

ON THE APTITUDE OF APTITUDE TESTING

By

John M. Dodds

SSLM, University of Trieste

Catherine Stenzl during the discussion on aptitude testing that took place in the course of the Symposium on the Theoretical and Practical Aspects of Teaching Conference Interpretation threw down the gauntlet, so to speak, when she said that *"the validity of our entrance tests has not been proved yet, because binding admission tests preclude an assessment of whether the failed candidate might in time have become a good interpreter. The work of the schools where the test is not binding, therefore, becomes essential. They will be able to inform us about the performance of students who failed the test, enabling an evaluation of the entrance tests."*¹

Now, the Trieste School, as is well known, has and has to have a non-binding aptitude test as Italian law does not permit any form of selection at the beginning of the interpreting or translating courses. So research into aptitude testing has been organized (notably that carried out by Angela Paola Gringiani², Chiara Russo³, and Christopher Taylor⁴) and quite frankly some of their results have confirmed what I and many others have long suspected, namely that these tests in their present form are unreliable, subjective and therefore unable to correctly predict the candidate's future academic and professional performance. For example, according

to Gringiani's figures referring to the Trieste school, in one year there was an appallingly high percentage of failed candidates who then went on to complete the course and become professional interpreters (36%) and an even higher percentage of successful candidates "with aptitude" who then failed to complete the course (45%).

Now although the figures were not this bad in other years, these percentages must nevertheless be considered quite unacceptable as a reliable indicator of potential, although a great deal more work has to be done on the motivation/demotivation factor. Clearly, if 36% of the students were failed, i.e. were judged not to have the aptitude to become conference interpreters, then either the subjective or objective criteria of the examiners were wrong in 36% of the cases or the teaching during the course was so good that after a certain period of time the students "acquired" the necessary aptitude or the psychological make-up of certain candidates and not others is such that a failure is going to greatly increase their motivation. Conversely, if 45% of the candidates who were judged to have the necessary aptitude for the profession fail to complete the course, then either the subjective or objective criteria are wrong once again or the teaching is atrocious or the students rest smugly on their laurels for the duration of the course to be promptly failed in the examinations.

Clearly the reasons are varied and multiple and certainly depend on a combination of all of these factors. I would nevertheless like to take a closer look at testing and test criteria in general and also with specific reference to what we do in Trieste, as I myself have been a rather perplexed "aptitude tester" for some years now. I shall not, however, be going into the whole question of successive teaching quality or the possible lack of quality as there are clearly far too many

¹ Catherine Stenzl in *The Theoretical and Practical Aspects of Teaching Conference Interpretation*, (eds. Laura Gran & John Dodds), Campanotto editore, Udine, 1989, p.266.

² Angela Gringiani, 'Reliability of Aptitude Testing: A Preliminary Study' in *Aspects of Applied and Experimental Research on Conference Interpretation* (eds. L. Gran & C. Taylor), Campanotto editore, Udine, 1990, pp. 42-53.

³ Chiara Russo, 'Text Processing Strategies: A Hypothesis to Assess Students' Aptitude for Simultaneous Interpreting' in *The Interpreters' Newsletter N°2* (eds. L. Gran & J. Dodds), SSLM, Univ. of Trieste, Trieste, 1989, pp. 57-64.

⁴ Christopher Taylor, 'Coherence and World Knowledge' in *Aspects of Applied and Experimental Research on Conference Interpretation* (eds. L. Gran & C. Taylor), Campanotto editore, Udine, 1990, pp. 21-27.

variables involved but I would like to return afterwards to the psychological question of motivation and demotivation which is very important if not essential for the success of any teaching activity and also for the actual testing of the candidates themselves.

Before looking at the individual tests we use in Trieste, I must make one further yet nevertheless important generalization. To throw a child into the deep end of a swimming pool to see if he/she will sink or swim and then to say that the one who swims could become a professional while the one who sinks will never learn is, I hope we all agree, total and utter rubbish. So, why do so many of us, myself included, assess aptitude on the basis of tasks that many or indeed most of the candidates have never previously performed?

This is one of the reasons we in Trieste avoid shadowing tests (and particularly phonemic and adjusted lag shadowing) in that apart from being diametrically opposed to what interpreting is all about, it is also a totally artificial activity that has never been performed previously and hopefully will never be repeated in the future. Here I am being particularly and intentionally provocative in the hope that more discussion on the subject of shadowing might be elicited and more real research carried out in order to verify or falsify what I personally feel is very much a myth, namely the usefulness of shadowing as either an aptitude test or as an introductory exercise to simultaneous interpreting.

In this same issue of *The Interpreters' Newsletter*, the position for shadowing as a valid "screening examination" is put forward by Nancy Schweda Nicholson who claims that students taking this sort of test "may, for example, show no improvement during a five-minute shadowing passage on a screening examination. In such a case, these students may simply not be suited for interpretation. One of the great values of shadowing exercises is their ability to quickly identify those candidates who appear to be promising trainees". She then goes on to say that "shadowing has consistently proven to be a strong indicator of future performance".

Again, in this same issue of *The Interpreters' Newsletter*, the position against shadowing is expressed by Ine van Dam in what appears to be no less a statement of faith rather than fact when she calls shadowing a "pernicious exercise", stating that the harm it causes "is not only potential, it is real and it is often irreversible".

Now, as I have just said but I wish to repeat, both these stances are, I believe, based on faith and not on proven scientific fact or on the results of tried and repeated experimentation. However, where I feel the anti-shadowing advocates have

the upper hand is over the fact that shadowing and particularly phonemic and adjusted lag shadowing is simply the repetition of sounds and words and, as Ine van Dam says quite categorically "that's not at all what we conference interpreters do!" I cannot but concur, especially when Nancy Schweda Nicholson in her article uses so many hypothetical modals like 'can' and 'may' with reference to pre-course testing, to then state equally categorically that the usefulness of shadowing "has consistently been proven". I personally feel that we have not yet proven anything and that it is high time we did so either not to miss out on an invaluable if not infallible yardstick for measuring aptitude or avoid using an exercise that potentially could be very harmful in the training of young interpreters. I myself would be more than happy to conduct these tests in Trieste or to invite any of the pro-shadowing advocates to come and supervise their shadowing exercises in place of our normal aptitude test and then to monitor students' progress over their two-year course to see to what extent end results correspond to the initial shadowing tests. In this way, at least some first scientific steps towards the verification or falsification of the shadowing hypothesis will be made.

But what about more traditional aptitude tests based on summarizing, on sight translation and on general knowledge - the latter frequently being totally arbitrary and rather more reminiscent of a television quiz programme like 'Double Your Money' than a serious test which may well influence the careers of our students?

First of all then, summary writing. In Trieste this test consists of a semi-improvised speech delivered in English and then the candidates are asked to write a short summary of the speech in Italian (their A language), obviously after having taken notes during the delivery of the speech. A most useful and highly indicative test for potential interpreters, I'm sure we would all agree. It shows the candidates' ability a) to understand a foreign SL text, b) to select relevant information from redundant material or material of secondary importance, c) to be concise and to the point and d) to express themselves in their native tongue. This, at least, is the theory. In practice, however, what often happens is that the candidates almost always attempt to reproduce every word and if omissions are made this is usually because the SL text was not properly understood. Consequently, due to time restraints, candidates try to write out their "summary" at breakneck speed so as to reproduce everything, their written presentation is usually no more than a quasi-legible scrawl, there is obviously no concision whatever and their written

language is usually atrocious as the candidates simply do not have the time to check over what they have written simply because they have written too much.

So, from this mass of scribbled, uncohesive and incoherent summaries who then are the potential interpreters? Clearly, the "best" students in that it is they who manage to write everything in good Italian against the clock. But are they necessarily going to be the best interpreters? Clearly, from the statistics quoted at the beginning of this article, 45% of these so-called potential interpreters, in one particular year, failed even to complete the course. Conversely, "bad" students often performed quite well in this particular test because if they failed to understand something they would omit it, thereby giving the impression of concision, and consequently with less to write they had more time at the end to check their work. And what about the almost 40% of failed candidates who then completed the course and who are now professional interpreters?

Clearly, something is going wrong with this test! And what is going wrong, I think, is that the testers presuppose that candidates know what is expected of them and can perform accordingly. Nothing could be further from the truth! Students simply do not know what is expected of them when they are asked to write a summary, for at school in Italy there is no longer any formal training in *précis* writing and at school they are also discouraged from expressing themselves briefly and clearly but rather are expected to tend towards the obscure and the flowery so as to show a mark of refinery and culture. During the first two years at the University, things are little better for very little work is done on trying to teach students concision and clarity. So, I ask, how can they be expected to be concise and clear on the day of the aptitude test which is probably the very first time they have ever had to perform such a task? This is simply a continuation of the iniquitous policy of looking for the natural interpreter from among a mass of bewildered would-be interpreters who desire nothing else than be taught how to do what is expected of them.

Things are little better for the second test we use in Trieste, namely sight translation. Although little real research has been done on the subject, sight translation is generally recognized as being one of the most useful exercises to be employed in the training of the conference interpreter, even though Danica Seleskovitch claims that it is "*more*

akin to translation than to interpretation"⁵. Be that as it may, sight translation is very much an integral part of the job of simultaneous interpretation, "*a regular feature of professional interpreters' every day activity*"⁶ and consequently ought to be taught on all conference interpreting courses. The training of this professional activity involves improving students' visual perception, transforming written language into the oral form and presenting the text with a smooth, pleasant delivery.

Translating on sight for students at the beginning of a conference interpretation course means something quite different to them. It means primarily what they were used to doing in ordinary translation classes before starting an interpreting course. Translating on sight for them means to be presented with an unseen text, usually to read it through once before translating it and then to go more or less at their own pace through the text stopping and starting, repeating and changing things and with no regard whatever to presentation or delivery. This has little or nothing to do with sight translation as the interpreter knows it involving the immediate perception of long meaningful chunks to be translated virtually simultaneously into the target language in a clear, confident manner without confusing the receptor with a whole series of pauses, hesitations, repetitions, and modifications.

So when students take the aptitude test, it is not in the least bit surprising that the vast majority do the sight translation component very badly in that they are simply doing what they have always done. The fact that unseen translation and sight translation are two very different things is generally not explained to students before they take the test and even if it is, it does not help very much as sight translation takes a great deal of practice before any reasonable degree of proficiency can be achieved. So surely there is little point in testing students for a discipline that they know nothing about, for which their previous instruction is certainly detrimental and thus for which they cannot be expected to show any degree of proficiency or even potential. It is rather like testing a car driver for a pilot's course by asking him to fly. Obviously he is going to crash the plane because driving and flying have very different principles

⁵ Seleskovitch D., 'Interpretation, a Psychological Approach to Translation' in *Translation: Applications and Research* (R.W. Brislin ed.), Gardner Press, New York, 1976, p.114.

⁶ Viezzi M., 'Sight Translation: an Experimental Analysis' in *Aspects of English: Miscellaneous Papers for English Teachers and Specialists* (J.M. Dodds ed.), Campanotto editore, Udine, 1989, p. 109.

and techniques. You simply cannot test a person's aptitude for something by asking him to perform that activity, for if he can do it already he does not need a course in the first place.

As for the general knowledge component of aptitude tests, I am most relieved to see that in the round table meeting on interpretation research held in Trieste last year it was generally recognized as being the most subjective and least scientific of all, as clearly stated by Laura Gran and Christopher Taylor in their closing statements: "*In the past, some considerable weight had always been put on assessing the student's 'world knowledge' (or background knowledge, 'schema', etc.) and this was called into question by some participants. One member of staff, in trying to ascertain where to assign the limits to a student's potential world knowledge, asked whether the examiners would require to know of a student (a) the location of Antwerp, (b) the location of Guinea Bissau and (c) the extent of the gross national product of Guinea Bissau for the past year. Rather embarrassingly, the teachers called upon to answer the question were not unanimous in their response, indicating that real parameters do not exist for determining just what a prospective interpreter should know. Therefore, it was agreed that it would be extremely useful to formulate some more identifiable, scientific limits to this question rather than to continue to rely on the intuition of the teachers involved.*"⁷ I was in fact the member of staff mentioned and I was asking a question specifically designed to elicit a variety of responses on the part of aptitude testers, this not to discredit testers but simply to show that we are frequently breaking the fundamental rule of testing, namely consensus on the part of testers as to what is to be expected of candidates. You simply cannot have tests based on the individual personality of the tester who thinks that young 20-year-olds should know the GNP of Guinea Bissau or be able to describe the flags of every country in the world or what you will. Test experts are adamant on this point: "*objectivity is considered (...) a pluridimensional concept. As has already been seen, (...) a dimension concerns the objectivity of the testers, that is to say results must be independent of the examiner's personality*"⁸.

Yet, there can be no doubt as to the importance

of general knowledge in the everyday activities of the interpreter. One may of course claim that our world is becoming increasingly specialized and that the generalist interpreter in the not too distant future is going to have to specialize. This may well be true one day but, as Nancy Schweda Nicholson pointed out in the last issue of *The Interpreters' Newsletter*⁹, well over 80% of interpreters in her survey stated that they were non-specialists and listed over thirty different areas of subject matter specialization, the most common being politics, economics, finance, law, medicine and computer science. Interestingly, however, all of the participants in the Schweda Nicholson survey¹⁰ were over the age of thirty and almost 40% were over 45, in other words all interpreters with years of interpreting experience and exposure to the kind of world knowledge that we expect from our students. And Brown and Yule are also quite correct when they state that "*the interpretation of discourse is based to a large extent on a simple principle of analogy with what we have experienced in the past. As adults, we are liable to possess quite substantial amounts of background experience and knowledge.*"¹¹ So can we expect the same amounts of background experience and knowledge from a 20-year-old, a 30-year-old and a 40-year-old?

Clearly, the answer is no. But, in that case, what exactly should we expect from a 20-year-old or conversely what would not be expected of a 20-year-old that perhaps would be expected of an older, more experienced interpreter? In other words, how exactly can we delimit a student's general knowledge as being average or excellent or whatever if not totally subjectively? What is more, I agree with Gran and Taylor¹² when they say that the distinction between static and temporary data does not really help. Indeed, variable information may be much more important than non-variable as was certainly the case in one aptitude test when the young Italian candidate knew rather little about Libya or Gheddafi (non-variable) but what was much more unacceptable from the testers' point of view was that she was totally unaware that only ten days earlier missiles had been launched towards Italian territory (variable). Clearly, this lack of

⁷ L. Gran & C.J. Taylor, 'Closing Statement' in *Aspects of Applied and Experimental Research on Conference Interpretation* (eds. L. Gran & C. Taylor), Campanotto editore, Udine, 1990, p. 239.

⁸ M. Forzi, 'Considerations on Psychometric Tests' in *Didactic Support Materials N°8*, ANCIFAP, Trieste, 1984, p. 5.

⁹ N. Schweda Nicholson, 'An Interpreter Survey: The Specialist/ Generalist Controversy II' in *The Interpreters' Newsletter N°2*, (eds. L. Gran & J. Dodds), SSLM, Univ. of Trieste, 1989, p. 41.

¹⁰ *Ibid.*, p. 38.

¹¹ G. Brown & G. Yule, *Discourse Analysis*, C.U.P., New York, 1983, pp. 233-234.

¹² L. Gran & C.J. Taylor, *op. cit.*, 1990, p.239.

constant updating can only be because the candidate did not read newspapers or, to be kind, had not read the previous 10 days' newspapers.

However, once again, even over the question of newspaper reading, how can we as testers test aptitude on the basis of candidates' knowledge acquired from newspapers when we are perfectly aware that Italians generally do not read newspapers (with the notable exception of the *Gazzetta dello Sport!*) and that there is virtually no attempt made either at school or at university prior to an interpretation course to have guided reading of the press over a period of years. In Italy, general knowledge, current affairs, politics and economics seem to be systematically ignored at secondary school level and at the University may only be touched upon within the structure of a language course. Then students hear a couple of weeks before the aptitude test that newspaper reading is the best way to pass the general knowledge component and so start reading dozens of newspapers in the vain hope that they can fill the gap before the day of the test. Clearly, this sort of cramming never produces effective results even for restricted specialized subjects let alone for a subject matter as vast and ill-defined as "general knowledge". At best, it only lets students know how little they understand current affairs, politics, economics, finance and the like and awareness of the extent of their lack of general knowledge may in many cases be a determining factor in causing serious demotivation on the part of the student right at the very moment when motivation should be at its highest.

Now, motivation and demotivation are very complex psychological issues of the utmost relevance to teachers from all disciplines and most certainly to teachers of interpretation. Motivation and demotivation most frequently determine the success or failure of a course and thus must have primary consideration in a teacher's mind at all times. They can also be determining factors in the success or failure of a test and should thus have primary consideration in the tester's mind at all times. In the first case, for example, a few years ago the aptitude test was instrumental in destroying the unity of a class in that those who failed the test sent "to Coventry" the successful candidates, "those teacher's pets who had passed the test". This had serious psychological repercussions on mainly those girls who had been judged to have high aptitude for interpreting, with the result that a few students with high potential at the beginning of the course dropped out well before its completion.

As regards motivation during the test itself,

psychologists identify a whole series of factors that could possibly interfere with optimal test conditions. Kausler¹³ notes that the very name of the test may in some cases negatively influence results. His study shows that better test performances were obtained if the test were not referred to as an "intelligence test" but rather a "test of perception". Clearly, words like "general knowledge", "aptitude" and "intelligence" do have very strong connotations of ignorance and stupidity if the test performance is below standard or if there is fear of it being so.

Forzi¹⁴ goes on to mention many other "disturbance" factors in testing which explains why the verification of test results is such a complicated matter. For example, he includes the following factors as all being able to disrupt test results and consequently course efficiency:

- 1) feelings of success or failure;
- 2) the influence of praise or disapproval;
- 3) anxiety and stress during the test;
- 4) the influence of the examiner;
- 5) the feeling of being observed;
- 6) the influence of a time limit on performance;
- 7) the attitude of candidates towards exams;
- 8) the intentional falsification of results.

These disturbance factors are highly relevant for the testing of future interpreters for the student may well be doing the test against his will or the result of the test may determine whether he/she changes over to the translators course or whether to leave the school altogether and find another profession. Generally speaking, continuous feelings of success improve performance while the opposite is true for continuous feelings of failure. However, this seems to be more true of females than of males; the former generally react well to praise and feelings of success whereas the latter react well to stress, disapproval and even hostility on the part of the examiner.¹⁵ Time limits for the tests and also constant observation on the part of the testers may also lead to a poor performance but only when candidates are unused to such exam conditions and familiarity with the tester usually leads to a better performance than with somebody unknown to the candidates. Therefore, a friendly tester who is known to the candidates is preferable to a hostile stranger, at least as far as the female subjects are concerned. Lastly, there is always that old trick of intentionally

¹³ Kausler D. H., 'A Study of the Relationship between Ego involvement and Learning' in *Journal of Psychology* N° 32, 1951, pp. 225-230.

¹⁴ M. Forzi, *op. cit.*, pp. 5-6.

¹⁵ *qv.* M. Forzi, *op. cit.*, p. 8.

falsifying answers, of intentionally performing badly on the part of the candidate either as a sign of a rather scornful attitude towards the test situation or as a very common defence mechanism.

Clearly the whole question of testing and aptitude in general is a very complex one and it is our responsibility as teachers not simply to discard at random anybody who does not happen to have that innate talent that characterized the great interpreters of the past. To do this, we must find answers to several questions that as yet have hardly even been posed - otherwise we are simply playing around with the futures of our younger generation on the basis of subjectivity and intuition rather than on objectivity and science. The questions are the following. 1) Are the preparatory courses providing students with adequate training so as to familiarize them with the aptitude test? Is shadowing a valid way of screening candidates? Do we really know what we are looking for in terms of translation competence, general knowledge, psychological traits and motivation when we talk of a "potential interpreter"? I personally feel that, if the truth be known, the answer to all these questions must be no and that there is still a great deal work and joint research to be done before we can answer in the affirmative.

As far as shadowing is concerned, we ourselves - in collaboration with neurophysiologists - can do a great deal more in terms of research into its validity as an aptitude testing device. As regards summary writing and sight translation, closer cooperation with language and translation teachers will be necessary in order to prepare students for these two components and to in some way avoid training students before an interpreting course in a way that would be harmful to their performance on that course and during the aptitude test. Equally, we ourselves should really put our heads together over the whole question of general knowledge to ascertain at least broadly what we expect of them so that this information can be given to students and other teachers well before the aptitude test itself. Lastly, over the various issues connected with the psychological traits of candidates, motivation and demotivation factors and the like, here we really shall need the help of the experts, occupational psychologists, neuropsychologists, psychometric testers and so on. What I am sure of, though, is that if we do not do this work, aptitude tests in their present form will continue to remain rather low on aptitude themselves.