

” TESTING IS TOO IMPORTANT TO BE LEFT TO TESTERS”

by J Charles Alderson

Profesor of Linguistics and English Language Education, Lancaster University

Plenary Address to the Third Annual Conference on Current Trends in English Language Testing, United Arab Emirates University, Al Ain and Zayed University, Dubai Campus, May 1999

In this paper, I largely refer to large-scale testing, at institution or national level, rather than individual classroom testing. Much of what I discuss relates to school-leaving examinations at secondary level, although it also applies to tertiary institutions and their assessment procedures.

The background to the paper is the fact that in many settings, language testing is the work of specialists. People are appointed to be testing and measurement supervisors, or assessment coordinators; special committees are constituted that concern themselves with test development and administration; testing offices or even examination centres are often separate from curriculum development centres, or in-service training departments. But I argue that this is inappropriate and even damaging: testing specialists cannot work alone, they depend on the expertise and co-operation of others, even though these others may not consider that testing is their business, and might even prefer to leave testing to testers.

1) WHY IS TESTING TOO IMPORTANT TO BE LEFT TO TESTERS?

Assessment and feedback are central to teaching. It is impossible to imagine teaching without assessing how well learners are doing, or without giving feedback on performance. Learners need to know how well they are doing, formally or informally

It is important that tests and examinations be as close to the curriculum and to classroom teaching as possible. Achievement, progress and placement tests need to relate to curricula in order to be valid (this is less true for proficiency tests, although even with proficiency tests one hopes that there is no major mismatch between what is taught and what is measured). It is also hoped that such tests measure what and how teachers are actually teaching, on the assumption that they are teaching "properly". Curricular validity relates to how good teachers teach. If there is no correspondence, then the test does not reflect 'best practice'.

It is, however, equally important that teaching be close to tests and examinations. One hopes that the way teachers teach will bear a reasonably close relationship to the tests. We have to distinguish means and ends, of course. One may expect learners to learn dialogues by rote in class, to improve their confidence and fluency, but not expect them to regurgitate rote-learned materials in a test. However, if the test is good, it will reflect good practice and one would expect that the way teachers teach towards it would not be substantially different from the way they normally teach (this is the basis for the often-heard claim that the Cambridge tests are content valid: teachers do not have to change what they do, and test preparation books are just good textbooks).

Tests and examinations are often used as instruments of quality control, especially when the measurement of outcomes is as important as the measurement of inputs and resources. Thus if tests are inappropriate because they are remote from what teachers do or ought to do, a false picture will be gained of the quality of education.

Tests have impact on teachers, learners, textbooks, and so on. This impact is complex but often deliberate: tests are often used as levers for change. Whilst this impact is evident in high-stakes tests like TOEFL, or university entrance tests, what appear to be relatively low stakes may be perceived as important by students: end-of-year exams, placement tests, and even progress tests can impact on a learner's self-image or the views of their peers. Tests that are remote from teachers and learners are more likely to have negative impact than those that involve them.

But the impact of tests will be minimal if teachers are not involved in testing. Teachers mediate tests: they usually prepare students for them in some way or other. If teachers are not involved in the testing in some way, minimally in terms of understanding the nature and philosophy of the test, maximally in terms of construction, piloting, marking, reporting, then their test preparation practices may be inappropriate, and less effective than desired.

In sum, the validity of achievement tests crucially relates to the extent to which they reflect the curriculum and classroom teaching, and the extent to which they are within the reach (ability level) of the students. Testers alone cannot determine this: they need insights from teachers - educated, insightful teachers.

2) WHY IS TESTING LEFT TO TESTERS?

Many teachers simply believe that testing is about statistics and numbers, its jargon is unfamiliar and forbidding, the measurement culture is different from the humane culture of a liberal arts education: it is felt to be alien, and threatening to a teacher's self respect and professional knowledge.

It may be because the field is indeed arcane: testing journals are not accessible to the average classroom teacher, but are aimed at specialists who are comfortable with statistics and theorising, and they rarely address practical classroom concerns. However, other publications, like Language Testing Update, the newsletter of the International Language Testing Association, is aimed more at teachers, although admittedly it tends to be read by those with a special interest in testing, rather than the "average classroom teacher".

It may be because the field has concentrated on high-stakes proficiency testing, like TOEFL and IELTS. This is understandable, since it is important to do high-stakes testing well, and academic testers focus on proficiency testing since examination boards can afford to finance research into reliability and validity. Interestingly, however, equally high stakes testing - national school-leaving examinations - are just as important but rarely attract funds to establish their validity and reliability.

It may be because teacher training in testing has been too short, too superficial, or irrelevant. In my experience, a lot of training in testing gives teachers a taster, but leaves them feeling it is a huge and somewhat alien field. It rarely deals with classroom testing and so-called alternative assessment, and concentrates on techniques that are more relevant to professionally designed and validated tests - the high-stakes proficiency tests.

The poor training, the forbidding nature of the field, its irrelevance to classroom testing and concerns close to teachers' daily lives, may combine to create insecurity and lack of confidence. It may also simply be the case

that whilst teachers cannot ignore teaching - they have to face their classes day in and day out - it may be easier not to have to worry about assessment. They can take published tests, the end-of-unit tests in their textbooks, or tests written by colleagues, instead of devoting previous time and energy to something they feel is neither central to their profession, nor something they feel able to do anyway. Testing is also something that is not normally recognised in the allocation of teaching-related duties by administrators. Developing expertise in testing is something that requires a lot of hands-on experience, as well as attendance at courses for which there may not be time.

A lack of time, a lack of money to pay for substitutions, or to reduce teacher workloads in order to allow them to devote time to contributing to test development, doubtless conspire against teacher involvement. In addition, however, there is the issue of trust: can teachers be trusted to write tests or to contribute to test design and development? Might they not cheat - leak the tests to their students? Might their students not thereby have an unfair advantage, or special preparation? Are teachers capable of writing good tests anyway? Why not just leave it to the professionals?

I suggest that a lack of trust in teachers' professionalism, their honesty, fairness, or ability contributes to a vicious circle in which teachers do not learn about testing because they are not involved in it and they are not involved in it because they are not felt to be sufficiently knowledgeable. Thus tests become something that teachers are not involved in and do not own. It is therefore perhaps understandable if they are tempted to assist their own students in performing better by whatever means seem possible, since the test has nothing to do with them, and they may well be judged by the test results anyway.

3) WHY DO TESTERS NEED TEACHERS?

This is the other side of the coin to section 1. If tests are to be appropriate to learners' levels of achievement, they need to be informed by the experience of those who see this development on a day-to-day basis. Often teachers have a rather gloomy idea of what students are capable of, but nevertheless their views are important, and they should be consulted about the difficulty and intelligibility of test tasks (although, of course, tests should be carefully pretested to check assumptions of difficulty against actual performance).

Teachers know what is delivered in classrooms, rather than what is intended by curricula and authorities. The curriculum does not tell all. Curricula are idealistic planning documents, rarely adjusted in the light of piloting or experience. Teachers usually know how much material they can get through in a year, and under what conditions, and they often have a good idea of how much and with what quality learners are typically able to manage. However, teacher estimates and views should always be supplemented by empirical trials of tests to see what students themselves do actually achieve.

Teachers have to implement testers' intentions in any case. If teachers cannot interpret what the test is intended to achieve, if they do not understand the item types, the instructions, the criteria for marking, the marking schemes, they will not be able to prepare students adequately for the test. The better they understand what is being tested, the more likely it is that their teaching will be in tune with the test's intentions. The less threatened they are by the test's consequences, the less likely they are to cheat, to help their students to cheat, or to engage in unprofessional or undesirable test preparation practices.

4) WHY DO TESTERS NEED POLICY MAKERS AND RESOURCE PROVIDERS?

There is a tendency in testing and teaching circles to believe that if testers and teachers are in agreement and harmony, all will be well in the assessment world. Nothing could be further from the truth. Crucial partners in all this are the bureaucrats, the administrators, the heads of department or the advisers to Ministers, who develop policies and implement them, and release or restrict resources

Bottom-up approaches to innovation are not enough on their own. All innovation needs resources and the support of policy, adequate infrastructure and a guarantee of teacher involvement (release time, compensation, etc). Whilst grassroots, teacher-based initiatives are fine and important, if those in authority and with responsibility are not in agreement, simply do not know what is going on on the ground, or do not or cannot release resources or provide adequate conditions, there will be frustration and failure. If teachers are to be involved in testing, they will need to be compensated, or their school compensated, release-time will need to be ensured, adequate reimbursement of expenses and possibly extra pay will all need to be considered and secured. Good tests also need adequate delivery mechanisms, adequate arrangements for external marking and monitoring of quality, adequate arrangements for registration of candidates, reporting of results, issuing of certificates, and guaranteeing the security and integrity of test papers: all too often innovations in test content and method have not been accompanied by the necessary infrastructure, for example an Examinations Centre, to ensure the quality of the whole assessment system and procedures.

Curricula are typically determined centrally, and it is important that examinations and curriculum be in harmony rather than in conflict or contradiction. Most policy makers and their colleagues understand the need for curricula: curricula are planning documents. However, the crucial link between curricula and examinations is often ignored, forgotten or neglected. The effect is all too often dissonance between the curriculum on the one hand and the examination system on the other. This is often made much worse if one institution is responsible for the curriculum, and another institution is responsible for the examinations: such institutions frequently guard their territory jealously and do not communicate with each other. The Examination Section or Centre sees itself as the guardian of standards, and the Curriculum Development Authority sees itself as innovatory, in contrast to the conservatives in Examinations. Even in so-called unitary curriculum and assessment authorities, internal rivalries and resistances are common. Decision makers and those who implement decisions need to be aware of the importance of integration, of common goals and objectives, and they need to monitor to ensure such integration is achieved. It cannot be assumed.

If policy makers are not involved in, and do not understand the nature of, testing, there is a danger of a contradiction between the aims of examination reform and the manner in which such reform is implemented. Decision-makers are apt to see tests and examinations as a means of bringing about change: tests as levers for change. Administrators are apt to overlook, or simply not to be aware of, the fact that good examinations need infrastructures to develop and maintain them. One cannot just give some teacher or tester the responsibility of designing a test, and then pass it on to all the schools for teachers to administer and mark. Reforming examinations necessarily involves inspecting all aspects of the examination process, from specifications to the issuing of results and certificates, to see if they are adequate for the new tasks. Typically, new administrative procedures are needed when the content and methods of examinations are revised.

Laws and legal regulations are needed to provide a framework for assessment: litigation is increasingly common, and test administration and the issuing of results need to take place with a clear understanding of the rights and responsibilities of all parties. In many parts of the world, students and their parents are increasingly inclined to complain to the courts if their test results are not as expected - ie good (they do not complain if they are better than expected!). This means that those responsible for tests and examinations need a legal framework for their work, to govern codes of practice, what is fair treatment, how cheating is

defined, what students can and cannot expect, how they should behave during examinations, what their rights and obligations are. Testers cannot be expected to draw up such regulations themselves, although it is crucial that they contribute to the drafting of regulations and laws, since they will affect their professional work.

5) HOW TO INVOLVE POLICY MAKERS AND RESOURCE PROVIDERS?

Dialogue and mutual understanding and respect are essential. In most situations I have been involved in, teachers, teacher trainers and Ministry officials treat each other with mutual dislike and mistrust. The universities which produce the teachers complain that the Ministry does not understand teaching and teachers' needs, nor the need for adequate teacher-training programmes. Ministry officials often see universities as ivory towers unconnected to the real world of the classroom, and regard the teacher training they provide as inadequate. In many settings, universities have autonomy from Government "interference" and jealously guard this. Ministry officials think that teachers and university people believe they are a special case, and that rules and regulations should not apply to them. University people, trainers and teachers are often seen as pressure groups with their own self-interest in mind, and in such conditions of suspicion, collaboration is difficult if not impossible.

Indeed, it is unusual for Ministry officials and teachers / trainers ever to meet in the same forum. Yet dialogue and communication, understanding of the pressures which the other side experiences, and mutual respect for each other's concerns are essential. Above all there should be, within reason, regular meetings to discuss policy and its implementation and the legal framework that governs education and assessment. This rarely happens.

Decision and policy makers are concerned with more than just one language: English may be the most widespread and important, but other languages are also taught and tested in most settings. It would be invidious to have one set of rules and practices for one language, and a quite different set for other languages. As far as possible, respecting, of course, each language's particularities and teaching traditions, innovations in formal assessment procedures are more likely to be accepted if they are common to all languages. It would be odd, for example, for English to have an oral component which is externally assessed by trained examiners, whilst French or Spanish did not have an oral component, or allowed untrained teachers to mark the students' performance.

Similarly, policy makers are inclined not to see languages as a special case, and it is easier to institute reform if there is cooperation across the curriculum. Obviously, at least to language testers, languages are indeed a special case. The nature of language learning is quite different from the learning of history or science, in that it is both skill and knowledge learning, which takes time and, since language is a vehicle for our personality and culture, language learning affects all parts of the individual. Thus language assessment, especially performance-based assessment, is different. And language testing is the one branch of educational measurement and assessment which has developed its own discipline, its own philosophy, procedures and research and development traditions. It is also the only discipline which is truly international in the availability of many examinations. However, it makes life for administrators easier if language testing is in line with general educational assessment, as they are then more likely to cooperate and to understand the needs of assessment. If language testing is in line with other curricular subjects, then a special case does not need to be made for different procedures, special resources or logistics are not needed, and so on.

Having said this, it is also the case that language testing has at times managed to pioneer innovations in assessment that the other curricular subjects have then been able to follow. This has happened, for example,

in the Baltic States, where first English, then German and French, introduced radical changes in test methods and content which implied major changes to the infrastructure, the creation of exam centres, and the development of a different way of reporting results, which have since been copied across the curriculum.

It is also important to stress that tests that are believable and believed, that have currency, are relevant and have benign impact, are not produced without effort and quality control.

Administrators are not necessarily experts in assessment, and for many lay people, examinations are automatically "valid" and have authority. However, in reality, many school examinations, and especially school-leaving examinations, are rather poor: they are produced without specifications, quality control is minimal if it exists at all, control over marking is weak, and often results vary from school to school without reason. Such examinations are worthless. Examinations that are not recognised by employers or universities have limited value. Administrators and policy makers need to be told what the currency of their examinations is, and they need to be persuaded that educational systems should only issue certificates based on good, relevant examinations that have a positive impact on teaching and learning.

Policy makers and administrators are necessarily concerned about the cost of innovations, especially when it may appear that the current situation whilst not perfect is not disastrous. But they need to be educated to understand that all examinations cost money: even the worst ones. Bad examinations, however, waste money, since either the results are not believed, or they have an effect on teaching and learners that is undesirable and which wastes time and talent. Whilst it is true that good, communicative, relevant testing takes resources (oral tests are more expensive than multiple-choice written tests, at least to administer) their impact is potentially hugely beneficial. Even if good exams cost money, it is money well spent if it results in better teaching and learning or more valid and reliable assessment.

One common objection to change is to say that it is impractical, that it will not work 'here', that it all takes too much time and effort. It is extremely helpful, in my experience, to cite examples from other countries or institutions that have worked. By referring to innovations in the Baltics, for example, we were able in Hungary to persuade people that things are also possible in Hungary (the Hungarians do not like to think that they cannot do something that smaller countries like Estonia can do). Of course, it helps if those examples can be seen to be clearly related, or in a similar setting, to your own setting. But in my experience, the difference between settings can be greatly exaggerated. I remember once giving a talk in Egypt in which I described a nameless country, its examination conditions, and the educational context. The audience all believed that I was talking about Egypt. I wasn't: I was describing a quite different country: Sri Lanka.

It also helps if it can be arranged for policy-makers and implementers to visit the country or institution where the innovation has taken place, and study it at first hand. They will probably be motivated by the trip, and they are much more likely to be sympathetic if they can see things working for themselves.

Inevitably, Governments do not last for ever, and so politicians want to be able to show results quickly, rather than in the long-term. Education, however, is a slow process, and it takes time for innovations to bed down and yield results. Policy makers need to be encouraged to take a long-term view, and need to be patient in waiting for the desired results. The temptation to get quick results should be resisted. Policy makers often think that top-down solutions like examination reform will bring quick results. They won't. It takes time for teachers to understand the new examinations, and to learn new ways of teaching, in order to produce the results that are desired. Examination change must be accompanied by changes in thinking, and changing thinking takes time. Quick fixes do not work.

6) CURRENT REALITIES

The facts as established by research show that teachers do let exams dictate what they teach and to some extent how they teach. So far, though, research has not clearly established why they do what they do. Why don't teachers either resist examinations completely, or also change their methods of teaching?

Common-sense and the washback hypotheses (Alderson and Wall, 1993) suggest that this has to do with the importance of the exam.

One major study of the impact of testing, in Arizona (Smith, 1991), claimed to show schools dominated by worries about high stakes testing: tests distorted the curriculum and the teaching, caused anxiety and stress, etc. Yet this conclusion ignored the fact that tests were being used in the schools studied on a daily basis to place students, to diagnose problems, and to help refer students for assistance. None of this was questioned or criticised by teachers, or by the researchers: the test was seen as being there to help. Only when tests were tied to budgets and resources, to school and teacher reputation, to parental opinions, did they appear to become salient to teachers and even to the observer, and then became the object of criticism.

Conventional wisdom has it that teachers are more influenced by high-stakes tests. Actually the Arizona study also reports some teachers (admittedly a small minority) refusing to bend to that pressure, and insisting on doing what they thought was professionally responsible. So not all teachers are influenced even by high-stakes tests, and I think we need to investigate why teachers do bend: exactly what is their rationale for compromising professional standards and practice, if that is what they do?

If they change their methods, as the TOEFL study (Alderson and Hamp-Lyons, 1996) shows, why do they do this? One suggestion is that they do not do this consciously. Because, like most people, they want an easy life, they focus on test content, and don't think about methods for test preparation, unlike in their regular teaching. Test preparation textbooks exist (often directed at students, not teachers), and guidance to teachers on how to use them is very rare. But in many cases, as we saw in the TOEFL study, teaching methods do change, as classes become more teacher-dominated, traditional, less student-sensitive, more metalinguistic, more chalk and talk. No research that I know of has yet challenged teachers on this, although in the TOEFL study we did ask one teacher whether TOEFL could be taught communicatively and his response was "*I never thought of that*". We also asked teachers to complete the sentence: "*What I love about teaching TOEFL is.....*" and a common response was "*I don't have to plan lessons, and I don't have to give feedback*". Imagine: the two most important things teachers do - thinking about what and how to teach, and giving students feedback on what they have done - and TOEFL teachers don't think about either!

Teachers are busy, underpaid, have second or third jobs or other interests, or think that their real concern is regular teaching, and teaching for exams is somehow unprofessional, or cheating, and *coaching* is a dirty word. In addition, teacher training never focuses upon exam preparation. In fact, in the Sri Lankan study (Wall and Alderson, 1993) we discovered that those responsible for teacher training did not even know what the new exams were like, and I suspect that this is true elsewhere: universities and colleges can be rather isolated from the real world of primary and secondary classrooms and their day-to-day concerns, including what textbooks are used, and what examinations and tests used daily actually look like. I shall come back to the implications of this for teacher training below.

The real issue is what do teachers do and why do they do it? The fact is that at present this is a mystery, and we need to explore this: not just examining how teachers prepare students for examinations, but also how they behave during examinations. On a recent visit to a project in the Baltics States, where new English examinations had been introduced, I was shocked to discover how much cheating occurred, and how teachers

condoned it or just ignored its happening in front of their eyes as they administered the examinations. And indeed I have written about my own experience of cheating by my teacher when I was a secondary school student (Alderson, 1998).

In Hungary, in the school-leaving examinations, most students get the highest mark possible (5 on a five-point scale). English teachers tell me that if they give lower marks (less than perfect), their colleagues will conclude that English teaching is poor! In the oral exams, the examining panels examine each student in all school subjects and as a result the panels often include people who do not speak English. It has been reported to me that the Panel Chair may think a Reading-Aloud performance was fluent - it may have been rote-learned - and pressures the English teacher to give a 5. Many teachers report increasing the marks given on pressure from the Chair: they excuse this by saying it benefits the students! They thus feel under pressure to give higher marks, at least during the oral examination, than their pupils deserve, and they feel that their classroom assessments, free from external pressures like these are more valid and reliable! But why do they do this at all?

In Hungary, whilst planning the examination reform, and wanting to identify possible sources of opposition to the innovations, we conducted a Baseline Study, including a survey of teachers' attitudes (Nikolov, 1999). What we found was sobering, and cause for concern. Teachers were agreed that the existing examinations were worthless. They are too easy, they have no currency, they vary from school to school and teacher to teacher, virtually nobody fails, and the examination results are therefore not accepted for university entrance or employment. Yet teachers like the examination precisely because everybody passes. Students, they think, are more motivated to prepare for the exam because they will pass, and teachers feel, bizarrely, that good results reflect on them. Or at least, they say that an examination where some students would fail would reflect badly on their teaching. Thus any new examination, which must fail some students in some sense or it will also be worthless, must face this dilemma. And teachers need to think through their attitude and its implications for standards, for teaching, and indeed for motivation and for the value of the new examination they unanimously say they want.

Are high-stakes tests seen as being somebody else's responsibility, and therefore no ownership is felt by teachers? This is speculation, but it needs investigation if we are to understand why teachers do what they do, and change that behaviour. In Central Europe it is quite possible that cheating is a habit born of resistance to Authority, a habit acquired over many years of Communist dominance of behaviour and thinking, and resistance to it. If it is cultural, it will be deeply ingrained, and hard to change.

In a survey I conducted in Hungary, about which I report below (see Alderson, Nagy and Oveges, 2000), a very clear factor that emerged was cynicism on the part of teachers towards changes in policy and in governments. There was a tendency to believe that a new government would just reverse any changes made or attempted by the previous government. Many teachers have experience of very slow or a total lack of implementation of change, and this in turn leads to distrust. Teachers fear that the exam reform will not succeed because of changes in policy, because they have experienced conservative approaches to innovation and change on the part of policy makers and implementers, and they have become cynical and doubtful whether anything will finally happen. So they are not motivated to change habits, if the exams will not change anyway.

7) HOW TO INVOLVE TEACHERS?

I have discussed the need for teachers to get involved, and the consequences of their not getting involved, in testing, and have overviewed some aspects of what I consider to be current realities with respect to test preparation practice and attitudes to innovation and change in testing. I have argued that it is important to understand teachers' attitudes to and concerns about testing, if we are to involve them appropriately and fully. I would like briefly to illustrate what I mean by “involve” from my experience in Hungary, not because this project is perfect, but because it is the result of many discussions with colleagues in Hungary and of seeing how teachers have and have not been involved in other testing projects, or in testing more generally, in many countries.

The Examination Reform Project began informally in 1996, yet the new examinations are not due to be officially administered until 2005. However, already teachers are centrally involved in the reform, and every effort is being made to ensure that all teachers know about the new examinations and understand their rationale as well as think about what might be appropriate teaching to lead up to the new examinations. The Reform was partly inspired, certainly informed, by common views among teachers that the current exam was worthless, traditional, did not reflect changes in teaching and textbooks and needed rapid, urgent overhaul. These attitudes were then formally confirmed in the Baseline Study I referred to above (Nikolov, 1999), through questionnaires and focus group discussions with teachers. They have also been confirmed in presentations and discussions of the project at teacher conferences and seminars.

Teachers in secondary and tertiary education were then hired to begin the process of specification writing, and drafting possible test types and items. To assist them with this, they had training in testing in Hungary and abroad, most of them having attended two courses in the UK, one in Lancaster and one in Manchester, and three courses/ workshops in Hungary. They have written test items and seen these shredded and revised. They have been involved in piloting the draft test tasks on representative samples (it is crucial that item writers actually see how students respond to their tasks and do not just look at the statistical results). They have examined student performances on writing tasks, they have seen videos of oral performances, and they have developed draft criteria and scales for assessment. They have applied these scales to large numbers of scripts and videoed performances, and have revised them in the light of that experience, they have made proposals for the revision of the tasks as well as the scales, and they have developed training materials for raters of writing, and interlocutors and assessors of the oral tests.

Of course, this applies only to a limited number of teachers so far: 40 or so who are in effect being trained as testers, although they will remain classroom teachers also. New item writers are continuously being recruited and added to the pool of people who contribute to the piloting and banking of large numbers of test items. Involvement in item writing is clearly very beneficial for an understanding and ownership of the new tests.

In addition, more teachers are being involved in the piloting of of the examination: in April 1999, 23 schools and over a thousand pupils tried out the draft tasks and their teachers observed the piloting and gave their opinions of the test. In fact, schools clamoured to be allowed to pilot the tests, and many schools had to be turned down. It is hoped to pilot tests in those rejected schools in subsequent trials, as it is important not to alienate those who had volunteered and were disappointed the first time round.

More widely, teachers have been asked for their reactions to the tests as they develop, and to the various public documents. A publicity campaign about the progress of the reform is being developed, leading up to the official introduction in 2005. Project members are aware, however, of how much resistance can be whipped up by a hostile public and so they are already planning to explain at length why new exams are needed, what they will look like and what they will be worth. They intend regularly to publish articles in

newspapers and teacher newsletters about various aspects of the Project, and to give radio and TV interviews.

Already they have developed an in-service course in examination awareness for teachers, and the first version of this was piloted in Spring 1999 (even before the first pilots of the test were complete). The course focusses upon the new examination, its nature and its rationale. It illustrates and discusses the various tests and task types, it addresses what might be appropriate test preparation practice, and it relates this to good practice in language teaching generally. The emphasis is on how good teaching can be good test preparation, but it specifically focuses on the exam and how students can be prepared for it. And of course Project members also get valuable feedback on the test and task types from teachers attending the course and they can then take account of these comments in future test development.

The course takes sixty hours, over a two-month period: two days of intense contact, one month to try out ideas in class, and teachers then come back with the results for discussion in a further two days of workshops, a further one month to try out new ideas, and then a final day of round-up and certification. Teachers will get formal credit by the Government for doing this course, which will be accredited as a regular in-service course. The pilot training course was revised and tried out a second time in the autumn of 1999. It will shortly be available for anybody to teach, provided that they have first attended a trainer training course which is currently being developed. In this way, it is hoped to get wide coverage across the country.

In all this, what Project participants are trying to achieve is not just clever public relations: they are hoping that teachers will feel part of this project, that they will feel that they have contributed centrally to the examination reform, that they will understand what is being attempted and why, that they will feel comfortable with the changes and above all that they will feel that this is their project and their test, that this is not being imposed upon them but that they are centrally involved and treated with respect. Inevitably the Project will be limited in the number of people it can reach, which is why Project members have started as early as possible.

8) CONCLUSION: WHAT TEACHERS THINK OF THEIR INVOLVEMENT

I will conclude by letting teachers speak for themselves, by sharing with my readers opinions gathered in Hungary recently in a survey I conducted of the teachers directly involved in the Examination Reform Project, either as test writers, as members of the testing teams, or as teacher trainers (who also write items) responsible for developing the examination awareness course. In what follows, I reproduce verbatim the responses of one teacher who has written items and has been centrally involved in developing and teaching the examination awareness course. Her answers are typical of the majority of those I have received.

1. Why do you think teachers might want to get involved in the Examination Reform Project?

Because teachers are really fed up with the current system, and they feel the need for changes. At the same time, they do not like to be "left out", or simply to be told what to do or what not to do. If they are allowed (even better: asked!) to contribute to the changes, they will feel that the new system will be "theirs" in each and every sense. And an additional element of the whole process is that if a teacher gets involved in the Exam reform, he or she gains a lot personally, because taking part requires (and offers) professional development.

2. In what capacity do you think most teachers would like to be involved (item writing? administering tests? attending courses on test preparation, etc)?

In my opinion, most teachers would like to attend courses on test preparation, and of course they would like to administer tests. This includes training for oral examiners (interlocutor, assessor), markers. There are a lot of good teachers in this country who would like to do this gladly. I am sure a few of them would like to write items, but this is a much more difficult area.

3. Do you think teachers should get involved in this Project. If Yes, why? If NO, why not?

Definitely YES, and my answer to Question 1 partly gives my reasons for this. We will not be able to implement the changes without those people who actually will carry them out at the classroom and school level. It would be a mistake to leave them out.

4. What do you think might stop teachers getting involved in this Project?

- 1. Their "inferiority complex" (concerning their professional abilities)*
- 2. The school managements.(being involved in the Project means being away from school at times, etc.)*
- 3. The fact that teachers are overworked and earn very little (so they rather take a few private students than spend their precious time on the Project).*

5. What should teachers know about testing in order to be involved in a project like this?

They need to know at least as much as we did when we started to work with the Project, so they need a special training course (it does not necessarily have to be the Manchester module though, there are other possibilities, perhaps cheaper, and in Hungary ...).

6. Why did you get involved in this Project?

Because for many years I had been dissatisfied with the whole system of teaching English in Hungary. I wanted to change things in my everyday teaching practice, but I knew my knowledge and experience was not sufficient enough. That is why I applied for the Manchester testing module. I liked the things I learnt there very much, I started to produce videos and other in-service teacher training things, and I enjoyed them thoroughly. It became a kind of an "obsession". In October 1997 the British Council asked me and 5 other colleagues from that testing course to join the Project in Lancaster, and to work on the in-service teacher training aspect of the Project. Ever since I have felt this was the best professional decision in my career.

7. What benefits do you think you have got from your participation in the Project?

- 1. Professional development.*
- 2. Experience.*
- 3. Pleasure.*

8. What have been the biggest rewards - the most satisfying features - in your involvement?

When I first felt that my 1.5 years' work came true - after the first two days of the examination awareness course. And of course that was the first time when I really started to trust myself - I mean that I am really able to do such a thing.

9. What disadvantages have there been in your involvement?

A lot of time which should be devoted to my family is devoted now to the Project. And this is a hard decision, I often feel guilty. (On the other hand I have an extremely supportive family, and they are very proud of me. I think they would be disappointed if I suddenly stopped this.)

10. What are your hopes and fears for the future of the Examination Reform Project?

Hopes:

A much better system of teaching and testing English in Hungary in just a few years' time.

A good examination centre network throughout the country. (And perhaps a centre would need my experience

...

A good pre-service and in-service training framework throughout the country. (And perhaps I would be allowed to do in-service courses in the future ...

Fears:

Wrong government decisions will prevent all this.

Local councils and school managements will not help us.

I hope I have shown in this paper how important it is that testing not be left to testers, that other people should be involved in it, crucially administrators, decision makers and teachers. I hope also to have illustrated how it might be possible to do this, and the benefits of so doing.

References

Alderson, J C (1998) Testing and teaching: the dream and the reality. NovELTy, 5, 4

Alderson, J C and Hamp-Lyons, L (1996) TOEFL preparation courses: a study of washback. Language Testing, 13, 3

Alderson, J C, Nagy E and Oveges E (eds) (2000) English Language Education in Hungary, Part II: Examining Hungarian Learners' Achievements in English. Budapest: The British Council

Alderson, J C and Wall, D M (1993) Does washback exist? Applied Linguistics, 14,2

Nikolov, M (ed) (1999) English Language Education in Hungary. Budapest: The British Council

Smith, M L (1991) Put to the test: the effects of external testing on teachers. Educational Researcher 20, 5

Wall, D and Alderson, J C (1993) Examining washback: the Sri Lankan Impact Study. Language Testing, 10, 1