answers of 35 undergraduate students (12 interpreters and 23 translators) to two questionnaires. The authors found statistically significant differences in the ID profiles of interpreters and translators. These differences did not, however, help predict interpreting performance in a sight-translation assignment, where self-confident students (interpreters) did not outperform students with marked language-anxiety profiles (translators). This led the authors to call for some caution regarding admission testing practices that take candidates' ID profiles into account, as these did not seem to be accurate predictors of success, at least in the short term.

Macnamara et al. (2011) carried out a study combining hard and soft skills in low- and highly-skilled sign language interpreters (14 and 15 respectively). The authors focused on domain-general cognitive abilities, including working memory capacity, information processing speed, fluid intelligence, cognitive control, task switching ability, mental flexibility and psychomotor speed, as these are all relevant general cognitive skills that affect interpreting performance. They were also deemed relevant for spoken simultaneous interpreting. The authors also measured personality traits, including risk-taking orientation, emotion-cognition integration style and intrinsic motivation to undertake complex cognitive tasks. Significant differences were found between the two groups on both cognitive and personality measures. The authors proved that the mix of stable domain-general cognitive abilities and personality traits may help distinguish between highly-skilled and lesser-skilled interpreters and may, therefore, help predict individual future interpreting competence and level of expertise.

Starting from the premise that the ability to manage stress is a prerequisite for interpreters (see, among others, Kurz 2003; Moser-Mercer 1985), Bontempo and Napier (2011) were the first scholars to attempt to obtain a psychological profile from a representative sample of qualified interpreters by measuring emotional stability to ascertain whether this is a predictor of interpreter professional competence and, therefore, could be a useful predictor to be included in an aptitude test for interpreting. Using validated psychometric tools, they tested 110 sign language interpreters for personality aspects considered predictive of performance in a variety of contexts, namely self-efficacy, goal orientation and negative affectivity, and proved that emotional stability is the strongest predictor. Given, however, that the impact of the many cognitive, non-cognitive and environmental factors affecting interpreter aptitude and performance is still not fully understood, “it may be unwise to implement formal personality testing in program admission screening” (Bontempo & Napier 2011: 99).

### **Conclusive remarks**

Research on aptitude testing is a complex task as there are various constraints that make it difficult to validate the procedures used in such tests. Firstly, there are legislative, institutional, and organisational differences among CITPs, which make it hard not only to compare the screening procedures they use, but also to transfer successful screening procedures from one institution to another. Secondly, it is difficult to develop objective evaluations that make methodologies replicable and prevent subjective biases. Thirdly, unlike other disciplines such as psychology,

medicine or sociology, where population studies can rely on large numbers, in interpreting studies there are often insurmountable difficulties in enrolling large, homogeneous samples of interpreting candidates to obtain more accurate estimates of aptitude parameters. Lastly, CITPs may find it difficult to keep track of candidates' performances and correlate test results with final exam outcomes due to intervening changes within the training programme, teaching staff turn-over and lack of methodological continuity among educator-researchers. Nevertheless, available data do highlight a certain convergence as regards the predictive hard and soft skills to be looked for in potentially successful candidates, which were intuitively proposed in the 'ideal candidate profile'.

The predictive value of certain hard skills, such as the 'word fluency factor' (Carroll 1978) and re-expression, has been demonstrated by the results of studies on the production of synonyms and of synonymic expressions (Gerver 1984; Moser-Mercer 1985; Pöchhacker 2011; Russo & Pippa 2004; Russo 2014;). Similarly, above all after the 2009 Antwerp Symposium, the predictive value of certain soft skills, such as emotional stability and motivation, has been borne out by several studies (Bontempo & Napier 2009, 2011; Shaw 2011; Timarová & Salaets 2011).

The quest for the most efficient battery of aptitude tests to screen potential interpreters needs to strike a balance between criteria that are scientifically sound and institutional constraints linked to logistics, finance and human resources. Binding aptitude tests are high-stakes tests, given their profound ethical, personal, institutional and financial consequences for students, trainers and CITPs alike. Further empirical studies are, therefore, needed to achieve sound, evidence-based screening procedures. This, in turn, requires far greater use of statistical models to confirm the predictive power of these screening procedures in order to have reliable aptitude-based selection criteria for candidates. While, after almost 60 years of research, no single test or battery of tests is unanimously considered predictive, this chapter shows that certain empirical studies have at least laid the foundations for promising predictive tests. It also highlights how the debate on aptitude in the field of conference interpreting has profitably moved forward from the far too narrow, early binary distinction between the 'born vs made' interpreter. In their special issue of *Interpreting*, Shlesinger and Pöchhacker (2011: 4) wrote: "The results represent an important milestone but nevertheless suggest the need for much further work before researchers can hope to supply educators with a full range of reliable and practical methodologies for testing the aptitude of applicants to interpreter training programs". Ten years later, this statement still holds true.

### **Further reading**

Angelelli, Claudia V. & Holly E. Jacobson (eds) 2009. *Testing and Assessment in Translation and Interpreting Studies*. American Translators Association Scholarly Monograph Series, XIV. Amsterdam/Philadelphia: John Benjamins.

Tsagari, Dina & van Deemter, Roelof (eds) 2013. *Assessment Issues in Language Translation and Interpreting*. Frankfurt am Main: Peter Lang.

## References

AIIC 2019. Linguistic ability alone is not enough to make a conference interpreter. Have you got what it takes? <https://aiic.org/site/world/conference/whatittakes> (accessed April 30 2021)

Angelelli, Claudia V. 2007. Assessing medical interpreters: The Language and Interpreting Testing Project. *The Translator* 13 (1), 63–82.

Arjona-Tseng, Etiliva 1994. Psychometric selection tests. In S. Lambert & B. Moser-Mercer (eds) *Bridging the Gap: Empirical Research in Simultaneous Interpretation*. Amsterdam/Philadelphia: John Benjamins, 69–86.

Baigorri-Jalón, Jesús 2004. *Interpreters at the United Nations: A History*. Salamanca: Ediciones Universidad de Salamanca.

Bernstein, Jared & Barbier, Isabella 2000. Design and development parameters for a rapid automatic screening test for prospective simultaneous interpreters. *Interpreting* 5 (2), 221–238.

Bontempo, Karen & Napier, Jemina 2009. Getting it right from the start: Program admission testing of signed language interpreters. In C. V. Angelelli & H. E. Jacobson (eds) *Testing and Assessment in Translation and Interpreting Studies: A Call for Dialogue Between Research and Practice*. Amsterdam/Philadelphia: John Benjamins, 247–295.

Bontempo, Karen & Napier, Jemina 2011. Evaluating emotional stability as a predictor of interpreter competence and aptitude for learning. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 85–105.



- Bowen, David & Bowen, Margareta 1989. Aptitude for interpreting. In L. Gran & J. Dodds (eds) *The Theoretical and Practical Aspects of Teaching Conference Interpretation*. Udine: Campanotto Editore, 109–125.
- Brisau, André, Godjins, Rita & Meuleman, Chris 1994. Towards a psychometric profile of the interpreter. *Meta* 39 (1), 87–94.
- Busch, Karin & Reinhart, Ulrike 2005. *Theoretische Grundlagen zum Begabungsbegriff*. Akademielehrgang Begabungsförderung Modul 1. Linz.
- Carroll, John B. 1962. The prediction of success in intensive foreign language training. In R. Glaser (ed.) *Training Research and Education*. Pittsburgh: University of Pittsburgh Press, 87–136.
- Carroll, John B. 1978. Linguistic abilities in translators and interpreters. In D. Gerver & H. W. Sinaiko (eds) *Language Interpretation and Communication*. New York: Plenum Press, 119–130.
- Chabasse, Catherine 2009. *Gibt es eine Begabung für das Simultandolmetschen? Erstellung eines Dolmetscheignungstests*. Berlin: Saxa-Verlag.
- Chabasse, Catherine & Kader, Stephanie 2014. Putting interpreting admission exams to the test. The MA KD Germersheim Project. *Interpreting* 16 (1), 19–33.
- Clifford, Andrew 2005. Putting the exam to the test: Psychometric validation and interpreter certification. *Interpreting* 7 (1), 97–131.
- Cronbach, Lee J. & Snow, Richard E. 1977. *Aptitudes and Instructional Methods*. New York/London/Toronto/Sydney: Irvington Publishers.
- Davidson, Janet E. 2009. Contemporary models on giftedness. In L. V. Shavinina (ed.) *International Handbook on Giftedness*. New York: Springer Science + Business Media, 81–98.
- Dodds, John M. 1990. On the aptitude of aptitude testing. *The Interpreters' Newsletter* 3, 17–22.

Donovan, C. 2003. Entrance exam testing for conference interpretation courses: How important is it? *Forum* 1 (2), 17–45.

Gagné, François 2000. Understanding the complex of talent development through the DMGT-based analysis. In K. Heller, F. J. Mönks, R.J. Sternberg & R.F. Subotnik (eds) *International Handbook of Giftedness and Talent*. Amsterdam: Elsevier, 67–79.

Gerver, David, Longley, Patricia, Long, John & Lambert, Sylvie 1984. Selecting trainee conference interpreters. *Journal of Occupational Psychology* 57, 17–31.

Gerver, David, Longley, Patricia, Long, John & Lambert, Sylvie. 1989. Selection tests for trainee conference interpreters. *Meta* 34 (4), 724–735.

Gerver, David & Sinaiko, H. Wallace (eds) 1978. *Language Interpretation and Communication*. New York: Plenum Press.

Gran, Laura & Dodds, John (eds) 1989. *The Theoretical and Practical Aspects of Teaching Conference Interpretation*. Udine: Campanotto Editore.

Gringiani, Angela 1990. Reliability of aptitude testing: A preliminary study. In L. Gran & C. Taylor (eds) *Aspects of Applied and Experimental Research in Conference Interpretation*. Udine: Campanotto Editore, 42–53.

Keiser, Walter 1965. Admission dans les écoles d'interprétation. In AIIC (ed.) *Colloque sur l'enseignement de l'interprétation, Paris 18–19 Décembre*. Geneva: AIIC, 3–8.

Keiser, Walter 1978. Selection and training of conference interpreters. In D. Gerver & H. W. Sinaiko (eds) *Language Interpretation and Communication*. New York: Plenum Press, 11–24.

Kurz, Ingrid 1992. 'Shadowing' exercises in interpreter training. In C. Dollerup & A. Loddegaard (eds) *Teaching Translation and Interpreting: Training Talent and Experience*. Amsterdam: John Benjamins, 245–250.

Kurz, Ingrid 2003. Physiological stress during simultaneous interpreting: A comparison of experts and novices. *The Interpreters' Newsletter* 12, 51–67.

Lambert, Sylvie 1992. Aptitude testing for simultaneous interpretation at the University of Ottawa. *The Interpreters' Newsletter* 4, 25–32.

Lederer, Marianne 1975. Le colloque sur les examens d'entrée aux Écoles d'Interprètes. *AIIC Bulletin* 3 (2), 25–28.

López Gómez, María J., Bajo Molina, Teresa, Padilla Benítez, Presentación & Santiago de Torres, Julio 2007. Predicting proficiency in signed language interpreting. *Interpreting* 9 (1), 71–93.

Mackintosh, Jennifer 1999. Interpreters are made not born. *Interpreting* 4 (1), 67–80.

Macnamara, Brooke N., Moore, Adam B., Kegl, Judy A. & Conway, Andrew 2011. Domain-general cognitive abilities and simultaneous interpreting skill. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 121–142.

Mönks, Franz J. 1992. Development of gifted children: The issue of identification and programming. In F. J. Mönks & W. A. M. Peters (eds) *Talent for the Future*. Assen, The Netherlands: Van Gorcum, 191–202.

Moser, Barbara 1978. Simultaneous interpretation: A hypothetical model and its practical application. In D. Gerver & H. W. Sinaiko (eds) *Language Interpretation and Communication*. New York: Plenum Press, 353–368.

Moser-Mercer, Barbara 1985. Screening potential interpreters. *Meta* 30 (1), 97–100.

Moser-Mercer, Barbara 1994. Aptitude testing in conference interpreting: Why, when and how. In S. Lambert & B. Moser-Mercer (eds) *Bridging the Gap. Empirical Research in Simultaneous Interpretation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 57–68.

- Moser-Mercer, Barbara 2000. Searching to define expertise in interpreting. In B. Englund Dimitrova & K. Hyltenstam (eds) *Language Processing and Simultaneous Interpreting: Interdisciplinary Perspectives*. Amsterdam/Philadelphia: John Benjamins, 107–131.
- Pippa, Salvador & Russo, Mariachiara 2002. Aptitude and simultaneous interpretation: A proposal for a testing methodology based on paraphrase. In G. Garzone & M. Viezzi (eds) *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam/Philadelphia: John Benjamins, 245–256.
- Pöchhacker, Franz 2011. Assessing aptitude for interpreting: The SynCloze test. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 106–120.
- Pöchhacker, Franz & Liu, Minhua (eds) 2014. *Aptitude for Interpreting*. Amsterdam/Philadelphia: John Benjamins.
- Renzulli, Joseph S. 1978. What makes giftedness: Reexamining a definition. *Phi Delta Kappan* 60, 180–184.
- Rosiers, Alexandra, Eyckmans, June, Bauwens, Daniel 2011. A story of attitudes and aptitudes? Investigating individual difference variables within the context of interpreting. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 53–69.
- Russo, Mariachiara 1989. Text processing strategies: A hypothesis to assess students' aptitude for simultaneous interpreting. *The Interpreters' Newsletter* 2, 57–64.
- Russo, Mariachiara & Pippa, Salvador 2004. Aptitude to interpreting: Preliminary results of a testing methodology based on paraphrase. *Meta* 49 (2), 409–432.
- Russo, Mariachiara 2011. Aptitude testing over the years. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 5–30.
- Russo, Mariachiara 2014. Testing aptitude for interpreting. The predictive value of oral paraphrasing, with synonyms and coherence as assessing parameters. In F. Pöchhacker & M. Liu (eds) *Aptitude for Interpreting*. Amsterdam/Philadelphia: John Benjamins, 129–146.

Sanz, Jesús 1930. Le travail et les aptitudes des interprètes parlementaires. *Analys d'Orientació Professional* 4, 303–318.

Schweda Nicholson, Nancy 2005. Personality characteristics of interpreter trainees: The Myers-Briggs Type Indicator (MBTI). *The Interpreters' Newsletter* 13, 109–142.

Setton, Robin & Dawrant, Andrew 2016a. *Conference Interpreting: A Trainer's Guide*. John Benjamins Publishing Company.

Setton, Robin & Dawrant, Andrew 2016b. *Conference Interpreting: A Complete Course*. John Benjamins Publishing Company.

Shang, Xiaoqi & Xie, Guixia 2020. Aptitude for interpreting revisited: predictive validity of recall across languages. *The Interpreter and Translator Trainer*, 344–361.

<https://doi.org/10.1080.1750399X.2020.1790970>

Shaw, Sherry 2011. Cognitive and motivational contributors to aptitude. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1), 70–84.

Shlesinger, Miriam & Pöchhacker, Franz (eds) 2011. *Aptitude for Interpreting*. Special Issue of *Interpreting* 13 (1).

Sunnari, Marianna 2002. Aptitude tests and selection criteria for interpreting students. In *Teaching simultaneous interpretation into a "B" language. Proceedings of EMCI Workshop*. Paris, 23–26.

Paris: ESIT. [https://www.emcinterpreting.org/emci\\_drupal\\_data/EMCI-TeachingSimultaneousIntoB-vol1.pdf](https://www.emcinterpreting.org/emci_drupal_data/EMCI-TeachingSimultaneousIntoB-vol1.pdf) (accessed 30 May 2020)

Timarova, Šarka. & Salaets, Heidi 2011. Learning styles, motivation and cognitive flexibility in interpreting training. In M. Shlesinger & F. Pöchhacker (eds) *Aptitude for Interpreting*. Special Issue of *Interpreting*. 13 (1), 31–52.

Timarová, Šarka & Ungoed-Thomas, Harry 2008. Admission testing for interpreting courses. *The Interpreter and Translator Trainer* 2 (1), 29–46.

Timarová, Šarka & Ungeod-Thomas, Harry 2009. The predictive validity of admission tests for conference interpreting courses in Europa: A case study. In C. V. Angelelli & H. E. Jacobson (eds) *Testing and Assessment in Translation and Interpreting Studies*. Amsterdam/Philadelphia: John Benjamins, 225–246.

Zha, Jianshe 2016. The past, present and future of interpreting aptitude test. *Chinese Translators Journal* 4, 54–60.

Zannirato, Alessandra 2013. The quest for ‘perfection’. *The Interpreter and Translator Trainer* 7 (1), 107–127.

---

<sup>i</sup> But it would be abnormal to allow untalented candidates to embark on studying interpreting when it is clear that they will never become interpreters, or that they will only become interpreters after fifteen years of study, which would result in an unjustifiable disproportion of effort.