

Supplement and corrections to Tuning (Medicine) Validation Brochure

Prepared by Allan Cumming & Michael Ross, September 2007

Introduction

This document summarises the findings of the Tuning (Medicine) Validation Panel which met in Brussels on 22nd June 2007, and the actions subsequently undertaken by the MEDINE Tuning Task Force to address them. The Validation Panel findings consisted both of responses to the Tuning Validation Questions and of more general findings about the work undertaken and the Validation Brochure produced by the MEDINE Tuning Task Force.

Panel findings in response to the Tuning Validation Questions

1. Is the description of the subject area complete, clear, relevant: what do the Panel members think about it?

General description is appropriate, but has to reflect the actual situation in Europe more – therefore can remove parts referring to the two cycle model.

2. Degree profiles and occupations: how clear are they, is anything missing, etc.

a) Third cycle description, “**degree**” (p 10) requires further clarification. We recommend revision of the text to clearly differentiate between professional training and additional scientific training (e.g. in science doctorate programmes).

b) Since the information about graduates from the first cycle is still scarce, we are not able to validate their professional perspectives.

3. Relevance to subject specific competences: should certain competences be emphasised more, or less?

a) The Validation Group understands that the ranking was necessary in the process of Tuning. The results should be presented to medical schools as core competencies of graduates without the ranking figures to avoid misinterpretation, as they are irrelevant for curriculum planning. Also worth stressing that ranking is not related to the amount of time devoted to each area in a curriculum.

b) We are missing a link of this list with the clinical content as represented in the appendices. The Validation Group suggests that Appendices A and B [knowledge outcomes, clinical attachments and experiential learning] be incorporated into the text, as understanding the application of basic and clinical sciences are essential for

professional competence and the practice of medicine. [Panel highlighted that these were examples only however, and not complete list of knowledge / experience; .e.g. no knowledge of e.g. clinical presentations, diagnosis & management; or knowledge of general medicine / surgery]

c) We are missing health promotion for the individual patient (not for the population) as a core competency – should add this.

4. Do the Panel members think that the competences can (or should) be used in the process of professional recognition?

YES, but one needs good assessment tools to evaluate the competences.

5. Relevance of generic competences: should certain competences be emphasised more, or less?

Many of the generic competencies can also be identified as subject specific competences for medicine. Both generic and subject specific lists should be used by Faculties as a frame of reference.

6. In addition, the Panel members are asked what they think about the Tuning approach with regard to ECTS (workload) and of specific relevance for the health care sector: how can ECTS be used in a professional context? How does this help professional recognition?

It would be of advantage if ECTS and workload specifications were related to Learning Outcomes achieved.

7. Concerning teaching, learning and assessment, what do the Panel members think of the Tuning approach?

a) We endorse the idea of constructive alignment according to which formats of teaching, learning and assessment have to be matched to learning outcomes to be achieved.

b) These assessment tools should be scheduled within and justified by means of an integrated assessment strategy for evaluating student progress in the whole programme.

8. With regard to quality enhancement, what do Panel members think of the Tuning approach?

The Tuning approach, if appropriately implemented, can contribute to quality enhancement in medical education.

General findings of the Validation Panel

I) The Validation Group expresses great appreciation for the work of the TUNING (MEDINE) Task Force 1 and has high respect for the results presented in the report. The results are by and large consistent with comparable studies identifying competences in medicine.

II) The Validation Group endorses the interpretation of the Tuning remit as a consideration of the learning outcomes achieved by students by the end the undergraduate medical programme (up to six years).

III) The panel understands that further analyses of the data will take place. Conclusions should be reviewed in the light of these results. Such analysis could include:

- a) Bar graph comparison of level-1 outcomes for all respondents, German respondents, UK respondents and all except German and UK respondents.
- b) Distribution graph (and / or statistical representation) of the responses for each Level-1 Subject-specific outcome (i.e. how many E, VI, QI or NI).
- c) Statistical comparison of the distribution of responses by country (particularly looking at countries with large numbers of respondents)
- d) Look at median (vs average) responses when making sub-group comparison
- e) Specific questions - Do German responses skew results? (or other large numbers such as the UK). Do student responses skew results?
Is there a difference between northern and southern European responses?

IV) Would be useful to tag competences with 'Public health', 'Clinical', 'Social sciences', etc. if possible, as has been done with the Global Minimal Standards.

V) Some errors were identified in the Tuning (Medicine) document and should be corrected as follows:

- a) Change the order of Teaching, Learning and Assessment section so that the statement about alignment is in the first line.
- b) Need to highlight and better-define the relationship between subject-specific and generic competences (e.g. 'Application of knowledge in practice')
- c) Should include the removed level-2 competences on p31-34 for completeness (communicate orally, audit, stats, research...)

VI) Should add something about optional components, emphasizing that the core competences are “Necessary but not sufficient”. No percentage of the programme to have optional can be defined for this at present however.

VII) Worth stating that we regard all the generic, level-1 and level-2 competences as essential / core for medical graduates in Europe. Could also define more-detailed level-3 competences but would certainly not be core at present.

VIII) Missed opportunity for inter-professionalism at the Healthcare Validation meeting – should have exchanged results of four healthcare groups but no opportunity for this was created by the Tuning organisers.

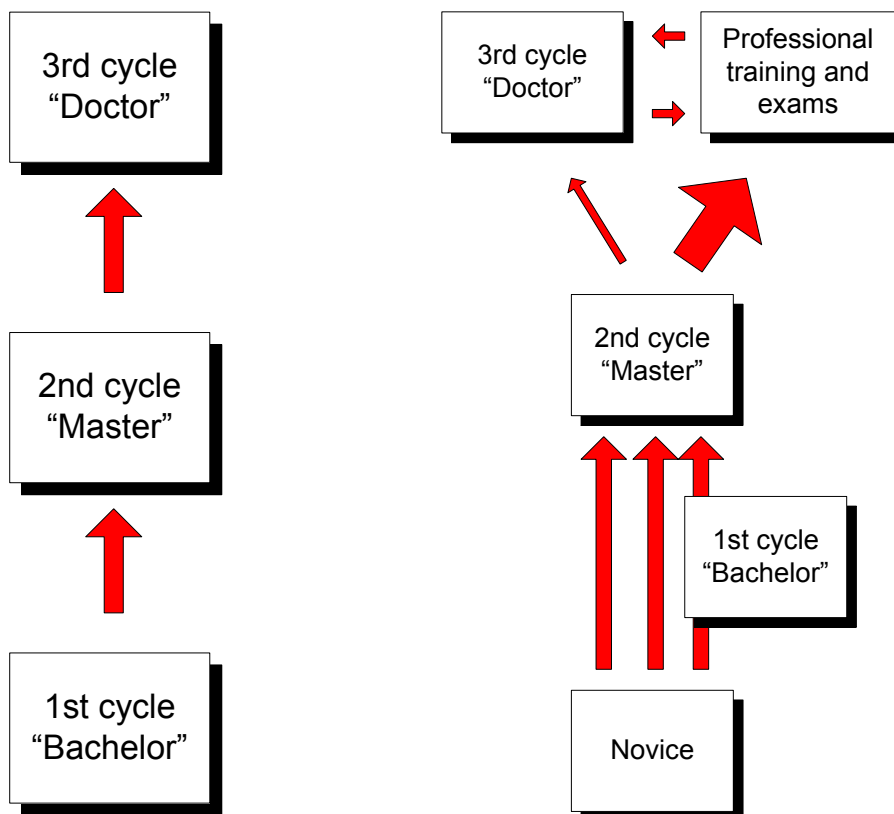
IX) May want to repeat the Tuning process in several years time to see if the outcomes change over time – perhaps it could be an ongoing process.

X) If funding for Tuning2 is granted, might want to work on what the Dublin descriptors mean for medical education; and also shape the competences identified in Tuning (Medicine) into a product which is easier to apply to curriculum development. Also need further work on defining core competences in Research, and core knowledge for clinical practice.

Tuning (Medicine) Task Force Response to Validation Panel

Tuning Validation Questions

1. We agree that the Bachelor of Medicine is very uncommon currently in Europe and that most undergraduate medical degrees are an integrated whole up to the level of the primary medical degree (Masters or 2nd Bologna Cycle). The brochure will be amended to ensure that there is no confusion about this.
2. We agree, and feel that diagrams are more useful in clarification of the cycles and how they relate to the current situation in medicine. The Bologna process is typically described as three logical cycles as represented in the diagram on the left below. The reality of medical education in Europe at present is much better represented by the more complex diagram on the right. Note that newly recruited students have been indicated as 'Novice', and the primary medical degree is considered to be Bologna 2nd cycle (Master). Undergraduate medical programmes are very diverse in structure and approaches, with only a few throughout Europe awarding students a 1st cycle (Bachelor) degree at an appropriate point during their undergraduate training. The vast majority of medical graduates (bold arrow) undertake rigorous further professional training and assessment in their chosen field (e.g. medical or surgical specialities, general practice, psychiatry) but do not undertake further study in a Higher Education context. Some medical graduates do undertake further HE training and research Doctoral studies, either directly after graduation or during / after further professional training.



3. Figures used for ranking will not be presented in the documentation for curriculum developers, and we wholly agree that these numbers do not have any relationship with the amount of time which is being devoted to the development of these competences within undergraduate medical curricula. We will highlight in the dissemination documents that knowledge outcomes & experiential learning / learning situations are integral to student development of the core competences. We acknowledge that these additional items are incomplete and were not part of the Tuning method so we feel they cannot be fully integrated into the final competency framework. We will add a competency of ‘Ability to undertake health promotion with patients’.
4. We agree that good assessment tools are required to evaluate these core competences, and have applied for funding to identify best practices in learning, teaching and assessment for each of these competences.
5. We have considered the generic competences to be an integral part of the core competences for medicine, and relate to ‘Personal Professional Development’ or ‘Professionalism’ in many undergraduate curricula. We will combine generic and subject-specific competences for the final document for dissemination to curriculum developers
6. The relationships between the competences and ECTS / workload specifications were beyond the scope of the current project. We agree that this would be important work and support an application for funding to undertake this.
7. We agree with these findings, and have applied for funding to identify best practices in learning, teaching and assessment for each of these competences.
8. We agree that the Tuning outcomes could fit in with a developing European dimension to quality assurance and accreditation in medical education. Reference has been made to the Tuning (medicine) outcomes in the new European specifications of the World Federation for Medical Education Global Standards document (available: www.wfme.org).

General findings of the Validation Panel

- I) The MEDINE Tuning Task Force would once again like to thank the Validation Panel for their work and contribution.
- II) We have considered in the current phase of Tuning (Medicine) the core learning outcomes / competences at the end of the undergraduate medical programme (primary medical degree, Bologna 2nd cycle). We intend to consider the situation of Bologna 1st and 3rd cycle degrees in Medicine as aspects of future research.
- III) Professional statistical help (Dr Margaret McDougall, Senior Lecturer in Medical Statistics, Department of Public Health, University of Edinburgh) was enlisted to undertake further statistical & subgroup analysis as suggested by the Validation Panel:

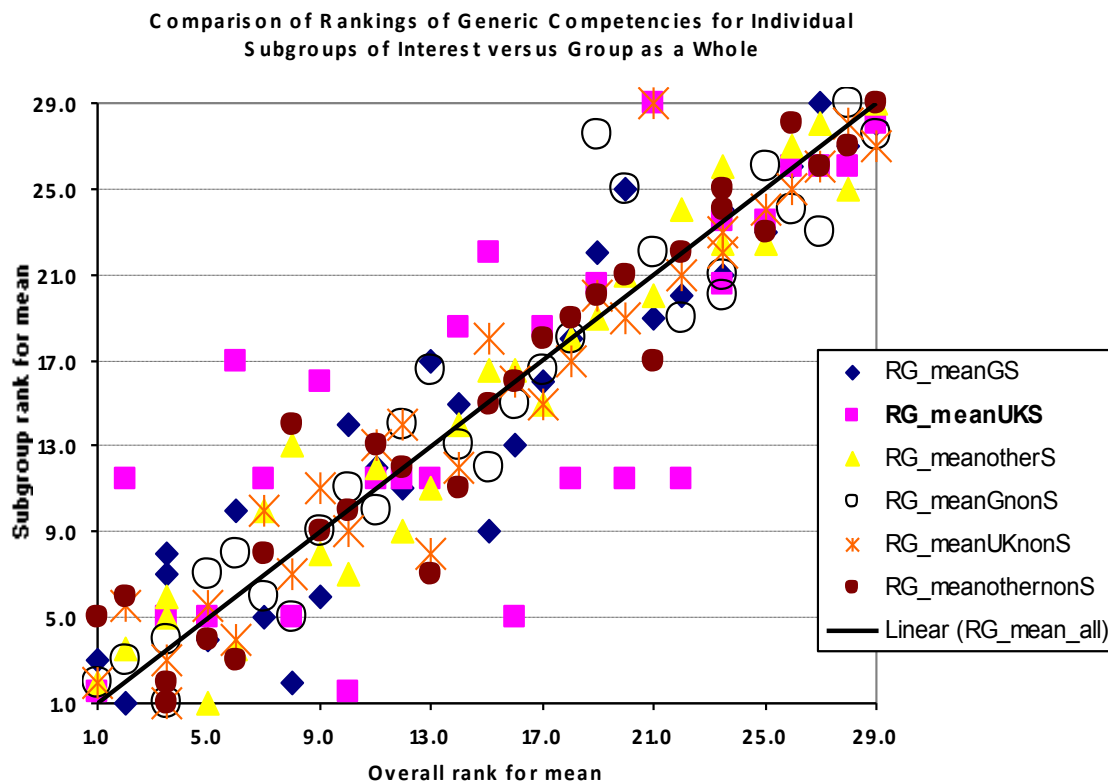
a,c,e) Subgroup analysis

i) Intra-class correlation (absolute agreement) obtained from comparison of subgroups with overall groups (note 'other' means 'neither UK nor German')

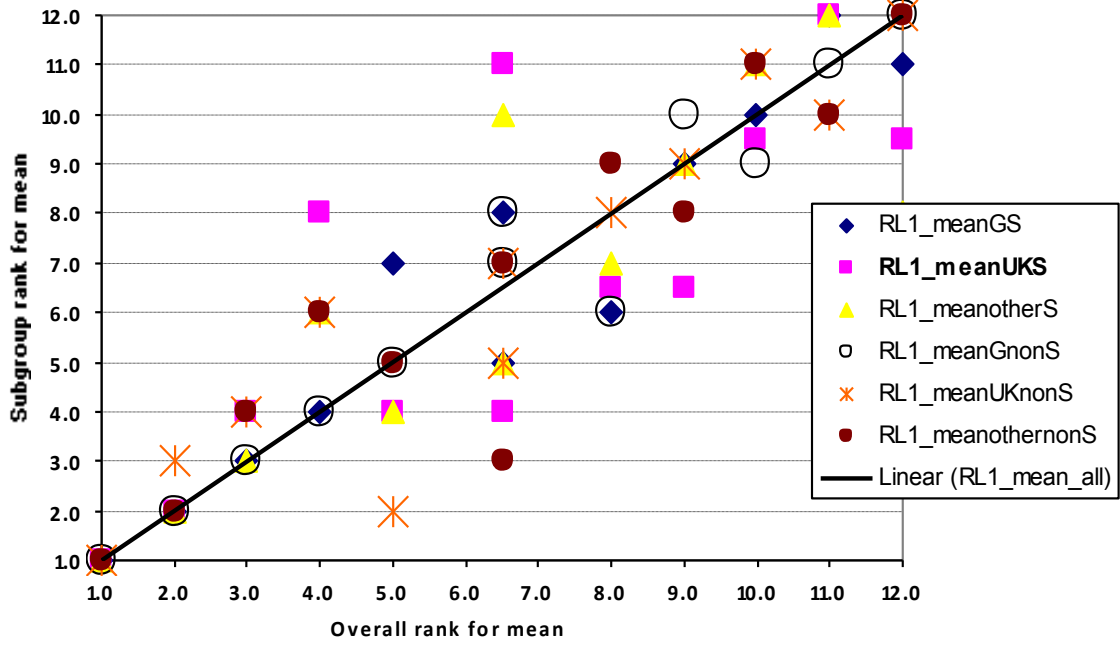
Type of competency	Rankings Compared	ICC	95% CI for ICC
Generic	German students vs entire group	0.942	(0.880, 0.972)
	UK students vs entire group	0.789	(0.596, 0.895)
	other students vs entire group	0.969	(0.936, 0.985)
	German non-students vs entire group	0.951	(0.898, 0.977)
	UK non-students vs entire group	0.960	(0.917, 0.981)
	other non-students vs entire group	0.960	(0.916, 0.981)

L2	German students vs entire group	0.889	(0.830, 0.929)
	UK students vs entire group	0.609	(0.442, 0.735)
	other students vs entire group	0.894	(0.837, 0.932)
	German non-students vs entire group	0.931	(0.893, 0.956)
	UK non-students vs entire group	0.880	(0.815, 0.922)
	other non-students vs entire group	0.957	(0.933, 0.973)

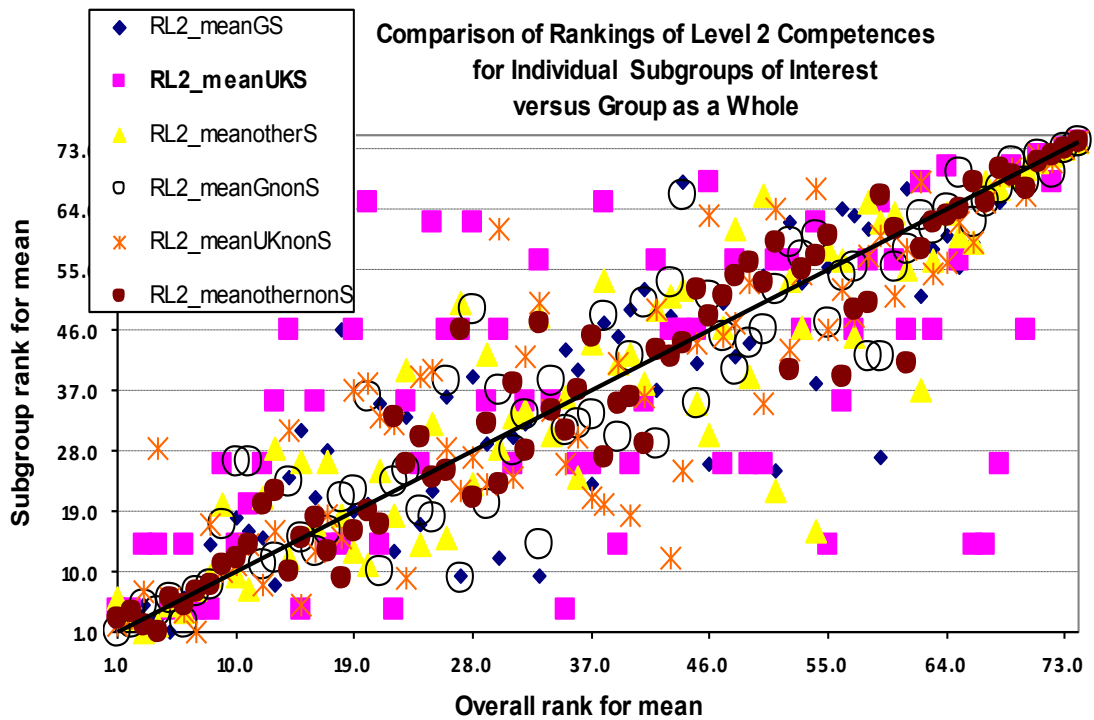
ii) Comparisons of rankings of competences for individual subgroups (coloured markers) vs all respondents (represented by the line). (RG – ranking generic; RL – ranking level (1 or 2); G German, UK – UK; S – students; NS - non-students. The identity of the competency which each marker represents for Generic and Level-1 competences can be found by looking-up the overall ranking (X-axis) in the tables on pages 24-5 and 29 of the Tuning Validation Brochure)



Comparison of Rankings of Level 1 Competences for Individual Subgroups of Interest versus Group as a Whole



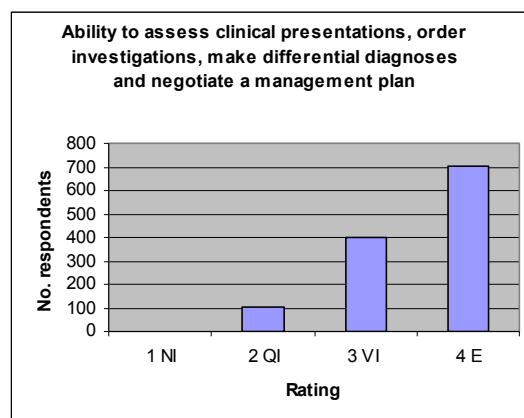
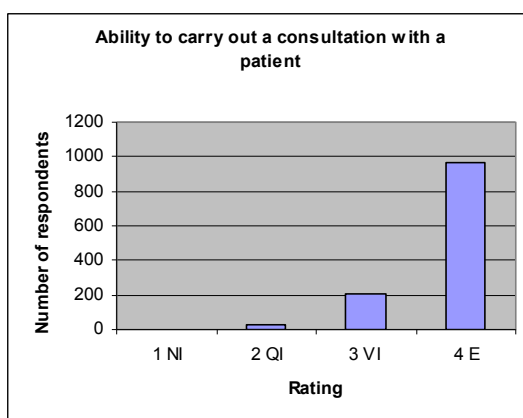
Comparison of Rankings of Level 2 Competences for Individual Subgroups of Interest versus Group as a Whole

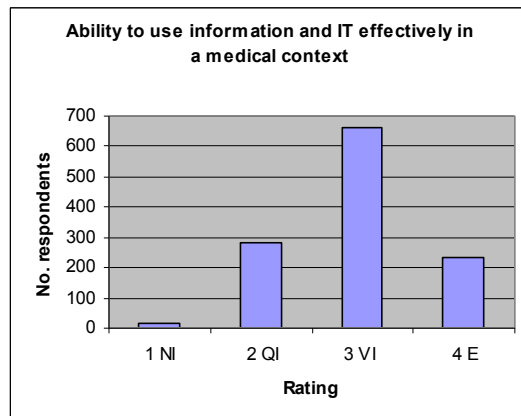
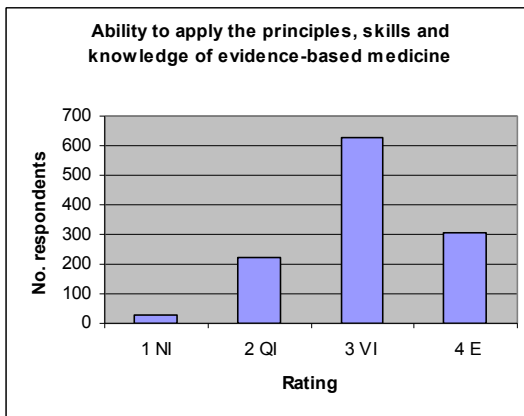
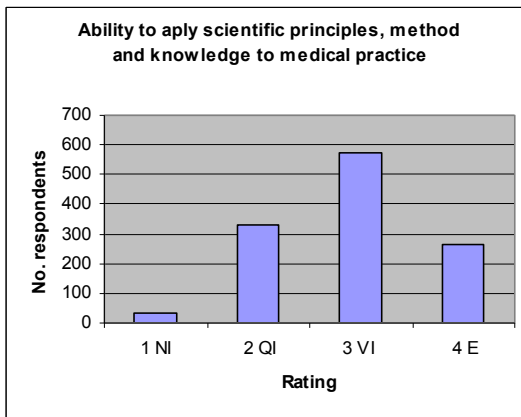
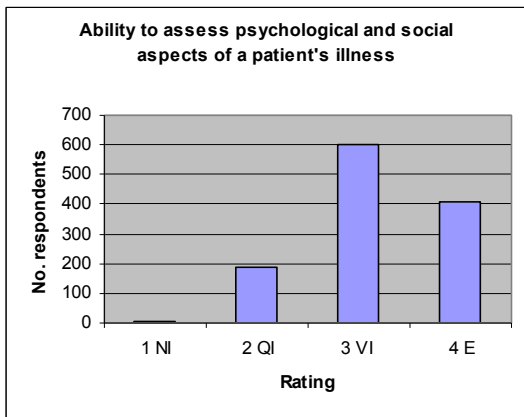
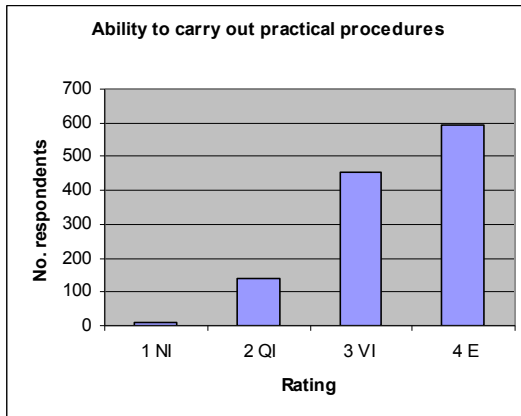
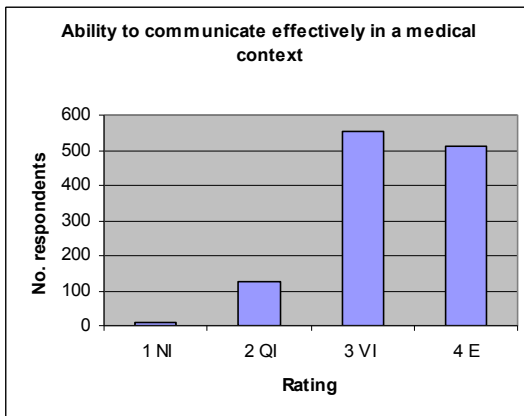
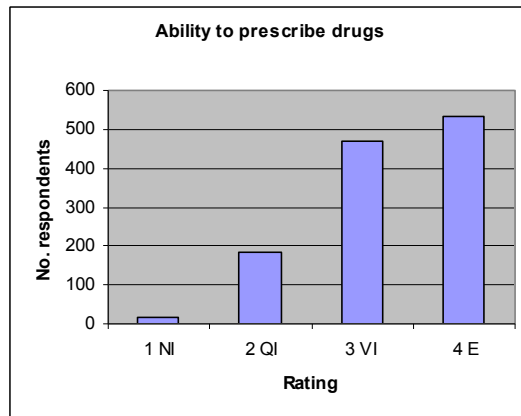
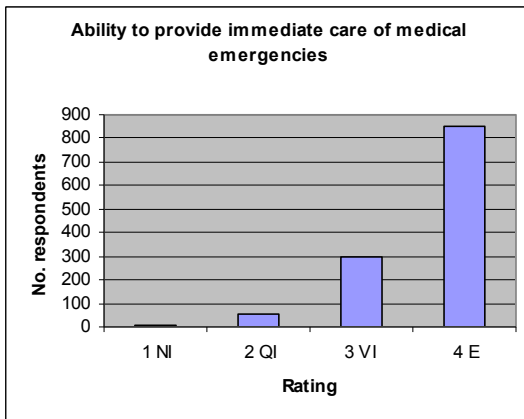


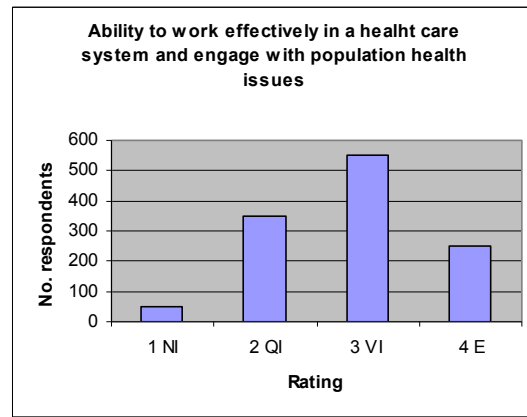
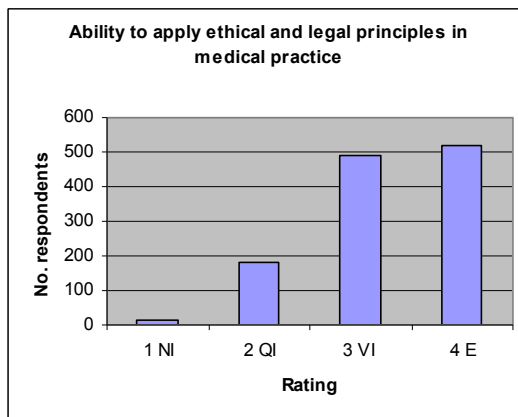
The intra-class correlation (ICC) coefficient for absolute agreement is intended to assess conformity to the 45 degree line highlighted in the scatter plots within the file Agreement charts. Note the ICC ranges from 0 to 1 (therefore we do not have agreement which is less than we would expect by chance), and that the ICC has been used with continuous data here (that is, with means). The idea of the ICCs is that of isolating individual subgroups within the complete dataset and seeing if this changes the overall ranking of competences. If the ICC is high in any one case then the groups which were being ignored within the subgroup could not have been having a strongly dominant influence on the final ranking. Thus, for example whilst the generic ICC for German students vs the entire group is very high (0.942), the generic ICC for other non-students vs the entire group is ALSO very high (0.960) - thus German students are not being more strongly represented than other non-students in the final ranking for generic competences. Similar arguments can be made with other parts of the table above. The ICCs are all very high (close to 0.9) with encouraging confidence intervals except in the case of the UK students. In the graphs this is seen by a very high degree of clustering around the line for all groups with only UK students (pink squares) as outliers. We also used an independent measure of consensus which allowed a higher degree of confidence in use of mean values for non-linear results. Results show that if UK, German, student or combined groups were removed however, all other groups still correlate very highly with the overall results, therefore there is no statistical indication to remove them. Thus, these large subgroups are not having a significant effect on the overall ranking of the Tuning core competences. One might infer, however, that UK students collectively seem to have a different opinion to the average in Europe, as taken on their own they would produce a different ranking. This would suggest that it would be a fertile ground for more qualitative inquiry into their perspective and opinions on these competences.

b) Distribution graphs of responses for each level-1 subject-specific outcome

Bar graphs are reproduced below which demonstrate the distribution of responses for each of the Level-1 Subject-Specific competences (in rank order from high to low). These indicate in graphic form the spread of responses and the degree of consensus in relation to each competence. As would be expected, the highest-ranked competences show the greatest degree of consensus.







d) Medians

We attempted to use medians, as advised, because of the potentially skewed data (as the mean may not identify outliers). However we could not use medians to rank the competences as required for Tuning as medians were not adequately discriminatory (e.g. only 3 values for median of 29 generic competences – 2.0, 3.0 or 4.0).

As an alternative, the Leik measure of ordinal consensus was used, which is based on counting the frequencies of cases where each of the options 1 to 4 were selected from the Likert scale for any given competency. The results are presented in tabular form, followed by graphical representation of the rankings for mean and consensus for each of the competences. The following interpretation guide for consensus comes from Medical Statistics at a Glance (which refers to Cohen's Kappa) and shows that generally the consensus on the rating of each competency was 'substantial' for the generic and level-1 competences, and highly variable between 'good' and 'fair' for the level-2 competences:

- poor: less than or equal to 0.20
- fair: 0.21 to 0.40
- moderate: 0.41 to 0.60
- substantial: 0.61 to 0.80
- good: greater than 0.80

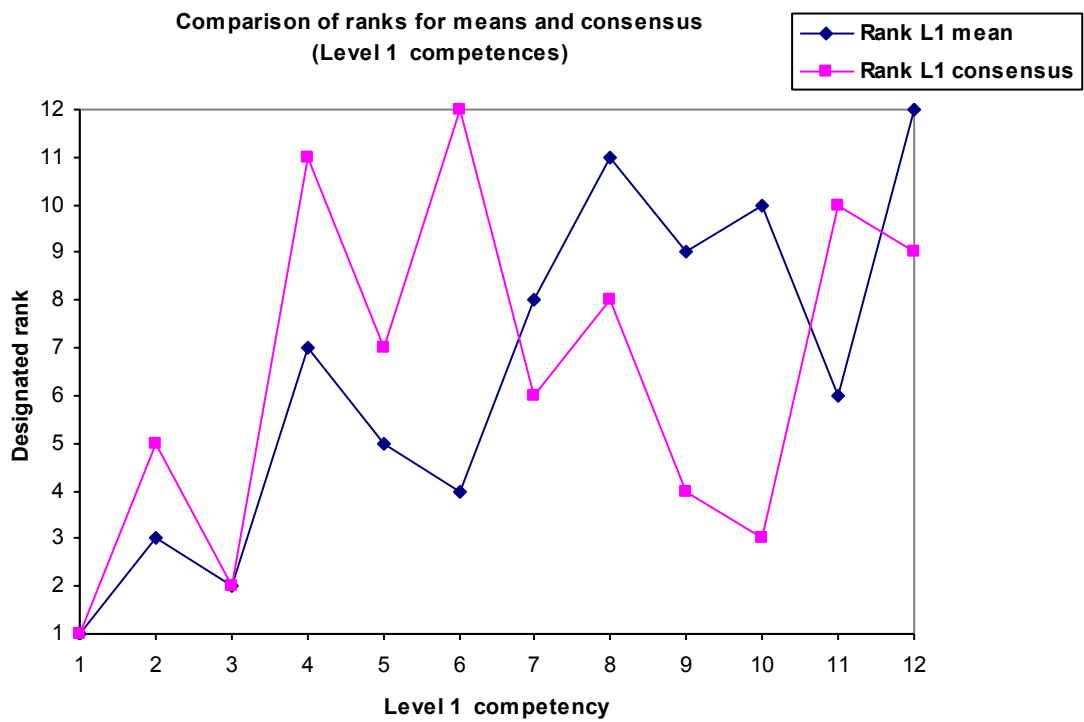
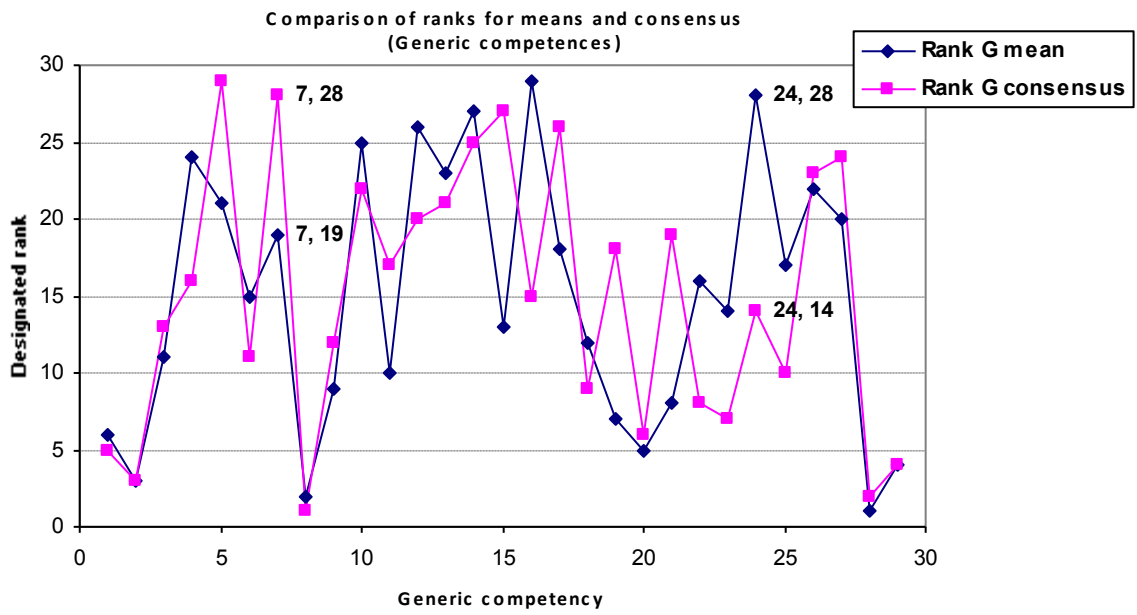
G category	G mean	G median	G consensus	Rank G mean	Rank G median	Rank G consensus
28	3.63	4.00	0.75	1	3.5	2
8	3.61	4.00	0.76	2	3.5	1
2	3.58	4.00	0.72	3	3.5	3
29	3.58	4.00	0.72	4	3.5	4
20	3.57	4.00	0.71	5	3.5	6
1	3.51	4.00	0.71	6	3.5	5
19	3.41	3.00	0.61	7	13.5	18
21	3.37	3.00	0.61	8	13.5	19
9	3.35	3.00	0.65	9	13.5	12
11	3.27	3.00	0.61	10	13.5	17
3	3.23	3.00	0.64	11	13.5	13
18	3.22	3.00	0.67	12	13.5	9
15	3.21	3.00	0.55	13	13.5	27
23	3.14	3.00	0.71	14	13.5	7

6	3.10	3.00	0.66	15	13.5	11
22	3.09	3.00	0.69	16	13.5	8
25	2.87	3.00	0.66	17	13.5	10
17	2.70	3.00	0.56	18	13.5	26
7	2.58	3.00	0.54	19	13.5	28
27	2.54	3.00	0.57	20	13.5	24
5	2.51	2.00	0.45	21	25.0	29
26	2.49	2.00	0.57	22	25.0	23
13	2.45	2.00	0.60	23	25.0	21
4	2.45	2.00	0.62	24	25.0	16
10	2.44	2.00	0.60	25	25.0	22
12	2.33	2.00	0.61	26	25.0	20
14	2.29	2.00	0.56	27	25.0	25
24	2.24	2.00	0.64	28	25.0	14
16	2.22	2.00	0.64	29	25.0	15

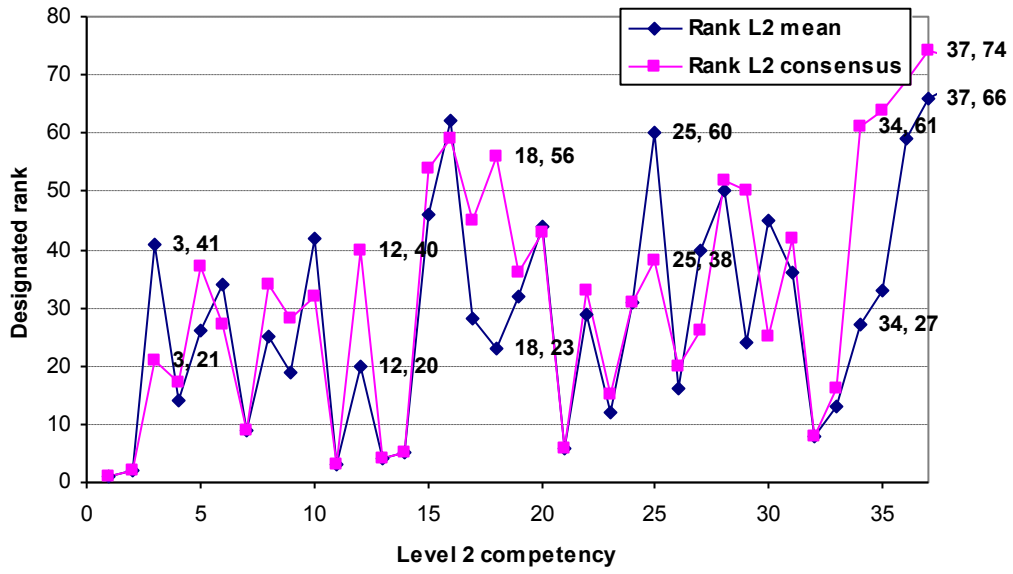
L1 category	L1 mean	L1 median	L1 consensus	Rank L1 mean	Rank L1 median	Rank L1 consensus
1	3.77	4.00	0.85	1	2.0	1
3	3.66	4.00	0.77	2	2.0	2
2	3.50	4.00	0.66	3	2.0	5
6	3.36	3.00	0.58	4	8.0	12
5	3.31	3.00	0.64	5	8.0	7
11	3.26	3.00	0.60	6	8.0	10
4	3.26	3.00	0.59	7	8.0	11
7	3.17	3.00	0.66	8	8.0	6
9	3.02	3.00	0.67	9	8.0	4
10	2.93	3.00	0.69	10	8.0	3
8	2.89	3.00	0.63	11	8.0	8
12	2.83	3.00	0.61	12	8.0	9

L2 category	L2 mean	L2 median	L2 consensus	Rank L2 mean	Rank L2 median	Rank L2 consensus
1	3.80	4	0.87	1	13.5	1
2	3.78	4	0.85	2	13.5	2
11	3.77	4	0.85	3	13.5	3
13	3.76	4	0.84	4	13.5	4
14	3.76	4	0.84	5	13.5	5
21	3.75	4	0.84	6	13.5	6
63	3.73	4	0.82	7	13.5	7
32	3.61	4	0.74	8	13.5	8
7	3.56	4	0.71	9	13.5	9
69	3.54	4	0.69	10	13.5	10
70	3.53	4	0.68	11	13.5	13
23	3.53	4	0.68	12	13.5	15
33	3.51	4	0.67	13	13.5	16
4	3.51	4	0.67	14	13.5	17
61	3.50	4	0.67	15	13.5	19
26	3.49	4	0.66	16	13.5	20
58	3.48	4	0.66	17	13.5	22

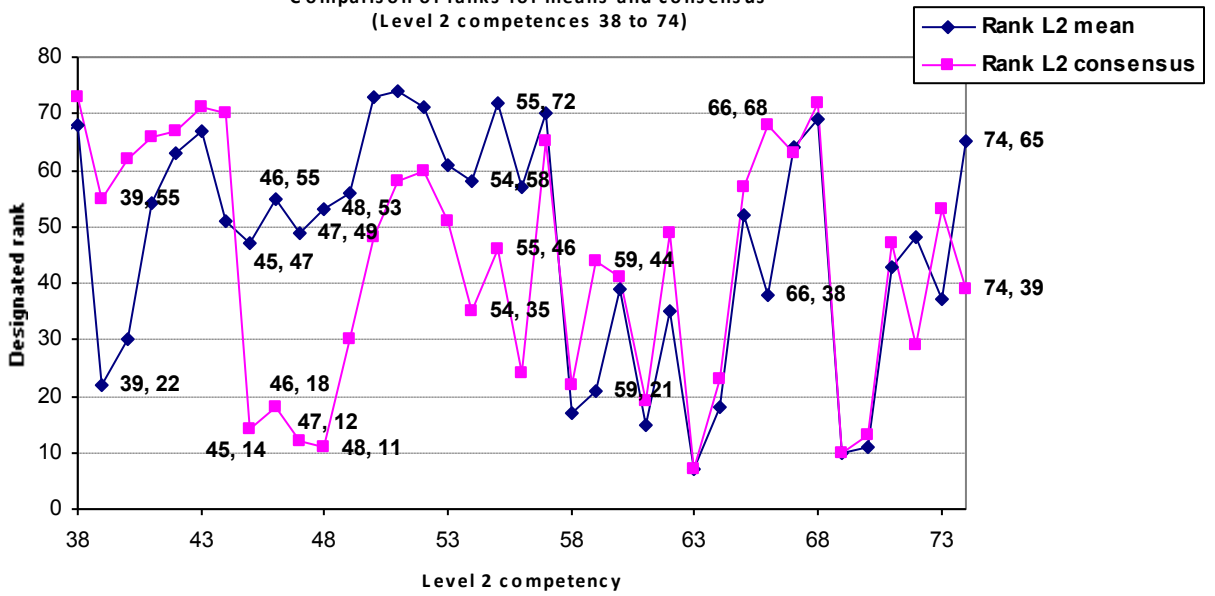
64	3.48	4	0.65	18	13.5	23
9	3.46	4	0.64	19	13.5	28
12	3.44	4	0.63	20	13.5	40
59	3.43	4	0.62	21	13.5	44
39	3.39	4	0.59	22	13.5	55
18	3.39	4	0.59	23	13.5	56
29	3.39	3	0.61	24	47	50
8	3.39	3	0.63	25	47	34
5	3.37	3	0.63	26	47	37
34	3.36	4	0.57	27	13.5	61
17	3.36	3	0.62	28	47	45
22	3.33	3	0.63	29	47	33
40	3.33	4	0.55	30	13.5	62
24	3.31	3	0.64	31	47	31
19	3.30	3	0.63	32	47	36
35	3.30	4	0.54	33	13.5	64
6	3.30	3	0.64	34	47	27
62	3.30	3	0.61	35	47	49
31	3.29	3	0.62	36	47	42
73	3.28	3	0.60	37	47	53
66	3.28	3	0.52	38	47	68
60	3.25	3	0.62	39	47	41
27	3.24	3	0.64	40	47	26
3	3.22	3	0.66	41	47	21
10	3.22	3	0.63	42	47	32
71	3.21	3	0.61	43	47	47
20	3.21	3	0.62	44	47	43
30	3.17	3	0.65	45	47	25
15	3.15	3	0.60	46	47	54
45	3.11	3	0.68	47	47	14
72	3.10	3	0.64	48	47	29
47	3.09	3	0.69	49	47	12
28	3.08	3	0.60	50	47	52
44	3.07	3	0.49	51	47	70
65	3.02	3	0.59	52	47	57
48	3.01	3	0.69	53	47	11
41	3.01	3	0.53	54	47	66
46	3.00	3	0.67	55	47	18
49	3.00	3	0.64	56	47	30
56	3.00	3	0.65	57	47	24
54	2.99	3	0.63	58	47	35
36	2.99	3	0.50	59	47	69
25	2.96	3	0.63	60	47	38
53	2.94	3	0.61	61	47	51
16	2.91	3	0.58	62	47	59
42	2.90	3	0.53	63	47	67
67	2.86	3	0.54	64	47	63
74	2.85	3	0.63	65	47	39
37	2.76	3	0.45	66	47	74
43	2.72	3	0.49	67	47	71
38	2.51	2	0.47	68	71	73
68	2.47	2	0.49	69	71	72
57	2.45	2	0.53	70	71	65
52	2.15	2	0.58	71	71	60
55	2.14	2	0.62	72	71	46
50	1.79	2	0.61	73	71	48
51	1.70	2	0.59	74	71	58



Comparison of ranks for means and consensus
(Level 2 competences 1 to 37)



Comparison of ranks for means and consensus
(Level 2 competences 38 to 74)



- IV) Competences will be tagged with appropriate themes when the framework is developed into a document for dissemination to curriculum developers.
- V) Highlighted errors in the Tuning (Medicine) Validation Brochure will be corrected in the final report as indicated by the Panel.
- VI) We agree that these core competences are “Necessary but not sufficient”. The remit of the Tuning (Medicine) project was to gain consensus on core competences rather than all competences, but we are very aware that there will be other competences specific to different national systems and HE institutions, and that in many institutions students are encouraged to follow their own interests in developing optional competences. Disseminated findings of this work will clearly state that these are ‘necessary but not sufficient’.
- VII) Level-3 competences could be defined as highlighted, but are beyond the scope of the current work.
- VIII) We have already fed-back to the organisers of the Healthcare Validation meeting that we would have appreciated opportunities to interact with those undertaking Tuning projects in other disciplines.
- IX) It would be very interesting to repeat the Tuning process again in a number of years as an ongoing project, but would require adequate funding / support.
- X) Funding from the European Commission was sought to continue the work of the MEDINE Tuning Task Force. This application proved unsuccessful, but further options for supporting and continuing this work are currently being considered.