

PRE-CONFERENCE WORKSHOP 2:

Standard Setting in Practice: How to cut the mustard?

Feedback Summary

<p>1. Did you receive all the necessary information (pre-course materials, timetable for the course, information about the venue, etc.) about the workshop prior to the event?</p>	<p>YES: 30 (88%) PARTIALLY: 4 (12%) NO: 0 (0%) MEAN¹: 0.94</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● <i>Couldn't open the web site</i> ● <i>I found the pre-conf. 'homework' very useful</i> ● <i>The homework did not reach me when it was first sent out</i> ● <i>For really busy (lazy) people a map from the hotel to the workshop would be helpful</i> ● <i>I didn't get the info about the hotels in Barcelona</i> ● <i>Reading material too late</i> ● <i>As a late applicant I received immediate answers to all my questions and clear guidelines as to what to do.</i> ● <i>Homework arrived late.</i> 	
<p>2. Did you find the training room and facilities satisfactory?</p>	<p>YES: 30 (88%) PARTIALLY: 4 (12%) NO: 0 (0%) MEAN: 0.94</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● <i>The long room with a long table was partly inadequate – some participants didn't see the screen properly.</i> ● <i>It was easy to see the whole group (big group).</i> ● <i>For certain tasks to be completed satisfactory, 2 rooms have been necessary (Very large group!)</i> ● <i>The facilities were excellent but the air conditioning sometimes distracted my attention.</i> ● <i>Excellent</i> ● <i>Satisfactory except air-conditioning problem (gets a bit too cold)</i> 	
<p>3. Did you like the format of the workshop as a whole?</p>	<p>YES: 33 (97%) PARTIALLY: 1 (3%) NO: 0 (0%) MEAN: 0.99</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● <i>Of course more time would have been even better ...but setting is perfect</i> ● <i>Very much!</i> ● <i>I think there was good balance between input and discussion.</i> ● <i>New comers had no opportunity to find out relevant info about both trainers & participants! (1) First things first; (2) Not too much, not too quickly</i> ● <i>I would have loved to go on ...</i> ● <i>Excellent atmosphere & interaction. Very interesting material. All stages well explained.</i> ● <i>Very much indeed</i> 	

¹ 1 – yes; 0.5 – partially; 0 - no

4. Do you consider appropriate the time limits for:	
4a. Familiarisation session	YES: 24 (70%) PARTIALLY: 5 (15%) NO: 5 (15%) MEAN: 0.78
4b. Test taking session	YES: 32 (94%) PARTIALLY: 2 (6%) NO: 0 (0%) MEAN: 0.97
4c. Training session	YES: 18 (53%) PARTIALLY: 13 (38%) NO: 3 (9%) MEAN: 0.72
4d. Judgment session	YES: 27 (79%) PARTIALLY: 7 (21%) NO: 0 (0%) MEAN: 0.90
4e. Setting cut-off scores	YES: 26 (77%) PARTIALLY: 8 (23%) NO: 0 (0%) MEAN: 0.88
4f. Internal validation	YES: 31 (91%) PARTIALLY: 3 (9%) NO: 0 (0%) MEAN: 0.96
4g. External validation	YES: 29 (85%) PARTIALLY: 4 (12%) NO: 1 (3%) MEAN: 0.91
Comments/recommendations: <ul style="list-style-type: none"> • <i>I would have appreciated more time for training and discussion internal valid.</i> • <i>Well, I tend to be a diesel machine – a bit slow ..., especially when statistics is concerned.</i> • <i>Obviously, I'd like to have the sessions longer but then it must take a week! It was pretty meaty!</i> • <i>More time for c); d) and e).</i> • <i>About 4a: As good as possible under circumstances – but really too short</i> • <i>More time for going through sample items is probably needed.</i> • <i>About 4c: maybe longer</i> • <i>For external validation, I needed more time as I didn't quite 'get it'.</i> • <i>Shorter presentations and more participants involvement; longer time for <XYZ>² and assessment</i> • <i>More time should be devoted to training CEF-level assignment of items.</i> • <i>Familiarization with CEF descriptors should be followed by looking at / working with benchmarked items.</i> 	

5. Do you consider the time limits for the discussions appropriate?	
	YES: 25 (74%) PARTIALLY: 8 (23%) NO: 1 (3%) MEAN: 0.85
Comments/recommendations: <ul style="list-style-type: none"> • <i>Discussions in large groups are always difficult / need much time; maybe, to some extent disc. in smaller groups would have worked well.</i> • <i>Discussions could always be longer, but only the last one about the final cut scores seemed too short.</i> • <i>Tricky, I know, but best not to let individual participants dominate (you did well on this).</i> • <i>It was especially interesting to follow experts' thoughts and considerations.</i> • <i>As there were so many people present, not everyone could probably have <i>their</i> say.</i> • <i>This is related to No. 4. The discussion took up a lot of time and added conflicting opinions. They were useful, but we needed more time then to guide us and focus our attention on the relevant <i>parts</i> of the CEF descriptors for the judgment task, and how to apply these features to perform the judgment task.</i> • <i>Having subgroups debates and then <XYZ> for agreement/disagreement on relevant issues would have helped.</i> • <i>I think anyone who wanted to contribute was given a chance and no one was forced to contribute.</i> • <i>We could have had discussions in small groups first and in the whole group.</i> 	

² <XYZ> – illegible handwriting

<p>6. Do you consider the time allocations of the workshop per day appropriate?</p>	<p>YES: 31 (91%) PARTIALLY: 3 (9%) NO: 0 (0%) MEAN: 0.96</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● Perfect time in <XYZ>, then sufficient <XYZ>! ● Perhaps 1.5 hours session would be better ● The training session could have been longer. ● The days ended too easily, with some material skipped due to the lack of time. ● Yes, in practice. The schedule given on the first day was a little heavy, ending at 5:30. Fortunately, we finished earlier. ● Even exhausting! 	
<p>7. Do you consider the length of the workshop appropriate?</p>	<p>YES: 29 (85%) PARTIALLY: 5 (15%) NO: 0 (0%) MEAN: 0.93</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● One week would have been perfect, but I am aware of the constrains and would like to thank all of you for your commitment and hard work - excellent! ● If longer time should be allowed to work on own data. Actually – EALTA could have courses on standard setting one or two weeks – where people could bring own data and work on own standard setting. ● We did many interesting things in just 2.5 days. Well done! ● As there was not time enough for proper familiarization it was somehow too short – but as we did not have to ‘really’ set cut scores it was acceptable anyway. ● A bit more time (1-2 hours?) needed. ● For the familiarization it was quite enough. ● Another half day would have allowed more time for training. ● I wouldn’t decrease the length! If possible, even longer (three days). 	
<p>8. Was this workshop as a whole useful for you ? Why?/Why not?</p>	<p>YES: 34 (100%) PARTIALLY: 0 (0%) NO: 0 (0%) MEAN: 1.00</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● Very helpful to get an overview in practice of different methods + applying them. Especially validation aspects and methods because it’s much clearer with hands-on experience. ● Especially the CEFR ‘aspect’ that I’m not that familiar with. ● Because my organization intends to tackle a standard setting process to relate its exams to CEFR in the near future. ● Will be looking at the standard setting from a much more professional perspective now! ● YES and Thank you ● Doing relevant research ● I have no previous experience with standard setting and this was an excellent introduction ● I understand the linking process better ● To tell truth, I’ve read about ‘cutting mustard’, but haven’t tried. ● It was very interesting to be part of this process with so many participants from different countries and with different background ● Extremely so!!! ● Very. It has given me the first hands on experience and the very practical knowledge on the <XYZ> requirements for validity. It has raised many, many questions too about what is really appropriate, but questions are good! ● Making comparisons with other SKA involved in workshop developing ● Because I have never done standard setting before. ● Very useful. Gives confidence to really discuss how this is done in our own country. ● Very much! ● It gave a practical step-by-step method I can use in my work. ● Elaborating my knowledge about standard setting. ● It is extremely useful to put into practice what one read about. A heterogeneous panel of experts is capable of generating highly ... discussions, the participants highlighting their own priorities & perspectives. => ● I could get the whole picture of the procedure. ● Lots of new ideas & broadening of horizons. The large size of the group was advantages for the statistical analysis & worked well. ● More discussions about relation between test-centered approaches (estimation of difficulty of test items) and nature of CEFR descriptors. ● Very useful and it provided much more information than in the manual. ● Very much 	

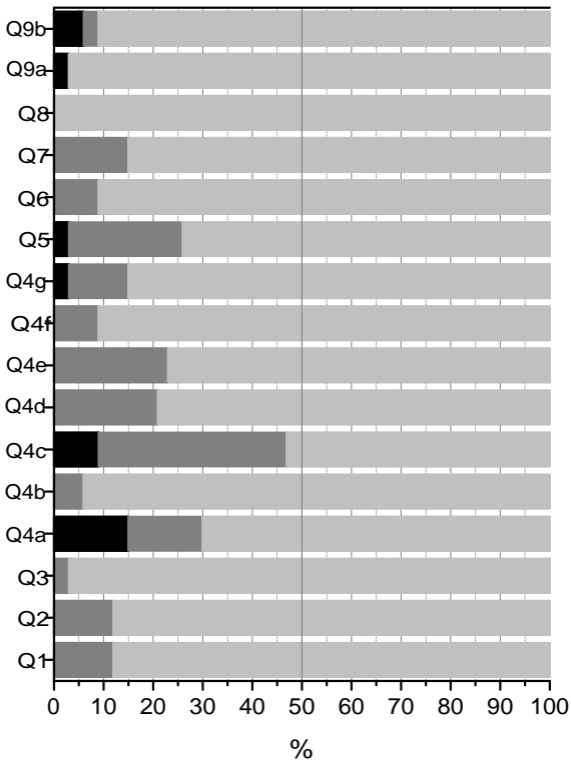
<p>9. a) Would you recommend this workshop to your colleagues / fellow professionals? Why?/Why not?</p>	<p>YES: 33 (97%) PARTIALLY: 0 (0%) NO: 1 (3%) MEAN: 0.97</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● Definitely – where else can you get first-hand expertise in such a profound way? ● Yes - There is a big difference between reading the literature and putting things into practice. ● Nothing similar done before in most (even high-stakes large scales!) contexts! ● It is useful to know some basic statistics. ● When you participate (not read literature), you realize the importance this task (in this case) ● I would recommend to anyone who needs to relate their exam to any scale. ● It is extremely practical, thorough and doesn't hide the problems (it highlights them)! ● <XYZ> ● Only to the professionals with some experience in statistics. ● Because people need to learn about testing standards. ● It is a pity that decision makers did not manage to come from my country. ● I could recommend this workshop for fellow professionals who already know something about standard setting and about CEFR and language testing. ● Very good in setting the basics! For the ones who don't have any experience. 	

<p>9. b) What about you, personally, would you like to participate in another workshop like this?</p>	<p>YES: 31 (91%) PARTIALLY: 1 (3%) NO: 2 (6%) MEAN: 0.93</p>
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● Yes – Always more to learn, deepen the expertise ● Yes – Wherever new methods etc. are presented and tried out ● Yes – For deeper studies ● Yes – This was a good start, but I feel there is much more for me to learn. ● No – I like having some unique experience! ● Yes – Maybe the next ... would ... bringing ... in the workshop, so that the same individual practice is carried out. ● Yes – Definitely! ● Partially – I don't know if another course would be useful. I know the basics now better, perhaps then more about statistics... 	

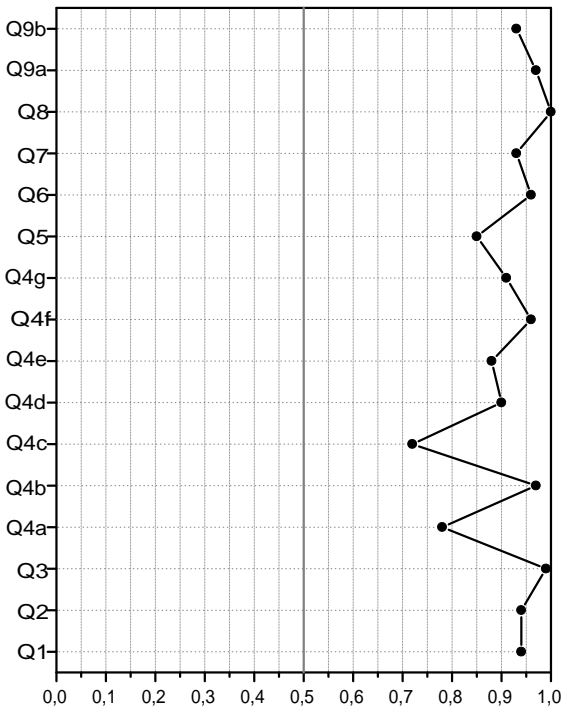
<p>10. What aspects of the workshop could be improved if any?</p>	
<p>Comments/recommendations:</p> <ul style="list-style-type: none"> ● Nothing really (except the time factor, but when we ever have enough of this?) ● If any at all, more discussion in smaller groups ● Maybe spend some more time on <XYZ> the stats part, time to clarify that we really catch what we were supposed to! ● Possibly more material to “take home” handouts ● More specific directions about how to do the homework (including guessing factor? Not use 100%? Etc.) ● Perhaps doing the Angoff test at home would have impacted on the quality of the judgments. ● Perhaps there would be a research oriented workshop on standard setting where we could go deeper into some specific issues (i.e. an alternative to people with some previous experience, i.e. for those who were in the workshop & who have worked on this home). ● You could run 2 workshops & let participants choose between reading/listening & speaking/writing. ● Thanks for your time & effort & enthusiasm & sharing your knowledge – it was (will be) useful! ● It is the first workshop like this for me; I was satisfied with all its parts. ● The first day was a bit rushed; otherwise everything was wonderful/excellent/superb! ● For the purposes of learning about the process and necessary conditions for good practice, it is good. ● Organization, <XYZ>, methodology, SKA <XYZ> in adult <XYZ>. ● A bit more theoretical explanation to contextualize what we are doing (achieved) and make sense of them ● It was perfect. ● Links with IRT ● More skills could be discussed ● Thank you! I think you have set high standards in facilitating, organizing a workshop and producing professional expertise. ● Some discussions in pairs/small groups, for example before the final decision of the cut-off points. ● More depth (hands-on) data analysis ● For me: more statistics => other procedures dealing with IRT data. ● Perhaps more time on familiarization and discussions. ● Thank you all of you! ● + a lot of flowers (🌸) & smileys (😊) 	

Feedback: Frequencies

■ No; ■ Partially; ■ Yes



Feedback: Averages



—●— Mean
(0 - no; 0.5 - partially; 1 - yes)