ASSESSMENT LITERACY OF NATIONAL EXAMINATION INTERVIEWERS AND RATERS – EXPERIENCE WITH THE CEFR

Ene Alas, Suliko Liiv

Abstract. The article investigates the training needs for the English language national examination interviewers and raters in Estonia in light of the new national examination that is set to measure students' proficiency at level B of the Common European Framework of Reference for Languages (henceforth CEFR) scale. A questionnaire study was designed to explore CEFR-related assessment literacy – the extent to which the target population employed the CEFR in their daily professional life, how accessible they felt the CEFR was for their professional needs and if they thought they could accurately place students on the CEFR levels. Additionally, the respondents assigned CEFR can do statements to the levels deemed appropriate. The results revealed ambiguity about the CEFR levels among the respondents, which lead the authors to propose a training sequence designed to empower the interviewers and raters to function more efficiently within the new national examination speaking test in Estonia.*

Keywords: assessment literacy, interviewer, rater, oral proficiency interview, CEFR, training

1. Introduction. Background to the study

The national examination in the English language in Estonia up until 2013 attempted developed measure to what extent the National Curriculum requirements concerning the learning outcomes in the English language have been met with a view that by the end of gymnasium students should reach B2 level on the CEFR scale as stipulated in the national curriculum of 2002. The exam developed was a 100 point exam, functioning as a B2 level achievement test. 'However, no research has been carried out to determine whether this is really the case. Anecdotal evidence suggests that, at least in the case of the English examination, the language competence

of those students who have gained at least 80-85 points out of the possible 100 (about 25-30 per cent of the test-taking population) is at level C1.' (Türk, Tender 2013: 233) which was also corroborated by SurveyLang findings reported in 2012 (European Commission 2012: 212–213). Like many other national examinations, the current exam has also acted as a proficiency measurement tool generating data that serve gate-keeping purposes in Estonia. Most universities in Estonia have set a national examination standard that has to be met in order to be admitted to a university programme.

In 2011, a new national curriculum was adopted by the Estonian Ministry of Education and Research (cf. Gümnaasiumi riiklik õppekava), which stipulates that students finishing Gymnasium should have (CEFR) B level proficiency in two foreign languages. Moreover, the new curriculum maintains that 'ainevaldkonda kuuluvate võõrkeelte õppe kirjeldus on üles ehitatud, lähtudes keeleoskustasemete kirjeldustest Euroopa keeleõppe raamdokumendis [---]. Kõikide võõrkeelte [---] õpitulemused on raamdokumendile toetudes kirjeldatud ühtsetel alustel' (ibid.: Appendix 2, p. 1) ['description of the learning of foreign languages belonging to the subject area has been developed proceeding from the level descriptors of the Common European Framework of Languages [---]. The learning outcomes of all foreign languages have been described in a unitary fashion relying on the CEFR'.] Thus a conscious decision has been made to align foreign language instruction and evaluation to the guidelines of the CEFR, implying that also language proficiency testing via national examinations would have to follow suit.

As the national curriculum does not specify clearly which of the two Independent User B levels – B1 or B2 – is to be achieved, and the students may in actual fact reach either, the new national examination in the English language has been developed to allow the student to be placed at either level. The new 2014 exam is designed to be a 100 point exam like its predecessor, comprising 4 tests: writing, listening, reading and speaking. A foreign language national examination will be a compulsory graduation requirement as of next year. Although English has always been the most popular foreign language in terms of how many students choose it as one of their school-leaving examinations (the number varying between 7500 and 9500 students annually), the amount of test-takers may increase as a result of the new regulation, creating added challenges for test administration and marking.

The process of developing a new test has meant overcoming a number of challenges in terms of overall approach (a bi-level examination vs. a single level examination), validity and practicality issues like the number and nature of tasks (how many texts/ tasks/ items would be sufficient to be able to draw informed conclusions about a person's language level without completely exhausting the test-taker/ interviewer/ rater), new marking scales for writing and speaking, benchmarking, score reporting, etc.

As a new format was proposed for the speaking test, additional challenges emerged concerning interviewer/ rater awareness and behaviour during the interview, i.e. interviewer/ rater training to tackle reliability issues. For the interviewers, the novelty is predominantly procedural: familiarising themselves with the new interviewer script, adhering to the interviewer frames, understanding the nature of tasks and managing them effortlessly. Relying on earlier research on the subject (cf. Alas 2010, Reemann et al. 2013), a substantial amount of awareness-building would

also have to be spent on the issue of time management during the conduct of interviews. For those who rate the student performances during the speaking test, the challenge lies mostly in learning to implement the new marking scale appropriately and consistently, which presupposes an ability to recognise CEFR criterial levels of performance, distinguishing between all CEFR levels but more notably between levels A2/B1/B2/C1. As can be seen from the above, interviewer and rater training has assumed notable significance for the speaking test to be conducted consistently and marked validly and reliably. The task is made ever more daunting by the fact that the number of people involved in the speaking test as interviewers / raters in Estonia is 998 and they all require re-qualification before the implementation of the new examination in May 2014.

Before embarking on planning a re-training programme, information was sought as to the interviewers'/ raters' degree of familiarity with the level descriptors of the CEFR as well as their attitude to level descriptors' usefulness in their day-to-day teaching and assessment practices, i.e. the usefulness of the CEFR as an assessment/ testing instrument. The current research is thus attempting to study assessment literacy with respect to the CEFR.

The term 'assessment literacy' was coined by Stiggins (1991, 1997) and has been defined by Malone as 'stakeholders' (often with a focus on instructors') familiarity with measurement practices and the application of this knowledge to classroom practices in general and specifically to issues of assessing language' (Malone 2013: 330). There seems to be agreement (Brindley 2001, Malone 2008, Popham 2009, Fulcher 2012) that many test users – and here we specifically include teachers who prepare students for tests, develop tests and act as interviewers and raters during high stakes tests like national examinations – have limited knowledge about language testing, sometimes irrespective of the fact that language testing and assessment have been part of their teacher education. Research into assessment literacy has focused on the kinds of textbook that are available and their relevance for particular groups in language education population like teachers, testers and researchers (Davies 2008, Fulcher 2012), suggesting that more accessible testing/ assessment materials be developed for educators who are not necessarily testing experts but need to make informed assessment related decisions in their daily professional life. Another area of research are the characteristic features of language testing courses (Brindley 2001, Brown, Bailey 2008, Inbar-Lourie 2008, Malone 2008) where the findings point to 'the presence of a stable knowledge base [among test users] that is evolving and expanding, rather than shifting radically' (Brown, Bailey 2008: 371), still emphasising that a framework of core competences of language assessment be established for language testing courses (Inbar-Lourie 2008: 396–398) with the competences involving three areas: language testing/ assessment skills, language testing/ assessment knowledge and language testing/ assessment principles (Davies 2008: 328). Malone reminds the reader of the fact that language testing textbooks are not the only way to convey language testing/ assessment related expertise (Malone 2013: 332). In fact, other modern approaches, such as 'traditional as well as face-to-face workshops, online and downloadable tutorials, materials produced by professional language testing associations, reference frameworks, such as the CEFR, video projects, pre-conference workshops and series of narrative accounts about developing assessment literacy' (ibid.) probably outweigh the importance of textbooks in developing assessment literacy

among test users. With the CEFR and Manual for Language Test Development and Examining related to it being undoubtedly the most 'influential publications [---] in the field of language learning and language testing in Europe' (Figueras et al. 2005: 261), with research literature abounding in articles discussing the process of relating examinations to the CEFR (Figueras et al. 2005, O'Sullivan 2010, Kecker, Eckes 2010, Martyniuk 2010, Noijons et al. 2011) and how the CEFR can serve as a guide to develop assessment instruments (e.g. Little 2005), there is relatively little research into the how CEFR literate test users are. The current study aims to bridge some of that gap.

2. Method

The questionnaire study, developed for the purposes of discovering how extensively teachers employ the CEFR in their professional life and how accurate they are in describing particular CEFR levels of performance, invited the participants – teachers training to become national examination speaking test interviewers/ raters – to place 6 statements (I use CEFR for test development/ marking student work/ textbook evaluation/ curriculum design/ course evaluation, I can successfully distinguish between different CEFR levels) on a Likert scale, indicating their use of the CEFR. This closed-response section was supplemented by 5 open-ended questions, asking about the usefulness of the CEFR for the participant as a language teacher, about a particular section of the CEFR that was most useful, about the ease of use of the CEFR, about having obtained any CEFR-related training and commenting on the usefulness of the latter, and about any training that the respondents thought they might need as national examination interviewers/ raters. In each case, the respondents were also requested to account for their opinion. In the third part of the questionnaire, the respondents had to place 20 speaking sub-skills (can do statements) drawn from the CEFR on the correct level. The statements for the questionnaire were chosen from the CEFR levels A2, B1, B2 and C1 as it is these levels that the students will potentially be demonstrating during the national examination speaking test.

The questionnaire was completed by 64 respondents who were starting their interviewer/ examiner initial or re-training. The respondents fell into two groups with regard to the amount of experience they had had as national examination interviewers and trainers: novice interviewers/ trainers who had had no interviewing/ rating experience, comprising 27 people; and the experienced interviewers/ raters – 37 people who had had at least 1 year of respective practice, i.e. they had been interviewing/ rating the national examination speaking test during at least one examination session.

3. Analysis and results

This section presents the findings of the current study along with some interpretation. Although the number of participants is fairly small, the results are given in percentage value to allow a more generalised view of the outcome. An attempt was also made to see if the results of both groups differed significantly from each other

using the chi-square test. The interviewers'/ assessors' confidence level about their ability to distinguish between CEFR levels is reflected in Figure 1 below.
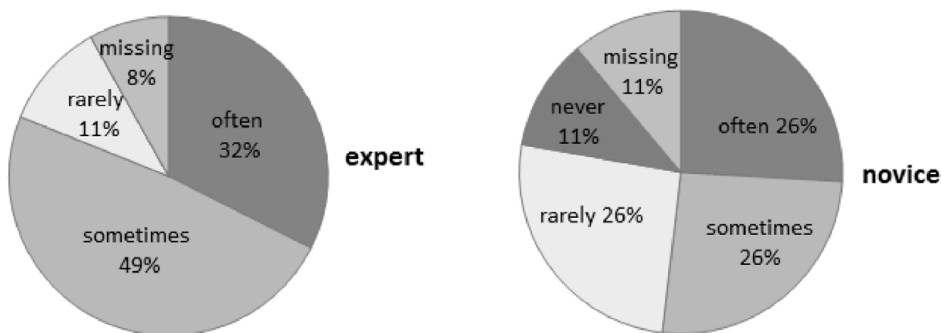
Figure 1. Interviewers'/ raters' confidence of their ability to distinguish between CEFR levels

A higher confidence level can be detected among the expert group as expected, as about a third (32%) believe themselves to be able to successfully distinguish between CEFR levels and a further 49% claim to be able to do that sometimes ($p \leq 0.05$). There is nobody in this group who said they could never do it but there is a fairly noticeable group of 8% who have not responded to this question, which may be an indication of lack of ability to separate levels. The novice group falls roughly in half with 52% of the respondents claiming to be able to identify levels successfully (26% often and 26% sometimes) and 48% needing more practice with level identification, as they can either rarely (26%) or never (11%) identify the level with a further 11% not responding to the question.

It was assumed that the interviewers and raters involved in the study who in their daily life are gymnasium level English teachers would be familiar with the CEFR and use the document in their professional life for various purposes. Five potential uses of the CEFR were proposed in the questionnaire with the request to rate the frequency of those uses. Figures 2 and 3 below reflect the results of this query for expert and novice interviewers/ raters respectively.
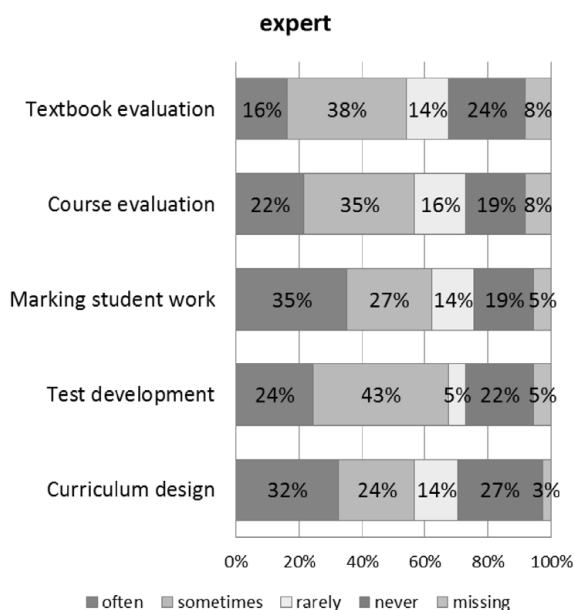
Figure 2. Using the CEFR in their professional life. Experts

It appears that the most frequent uses of the CEFR among expert interviewers/ raters are for test development (24% often and 43% sometimes) and marking student work (35% often and 27% sometimes) which is to be expected, as the CEFR is probably considered by the majority of teachers as a testing document. It is surprising, however, that more than about a third of the respondents admit either consulting the CEFR rarely or never, including the 5% who have not responded to the question. Other aspects of evaluation pertaining teaching – textbook evaluation (16% often, 38% sometimes), course evaluation (22% often, 35% sometimes) and curriculum design (32% often, 24% sometimes) are represented slightly less compared to test development and marking student work but still to an almost equal degree. Here, the group of rare or non-users is over 40% in all categories.

The comparison of the above data to that of the novice group gives a slightly different picture (cf. Figure 3 below).
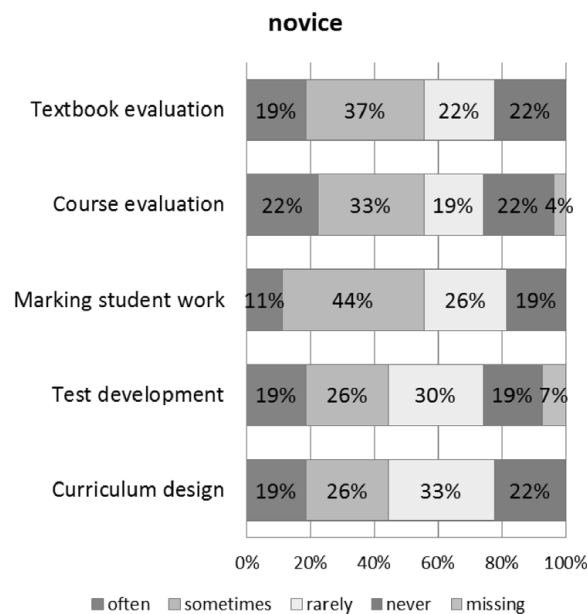
Figure 3. Using the CEFR in their professional life. Novices

The most frequently noted uses in this group are for marking student work (11% often, 44% sometimes), textbook evaluation (19% often, 37% sometimes) and course evaluation (22% often, 33% sometimes), whereas it can be noticed that the group who claims to use the CEFR frequently for textbook evaluation is larger than that of the expert group. Overall, the number of rare or non-users noticeably exceeds that of the expert group ranging from 44% for textbook evaluation to 56% in test development. Being novice national examination interviewers/ raters usually means that those respondents have also had somewhat less teaching experience, and thus they may have had less exposure to such activities as designing a course, planning a curriculum/ syllabus or evaluating the appropriacy of a textbook.

The CEFR is one of the key documents in managing (foreign) language instruction and evaluation in Europe on the governmental level. The questionnaire enquired in one of its open-ended questions how useful the respondents considered it as language teachers. The results are displayed in Figure 4 below.
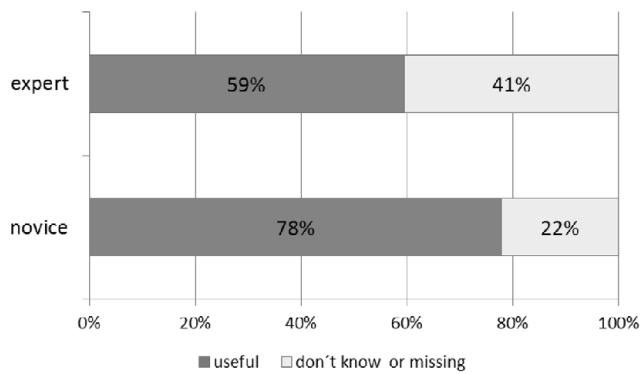
Figure 4. Usefulness of the CEFR

As can be seen, novice interviewers/ raters consider it useful more often than experts with as many as 78% of the novice respondents saying so, compared to 59% of the expert group ($p \leq 0.03$). It was speculated at this point of the analysis that the result would have something to do with the recency of the novice group's teacher education, being more familiar with the CEFR through teacher training, knowing its role and impact in planning language instruction/ evaluation. In their responses, the novice examiners/ raters pointed out CEFR's usefulness in terms of assessing the level of students, setting long-term goals, evaluating text-books and preparing for the exams; the respondents also valued CEFR because 'it is adaptable to every language situation'. Expert interviewers/ raters pointed out a larger variety of CEFR uses compared to novices: it is a guide for different kinds of evaluation, it allows the teacher to select/ customise resources, it serves as a common framework for assessment and feedback; it helps to find criteria, distinguish between proficiency levels and establish a uniform grading system; it explains different aspects of language testing. The respondents in this group valued CEFR in their development of placement tests and preparation for exams.

Although it was expected that the interviewers/ raters would consider CEFR useful for their professional purposes, the scope and the complexity of the document was predicted to cause problems for the respondent. The open-ended question concerning the ease of use of the document yielded the responses featured in Figures 5 and 6.
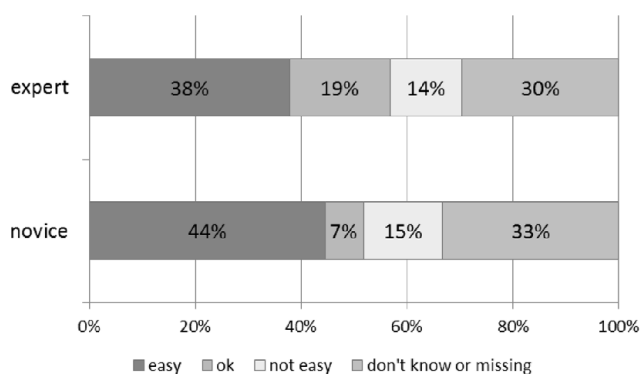
Figure 5. Ease of use of the CEFR
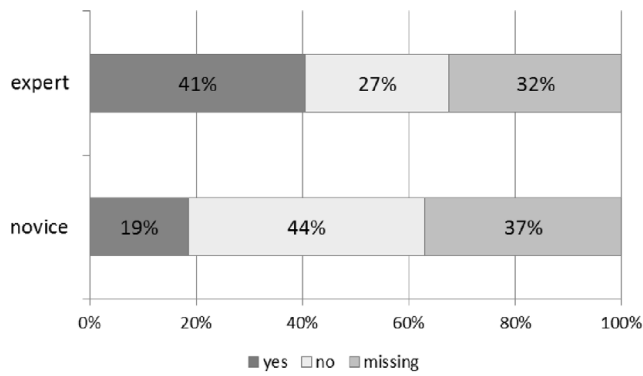
novices and experts together
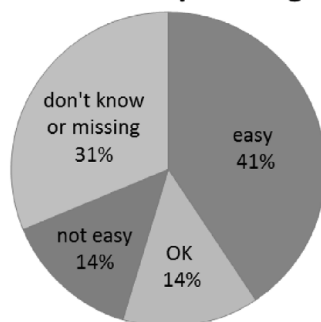
Figure 6. Ease of use of the CEFR. All respondents together

Comparing the groups that maintain that using the CEFR is either easy or mostly easy (OK) shows a slightly higher confidence level in the expert group (57% vs. 51%), the significance level of $p \leq 0.34$ showing no real difference between the groups though. While asked what made the CEFR easy to use the experts pointed out that the CEFR was well structured, logical, that the levels were clearly described but also the fact that they had gained experience with it (e.g. 'I have got used to it', 'I have studied it properly and introduced it to my teaching practice step by step'). Novices mentioned fewer reasons for the ease of use: the precision of level descriptors, clear tables and the fact that the skills were separately described. It is interesting, though, that the number of those who think it easy is larger among the novices (44% vs. 38%). This finding may be compatible with the results to the previous question: being more familiar with the CEFR through training, the respondents would be more confident using it. The fact that in both groups almost half of the respondents either considered using the CEFR hard to use (experts 14%, novices 15%) or refrained from answering the question, possibly because they had no experience using the CEFR (30% vs. 33%), points to the need of training with respect to the CEFR levels and their potential uses. Those experts who marked the CEFR not easy to use supported this by saying that language level evaluation was never easy, that the levels could be interpreted in various ways, that the evaluation was made more complicated by the difference in tasks and their own lack of practice in using the CEFR scales in their day-to-day practice. Novices mentioned similarity of level descriptors that caused confusion and their lack of experience with the scales.

All respondent were invited to share the amount and nature of their CEFR-related training so far. Figures 7 and 8 below indicate that, contrary to the assumptions made on the basis of responses to the earlier questions regarding CEFR-related training, the number of people in the expert group having gone through CEFR-related training noticeably surpasses those in the novice group (41% vs. 19%), $p \leq 0.051$.

The expert respondents who noted having had CEFR-related training indicated that although the training was useful, it had taken place quite some time ago and needed refreshing. Only a few respondents in both groups commented on the contents of the course and described it fairly ambiguously as 'working with levels'. It is noteworthy from the point of view of further training that a third of all the respondents claimed to have had no CEFR related training and 34% did not respond to the question which might be a further sign of no CEFR related training received.
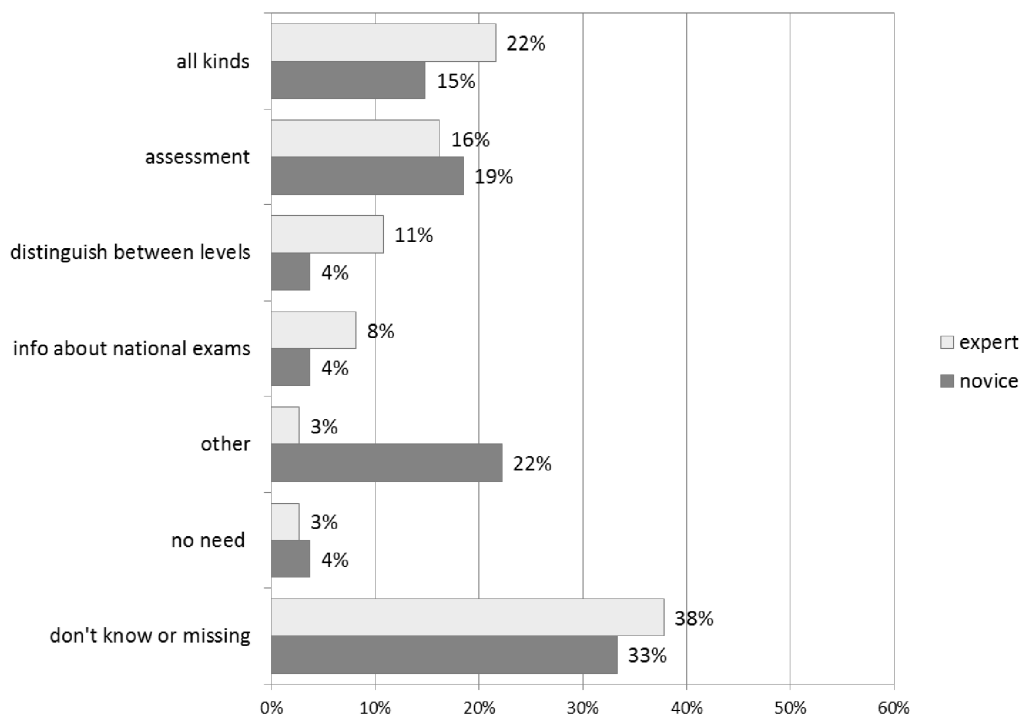
Figure 7. CEFR-related training so far

novices and experts together
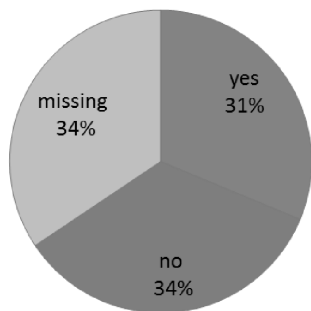
Figure 8. CEFR-related training so far. All respondents together

The final survey question concerning CEFR related training needs for the respondents as national examination interviewers/ raters provided further support for the research assumption that the target group was all but familiar with the CEFR and its potential. The responses to the question have been summarised in Figure 9 below.


**Figure 9.** Further CEFR-related training needs

There is a large proportion of respondents (38% of experts, 33% of novices) who have either indicated that either they did not know the kind of training that was necessary or did not responded to the question. The second largest group consisted of those whose response seems diametrically the opposite – 'all kinds of training' (experts 22%, novices 15%). The answer, however, is from the training point unhelpful, as it does not point to any problem areas that training might focus on or the interviewers/ assessors are aware of or admit. The response differs from the previous group's answer, though, in that it seems to indicate the respondents' interest in the suggested area of development. Also, not knowing, at the time of the survey, what the new national examination would be like, and to what extent working with the new marking scale would require familiarity with the CEFR levels, it is no surprise that the prevailing responses are somewhat vague. The more focused CEFR related training needs mention assessment of student work and being able to distinguish between CEFR levels, especially when rating speaking. The more specific training considered necessary are connected to the information about national examinations ('I need all the information connected with the new national exam as an interviewer and an assessor'), which seems more relevant to the respondents as language teachers, preparing their students for national examinations, rather than interviewers and raters. With the given group of respondents, the CEFR as a training focus is not mentioned. In this connection, it is perhaps also noteworthy that 3% of experts and 4% of novices indicated that there was no need for CEFR related training, without further specification, however.

In the final part of the survey, the respondents were asked to place 20 can do statements involving speaking derived from the CEFR levels A2, B1, B2 and C1 on the correct level. The success rate of the exercise is represented in Figures 10 and 11. The former (Figure 10) represents the overall result of assigning descriptors to the levels, whereas Figure 11 illustrates the success rate of the group of respondents who maintained that they could often distinguish between CEFR levels successfully. The figures show if the descriptor has been placed on a correct CEFR level (exact), or on a higher level (overestimated) or on a lower level than it belongs to (underestimated). As the respondents only had four levels to choose from, it was impossible for them to place A2 descriptors lower and C1 descriptors higher than they belonged to – hence the missing category of underestimated in section A2 and overestimated category in section C1.

As can be seen, the overall success rate is only marginally higher with the expert group compared to the results of all the respondents together. The level descriptors are assigned correctly on less than a half of the occasions with all the levels except C1 where the percentage of correct placement is 62 overall and 58 with the experts (sic!). It is also interesting that the descriptors are more frequently placed on a higher level than they actually belong. It may be that the respondents view the descriptors as illustrating perfect performances on that level, without allowing that although a student on A2 level would be able to 'describe plans and arrangements, habits and routines, past activities and personal experiences' (CEFR: 59) or a B2 level speaker would be able to 'give detailed descriptions and presentations on a wide range of subjects related to his/ her field of interest, expanding and supporting ideas with subsidiary points and relevant examples' (CEFR: 58) both students would still be struggling with language, A2 student with grammar as well as vocabulary and B2 student probably not so much with grammar but still with the range and appropriacy of vocabulary.
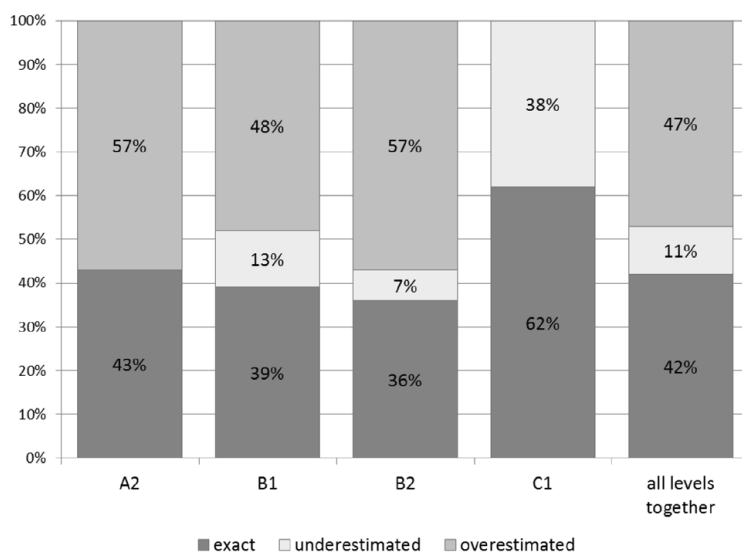
Figure 10. Assigning can do statements to appropriate CEFR levels. Overall
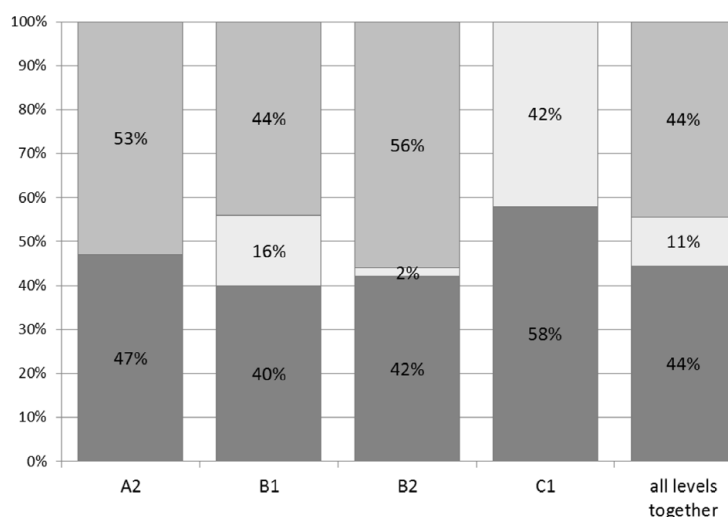
Figure 11. Assigning can do statements to appropriate CEFR levels. Experts

Having the above expectations might result in rating the students in the speaking test lower than they belong on a CEFR scale. Another explanation might be the one offered by Papageorgiou: 'judges tend to have their own preconception of the CEFR levels based on what examinees they know can do with the language and such understanding [can] have an effect on [their judgements]' (Papageorgiou 2010: 276).

4. Discussion and intended training

The above study was designed to investigate the interviewers'/ raters' degree of familiarity with the level descriptors of the CEFR as well as their attitude to its usefulness in their day-to-day teaching and assessment practices. The results show that, although there is an expectation that the national examination interviewers/ assessors start implementing a marking scale that is closely connected to CEFR

level descriptors, the degree of CEFR-related literacy varies considerably among the respondents. There is a fairly large group who seems to have had very little or no contact with the document whatsoever. And even those interviewers/ raters who admit having had CEFR related training, refer to the necessity of updating the knowledge about it and skills of working with it.

Although the degree of familiarity with the CEFR is generally low, many respondents expressed confidence with respect to being able to place students on appropriate CEFR level, the confidence being high both in the expert and novice group. The practical task of the survey of identifying particular descriptors with CEFR levels demonstrated a low success rate even among expert raters. Allowing that the CEFR descriptors are not without problems themselves (cf. Fulcher et al. 2011: 8), such a low success rate among the respondents calls for training that would aim at a more standardised understanding of what a student on a particular level can do, and how students on adjacent levels differ from each other in terms of their skills.

The need for CEFR related training having been established, the precise nature of the training is harder to be determined judged solely by the results of the questionnaire. The vast majority of the respondents refrain from making suggestions or respond with a fairly ambiguous request for 'any kind' of training. These findings are somewhat at odds with Fulcher who found that teachers who also act as testers want concrete help with 'a text that is not light on theory but explains concepts clearly [---], a practical "how-to" guidance [---], a balance between classroom and large-scale testing' (2012: 122). Though the current respondents' suggestions to incorporate assessment and distinguishing between levels into the training programme are helpful to some extent, the actual content of the training course is left to the discretion of the test developer with the hope that information would be obtained first and foremost about the national examination and also about the CEFR, if the latter has a bearing to the exam.

The interviewer/ rater training, bearing in mind the new examination format as well as the results of the given survey, has been planned relying on the recommendations given in Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual, chapter 5 (henceforth the Manual, Noijons et al. 2011), which has been adapted to the current needs. The training is intended to be divided into 4 stages:

- Phase 1: **Familiarisation.** This is a pre-training phase and consists in the participants working independently with the CEFR descriptors and completing task-sheets that focus the participants' attention on the differences between levels A2 to C1. Before coming to the training site, the participants should watch the sample interview of the speaking test and familiarise themselves with the marking scale developed for the national examination rating of speaking. The documents, tasks with keys and the sample interview will be available to the participant on-line, on the training web-page.
- Phase 2: **Illustration.** In this phase sample interviews will be watched in a group to identify salient aspects of interviewer behaviour on the one hand and sample performances of different CEFR levels, on the other. Here it will be important to view and discuss samples of both 'flat' and 'uneven' profiles (Manual: 42), benchmarked to the CEFR levels (Takala 2004).

- Phase 3: **Practice.** Here, the participants proceed to rate samples of different levels individually and in a group, to gain further confidence. In the process, lenient and strict raters will be found, followed by further discussion and adjustment of scores.
- Phase 4: **Individual assessment.** To complete the training, the raters will have to individually rate a set of speaking examination performances, complete rating grids and provide commentary if they feel the performance was extraordinary in some respect. Completed forms will then be handed in to the trainers. Individual feedback would be provided with regard to the success rate.

5. Conclusion

The article investigated CEFR-related assessment literacy among teachers who act as interviewers/ raters during the national examination in the English language in Estonia and studied their ability to successfully distinguish between CEFR levels. The findings show a fairly low level of familiarity among the teachers with the opportunities for the potential use of the CEFR for language learning purposes and confirm what research into the relevant literature had found earlier that 'respondents tend to say that everything presented to them is important, resulting in little variation' (Fulcher 2012: 118). The responses also indicated the need for a more substantial attention to be paid to CEFR-related pre- and in-service teacher training.

The study was conducted within one cultural context (Estonia) and involved a fairly small amount of respondents, still displaying significant differences on occasion between the two groups investigated. For more generalised conclusions, teachers assuming similar roles (interviewer/ tester working with the CEFR-related scales) could be involved in other cultural contexts. Further research should also focus on the success rate of the intended training to see if, as a result of a more focused CEFR-related interviewer/ rater training, more level-awareness will develop among the raters.

References

Alas, Ene 2010. The English Language National Examination Validity Defined by Its Oral Proficiency Interview Interlocutor Behaviour. Tallinn University Dissertations on Humanities 22. Tallinn: Tallinn University Press.

Brindley, Geoff 2001. Language assessment and professional development. – C. Elder, A. Brown, E. Grove, K. Hall, N. Iwashita, T. Lumley, T. McNamara, K. O'Loughlin (Eds.). Experimenting with Uncertainty. Essays in Honour of Alan Davies. Studies in Language Testing. Cambridge: Cambridge University Press, 126–136.

Brown, James Dean; Bailey, Kathleen M. 2008. Language testing courses: What are they in 2007? – Language Testing, 25 (3), 349–385. http://dx.doi.org/10.1177/0265532208090157

CEFR = Council of Europe 2001. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: Cambridge University Press.

Davies, Alan 2008. Textbook trends in teaching language testing. – Language Testing, 25 (3), 327–349. http://dx.doi.org/10.1177/0265532208090156

European Commission 2012. First European Survey on Language Competences: Final Report. http://ec.europa.eu/languages/policy/strategic-framework/documents/language-survey-final-report_en.pdf (11.2.2014).

Figueras, Neus; North, Brian; Takala, Sauli; Verhelst, Norman; van Avermaet, Piet 2005. Relating examinations to the Common European Framework: A manual. – Language Testing, 22 (3), 261–279. http://dx.doi.org/10.1191/0265532205lt308oa

Fulcher, Glenn 2012. Assesment literacy for the language classroom. – Language Assessment Quarterly, 9 (2), 113–132. http://dx.doi.org/10.1080/15434303.2011.642041

Fulcher, Glenn; Davidson, Fred; Kemp, Jenny 2011. Effective rating scale development for speaking tests: Performance decision trees. – Language Testing, 28 (1), 5–29. http://dx.doi.org/10.1177/0265532209359514

Gümnaasiumi riiklik õppekava. [National curriculum for upper secondary schools.] RT I, 28.8.2013, 8. https://www.riigiteataja.ee/akt/128082013008 (14.2.2014).

Inbar-Lourie, Ofra 2008. Constructing a language assessment knowledge base. A focus on language assessment courses. – Language Testing, 25 (3), 385–403. http://dx.doi.org/10.1177/0265532208090158

Kecker, G.; Eckes, T. 2010. Putting the manual to the test: The test-Daf – CEFR Linking Project. – W. Martyniuk (Ed.). Relating Language Examinations to the Common European Framework of Reference for Languages: Case studies and reflections on the use of the Council of Europe's Draft Manual. Cambridge: Cambridge University Press, 50–79.

Little, David 2005. The Common European Framework and the European Language Portfolio: Involving learners and their judgements in the assessment process. – Language Testing, 22 (3), 321–336. http://dx.doi.org/10.1191/0265532205lt311oa

Malone, Margaret E. 2008. Training in language assessment. – E. Shohamy, N. Hornberger (Eds.). Encyclopaedia of Language and Education, Vol. 7: Language Testing and Assessment. 2nd ed. New York: Springer Science and Business Media, 225–233.

Malone, Margaret E. 2013. The essentials of assessment literacy: Contrast between testers and users. – Language Testing, 30 (3), 329–344. http://dx.doi.org/10.1177/0265532213480129

Manual for Language Test Development and Examining. For use with the CEFR. Produced by ALTE on behalf of the Language Policy Division, Council of Europe 2011. Strasbourg: Council of Europe

Martyniuk, Waldemar (Ed.) 2010. Aligning Tests with the CEFR. Studies in Language Testing 33. Cambridge: Cambridge University Press.

Noijons, José; Bérešová, Jana; Breton, Gilles; Szabó, Gábor (Eds.) 2011. Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). Highlights from the Manual. Strasbourg: Council of Europe Publishing.

O'Sullivan, B. 2010. The City and Guilds Communicator Examination Linking Project: A brief overview with reflections on the process. – W. Martyniuk (Ed.). Relating Language Examinations to the Common European Framework of Reference for Languages: Case studies and reflections on the use of the Council of Europe's Draft Manual. Cambridge: Cambridge University Press, 33–49.

Papageorgiou, Spiros 2010. Investigating the decision-making process of standard setting participants. – Language Testing, 27 (2), 261–282. http://dx.doi.org/10.1177/0265532209349472

Popham, William James 2009. Assessment of literacy of teachers: Faddish or fundamental? – Theory into Practice, 48 (1), 4–11. http://dx.doi.org/10.1080/00405840802577536

Reemann, Edith; Alas, Ene; Liiv, Suliko 2013. Interviewer behaviour during oral proficiency interviews: A gender perspective. – Estonian Papers in Applied Linguistics, 9, 209–227. http://dx.doi.org/10.5128/ERYa9.14

Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual. Strasbourg: Council of Europe, 2009.

Stiggins, Rick 1991. Assessment literacy. – The Phi Delta Kappan, 72, 534–539.

Stiggins, Rick 1997. Student-Centered Classroom Assessment. Upper Saddle River: Prentice Hall.

Takala, Sauli (Ed.) 2004. Reference Supplement to the Preliminary Version of the Manual for Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching and Assessment. Strasbourg: Council of Europe.

Türk, Ülle; Tender, Tõnu 2013. Re-designing the school-leaving foreign language examinations in Estonia. – Jozef Colpaert, Mathea Simons, Ann Aerts, Margret Oberhofer (Eds.). Proceedings, "Language Testing in Europe: Time for a New Framework?". Antwerp: University of Antwerp.

Ene Alase (Tallinna Ülikool) teadushuvid on keeletestimine, testide koostamine ja nende kvaliteedi hindamine, õpetajakoolitus ja õppekirjanduse hindamine.
Narva mnt 25, 10120 Tallinn, Estonia
enealas@tlu.ee

Suliko Liivi (Tallinna Ülikool) uurimisvaldkonnad on kontrastiivuuringud, kultuuridevaheline suhtluspädevus, keelepoliitika, võõrkeelte õpetamise metoodika, testimine.
Narva mnt 25, 10120 Tallinn, Estonia
liiv@tlu.ee

RIIGIEKSAMI INTERVJUEERIJATE JA HINDAJATE HINDAMISALANE PÄDEVUS – KEELEÕPPE RAAMDOKUMENT JA SELLE TUNDMINE

Ene Alas, Suliko Liiv
Tallinna Ülikool

Artiklis vaadeldakse inglise keele riigieksami intervjueerijate ja hindajate väljaõppe vajadusi tulenevalt uue inglise keele riigieksami kehtima hakkamisest. Uus eksam on koostatud nii, et see mõõdaks Euroopa keeleõppe raamdokumendi iseseisva keeleoskuse B alatasemeid B1 ja B2 ühe eksami kaudu. Kuna inglise keele õpetamine ja selle hindamine on riiklikus õppekavas otseselt seotud Euroopa keeleõppe raamdokumendi tasemetega, püüdis käesolev uurimus välja selgitada, kui pädevad on inglise keele riigieksami eksamineerijad ja hindajad raamdokumendi kasutamisel: mil määral raamdokumenti igapäevases koolitöös kasutatakse, kui lihtne on dokumenti kasutada ning kuivõrd täpselt vastajad enda arvates õpilaste teadmisi raamdokumendi skaaladele toetudes hinnata oskavad. Lisaks paluti vastajatel raamdokumendi tasemekirjeldusi õigete tasemetega ühendada. Uurimusest selgus, et raamdokumenti kasutatakse peamiselt õpilaste ja õppekirjanduse hindamisega seotud tegevustes ning tasemekirjeldustes kiputakse kirjet õigest tasemest kõrgemale asetama. Tulemustest lähtuvalt on autorid välja pakkunud intervjueerijate ja hindajate koolituskava, mis aitaks Euroopa keeleõppe raamdokumendi tasemeid paremini tundma õppida ja Eesti inglise keele riigieksami suulist osa usaldusväärsemalt hinnata.

Võtmesõnad: hindamisalane pädevus, intervjueerija, hindaja, suuline keeleeksam, Euroopa keeleõppe raamdokument, väljaõpe