

QMS and The Setting of Minimum Standards: Issues of Contextualisation and Variation between The Testing Bodies.

Piet Van Avermaet - CnaVT

In the world of testing, "setting standards" is generally considered to be very important. However, as we all know, putting our standards into practice can be complicated.

From one point of view, setting standards is rather like "making rules" and, therefore, we have to be aware that power and influence come into play in this process. By taking their own norms as a point of reference, those who are in powerful positions are often able to make the rules and may exert undue pressure on others. However it could be argued that this need not be the case; as long as the other "partners" who are in the process of setting standards, have the opportunity to discuss the norms and the potential differences in existing norms, each of the stakeholders in the process can bring their own institutional, national, culture arguments, and a whole range of considerations and possible constraints can determine the outcome. (This links to post-modern views of cultural and ethical relativism as well as with Rawls and theories of justice and fairness). One outcome of a discussion of this kind could be that partner X adopts the norms of partner Y (or the other way around). Another outcome could be that both partners accept new, more neutral norms for both.

A third outcome of the discussion could be that the differences in norms remain, but that each partner accepts these differences as meeting the necessary criteria. Each partner would maintain his or her norms as being valid for their own particular situation. Indeed this outcome could have a positive effect on the future co-operation between the partners.

However, when one of the partners seeks to impose his norms on the others, and these norms then becomes the rule, the situation becomes much more demanding. In principle norms should be open to debate and set by consensus. Rules, however, may not be. A rule often takes on the status of something that is "set in stone" and is not up for discussion; once the rule has been made and has been established, it is often not felt to be appropriate to attack that rule.

This can, of course, be considered a subversive power mechanism - we often forget that rules were set by people who at some time in the past used their own position of power/authority to influence outcomes and used their own norms as a point of reference in setting those rules.

In ALTE we want to avoid a situation in which setting standards is created through these kinds of power mechanisms - our work on Quality Management Systems is part of our approach which seeks to prevent this happening.

There is a second reason why we have to be very cautious when setting standards. That is the attraction of those in power or who have the most prestige for the other members of the group:

a) If we suppose as a member of the group that we can gain (at least symbolically) by adapting to the rules that have been set out by the most powerful members of the group, there is a temptation to do so. We may try to bridge the gap between us and them - despite the

institutional constraints or different cultural contexts which may exist.

b) If, however, we suppose that we cannot meet the imposed standards or rules, and we believe that by trying to bridge the gap we may lose out (again possibly in a symbolic way), we may be inclined to praise and defend our own norms. Maintaining one's norms can be very good as we already said. However, if this is an irrational reaction against something one cannot achieve, the case is less obvious. In that case constructive discussion can be nipped in the bud.

This process of 'attraction of the powerful elite' has to be avoided. It is not because a Porsche or a BMW has a lot of prestige, that would make this type of car a suitable standard for drivers. For most people, the standard would be a reliable car which takes you safely from A to B and which is in proportion to financial possibilities and other constraints

Being a member of an organisation like ALTE could potentially enforce this mechanism of 'attraction of the powerful members'. But as a Code of Practice group within ALTE working on a Quality Management System, we have attempted to take these issues into consideration. As stated before, the differences between the ALTE members are large with respect to organisational, linguistic, educational and cultural contexts. For example, an international examination board like Cambridge ESOL has a long history, conducts exams all over the world and has many thousands of candidates a year. Other ALTE Members have limited numbers of candidates who take the exams within the country itself. There are also huge differences in knowledge and tradition with respect to all aspects of the examination cycle, including statistical and empirical issues like data gathering, data analysis, the equating of different examinations, and so on.

Despite these differences, all the members of ALTE share a commonly felt need for fairness in their examination systems and recognise that sound principles must underpin their work.

If it were to be the case that the cultural, logistic, psychometric, processing and other aspects of the larger institutions were to function as the only point of reference for the Quality Management System, such a system would become threatening for those members who know for themselves that they will not as yet be able to meet those standards. However, even for members that feel 'attracted' to the norms of the larger institutes and can meet them, adopting their norms might have negative effects were the specific characteristics of their own institution to be affected.

ALTE members would not wish to fall into the trap of believing that standards should be determined or imposed only by those that have symbolic power and that the others have to be 'attracted' by this and should try to achieve these standards at all costs. In order to avoid this, we have put a lot of emphasis on the step-by-step development of Quality Management Systems and the concept of continuous improvement. On the other hand, at the end of that process we are striving to ensure that minimum standards will be in place - based on the Code of Practice and a shared understanding of sound principles of good practice. These will be standards that have been selected and determined on the basis of in-depth discussions with all members, where every member's context has been taken into account. Of course, institutions that are able to and want to go beyond these minimum standards can do so and will strive for "best practice models". However, a quality-label can be achieved when the exams of an institution meet the minimum standards. The consequence is that every member, after a period of self-assessment and peer monitoring, should be able to meet these minimum standards and that individual members have the opportunity to go beyond that minimum level. In this way, the variation that exists among all the ALTE members, including the different cultural, national and

institutional contexts in which they each work, can be taken into account and can be acknowledged.

Some instances of this are as follows: one aspect of the Code of Practice for external exam providers concerns the roles of examiners and teachers. Good practice would suggest that a teacher of a candidate should not be that candidate's examiner. However, some smaller partners in ALTE occasionally have only a few candidates to examine in a country (e.g. in 2000/1, the Certificaat Nederlands als Vreemde Taal had only one candidate in Argentina) where there is only one qualified teacher. In that case it was inevitable that the teacher was also the examiner or the exam would never have taken place at all. If the minimum standard were to be that the examiner could never be the teacher, the CNaVT would never be able to reach that standard. Not because of the fact that it doesn't choose to meet it, but because of its own institutional constraints. If CnaVT, however, were to try to meet that standard persistently, the costs would not be relative to the benefits. This standard would not take into account the limitations of that partner. The minimum standard here can be that every member of ALTE has to strive for the test to be administered fairly by a person who is beyond all suspicion of bias or malpractice. There may be a variety of suitable checks which could be carried out to ensure this.

Another example concerns aspects of reliability, difficulty and discrimination - all considered to be important features of tests and which the test providers in ALTE should estimate for the components of their examinations (as noted in the Code of Practice). However, some countries do not have a tradition of systematically gathering item-level data in order to carry out statistical analyses - perhaps they should, but the current situation is different. Smaller members of ALTE do not have a psychometric section within their organisation to carry out this kind of work on a routine basis. This does not mean that these features are not recognised as important and to be accounted for in some way. A minimum standard for reliability should be that the reliability has to be estimated and reported in an appropriate way. How this is done may depend on the potential of the organisation to carry out data collection and analysis. If institutions have the possibility of doing more advanced analyses, this of course can only be encouraged.

Another instance might be that larger institutions often have the possibility of undertaking pre-testing and analysis every time they construct a new examination paper. They have enough staff and financial resources and there is a target group of candidates which is large enough to make up an appropriate sample of candidates to carry out the pre-testing. Smaller members often do not have these possibilities. A minimum standard cannot be established so that pre-testing is a precondition, for example, to ensure that differences in performance are related to the skills under assessment rather than irrelevant factors. A minimum standard in this case should be that adequate procedures have to be provided and described to ensure that differences in performance are related to the skills under assessment rather than irrelevant factors. This may be done, for example, at different stages of the examination cycle, starting with careful test design and task construction and then by carrying out appropriate post-exam analysis and grading procedures.

These are a few examples of how minimum standards can be set and at the same time the variation amongst the members can be respected.

The Quality Management System functions as a tool for members to enhance the quality of their examinations from the standpoint of fairness for the candidates. It can also function as a tool for discussions and debates amongst partners about the quality and the aspects of fairness of the different procedures and steps in the running of exams. And finally it can also be a tool to open

up discussions and negotiations with the funders of examination providers - it is often increased funding and other aspects of organisational change which lead to increased possibilities of enhancing quality and raising standards.

In conclusion, the COP working group now plays a central role within ALTE - Fairness and Quality Assurance is high on the ALTE agenda. However, we are always careful to respect the different contexts in which we all work and the various backgrounds from which every member comes.