

DESIGNING, IMPLEMENTING AND EVALUATING AN INNOVATIVE ONLINE INTERVIEWER AND RATER IN-SERVICE TRAINING PROGRAM

Ene Alas, Kristel Kriisa

Abstract. The current article reports on action research within an exploratory case study discussing the process of designing an online training program for English as a Foreign Language National Examination examiners and raters in Estonia, the first of its kind to be used in the context explored. The process is guided by Collins' Cognitive Apprenticeship Model and elements of the TPACK framework by Koehler et al. to determine the dimensions of the program, discuss its implementation and the trainees' response to it. The analysis of the process revealed clear stages in its evolution as well as its cyclical nature. The implementation of the resultant online program, consisting of video-based trainer input and interactive interview management and student performance assessment tasks, displayed notable training benefits – clear trainee satisfaction with the training quality and an increase in learner autonomy.

Keywords: assessment literacy, course content, task development, task sequencing, learner autonomy, training sociology, trainee response

1. Introduction

The national examination (NE) in English as a foreign language (EFL) in Estonia has been operational since 1997. As a result of continuous development and moderation, it assumed its current format in 2014. Parallel with the NE test development, close attention has been paid to the initial and continuous training of the examiners (interviewers) and raters (assessors) involved in managing the speaking test, one of the four subtests of the examination. As good assessment literacy is a key skill for any NE examiner and rater (Bijani 2018, Zulaiha, Mulyono 2020), all the examiners and raters in the Estonian educational context have been required to undergo initial certification training and after that get recertified every three

years. The recertification involved a three-hour instructor-led training in a NE training centre and focused on standardising both interviewing procedures and student performance rating. Given the scope of the training needed – approximately 300 examiners and raters to be retrained annually, the availability of trainers, and, mostly, the efficacy of large group training sessions – a more effective training instrument was called for.

First of all, it was deemed necessary to allow the trainees an opportunity to reflect on their interviewing and rating practices individually and make comparisons to the expected standard actively on their own, rather than be passive followers of the trainer presentation during on-location mass training. Another objective was to have a program that would allow the training centre a more individualised overview of the professional standard of the examiners and raters, which the final accounts of their individual training sessions would provide. The program would also provide a more standardised version of training, as opposed to the variation that inadvertently occurs during different trainer presentations.

An online examiner and rater training program was thus developed, its launch being spurred by the constraints involving any kind of face-to-face training, resulting from the 2020 pandemic restrictions. The discussion below follows the development of the training program in the light of the existing frameworks for online training in education.

2. Theoretical background

Online training programs aimed at standardising EFL speaking test examiner and rater behaviour seem to be gaining popularity (Savage 1990, Elder et al. 2007, Caulfield 2011, Knoch et al. 2016, Anderson 2018). Studies discussing the development of effective online examiner and rater training programs for the purposes of standard language testing are not numerous, however. Research outlines different advantages of online programs, for example, Elder et al. (2007) highlight the rater-training benefits – ‘facilitating access to training materials and rating samples and allowing raters to reorient themselves to the rating scale and self-monitor their behaviour at their own convenience’ (2007: 36); and Knoch et al. cite ‘flexibility of training at home in their own time [...] paper savings and the opportunity for reflection at an individual’s personal pace’ (2016: 91) as the assets of such training. Boling et al. (2012) second Knoch et al.’s flexibility and convenience argument but highlight a negative side of online training by citing students’ response to teachers’ requirement for synchronous online classrooms as challenging and frustrating. Knoch et al. also note that besides advantages, online training comes with its own set of problems – ‘technical issues, the strain of reading online and the lack of direct interaction with a trainer’ (2016: 91) to name a few.

In order to design technology-enhanced teacher-education programs, Koehler et al. (2014) propose the TPACK framework which they have found useful. The framework shows how the teachers’ use of technology is dependent on subject/content knowledge (CK), pedagogical knowledge (PK) and technological knowledge (TK). Thus, the three aspects need to be considered during the designing process and if successfully managed, additional benefits will emerge. The authors maintain that

while technology is being used for educational purposes, the three types of knowledge recombine to give rise to additional types of knowledge – technological content knowledge (TCK), pedagogical content knowledge (PCK), technological pedagogical knowledge (TPK) – which all together comprise technological pedagogical content knowledge (TPACK), see Figure 1 below.

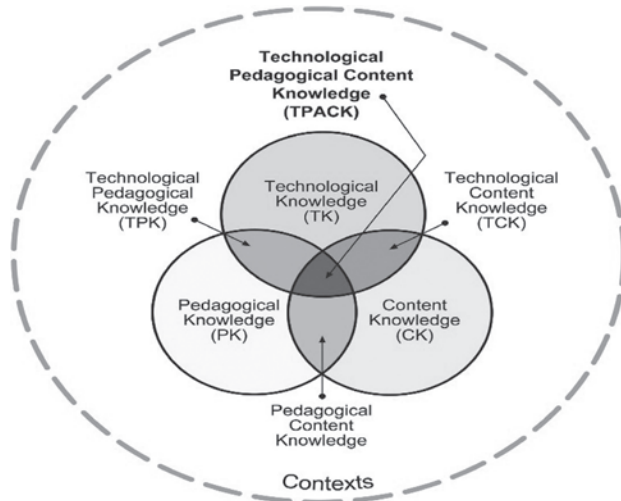


Figure 1. Koehler et al. 2014 TPACK framework

Koehler et al. define TPACK as ‘knowledge about the complex relations among technology, pedagogy, and content that enable teachers to develop appropriate and context-specific teaching strategies’ (Koehler et al. 2014: 102).

When starting to develop an online training program, awareness of the above complexity is essential. Research focuses on the different variables that need consideration in the design process, echoing the TPACK framework: the software and different applications (Malmsköld 2007), the scope of the subject content that needs to be imparted during training, the constraints of the task types chosen (Gustafsson et al. 2003, Boling et al. 2012), the learners’ computer literacy (Hsu 2011), etc. Bluemel et al. (2003) suggest three important online training program characteristics – interactivity, flexibility and easy authoring – maintaining that ‘the learning processes are stimulated when the user has to be active [---] the tool must be able to provide the user with a wide range of different training scenarios and [---] the system must be easy to author so that people with specific knowledge can be involved in the authoring process’ (p. 2002). It is an overall conviction that it is the ‘teacher attitudes towards technology [which is] a crucial element that determines the extent and ease of technology adoption’ (see Kadel 2005 and Canals, Al-Rawashdeh 2019 for a comprehensive review). An important angle to the discussion is proposed by Boling et al. (2012) who, relying on previous research, suggest that ‘the use of technology in education increasingly demands a shift from a teaching to a learning paradigm’. The model they propose is Collins’ (2006) Cognitive Apprenticeship Model (CAM), which requires the training program to include four dimensions: ‘*content, method, sequencing, and sociology*’. For a successful acquisition of the

proposed subject *content*, the authors recommend ‘modeling, coaching, and scaffolding’ as *method*, i.e. choosing tasks that are meaningful, require problem solving, and allow learner autonomy. The proposed tasks should then be logically *sequenced* which means planning the order of the learning activities carefully and ensuring their ascending complexity and diversity. The *sociology* of learning suggests that the training program be designed so that during training, the learner is allowed to set personal goals and cooperate with other learners in order to achieve those goals. The authors further emphasise the importance of a multi-modal approach in the training program, i.e., using ‘both visual and auditory modes such as text, graphics, audio, and video’ (Boling et al. 2012). The CAM model for program development, combined with the consideration of the TPACK framework implications for online training appears to present a viable path for an online training program design.

The current study was designed to find answers to the following research questions:

1. What kind of a process leads to attaining an effective online training program for the NE examiners and raters in the given context?
2. What is the trainees’ response to the proposed online training program?
3. What implications does the trainees’ response to the given online training have for the implementation of the program?

3. Methodology: Steps in the program design process

The discussion will follow the design of action research within an exploratory case study. The study investigates ‘work towards practice change during the research process’ (Coe et al. 2017: 71, Laherand 2018: 133) in a specific educational context (Yin 2009). As typical action research, the process involves a number of cycles (Cohen et al. 2007: 306–307): identification and defining the problem to be tackled, preliminary discussion among the interest groups, setting of the objective, generation and moderation of trainer input elements, generation and moderation of tasks, assembling the input and tasks into a workable training sequence, piloting the program, moderation of the program on the basis of piloting feedback, implementation of the program, reflection and overall evaluation of the program. The data obtained in one stage of the program informed the development of the next stage (Coe et al. 2017: 73), e.g. decisions about the necessary trainee input resulted from the existing interviewer guidelines, an Education and Youth Board document; the nature of tasks was derived from the behavioural expectations outlined in the trainee input.

Although seemingly linear, the development of the program and the respective research was iterative (Coe et al. 2017: 73): all stages of the cycle were almost constantly revisited as the training program evolved, e.g. the initial plans were augmented in terms of the scope and time of training, and the training tasks were moderated based on reflection on the pre-testing and piloting results. The development of the training program involved close collaboration between university assessment and evaluation experts (n = 2) and the NE development team (n = 5).

4. Training program design: Process

The process of designing the NE examiner training program followed the sequence shown in Figure 2 below.

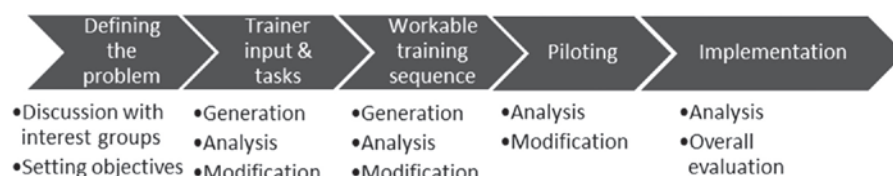


Figure 2. Program development process

Proceeding from the theoretical considerations outlined above, and with a special focus on the demands of the training program’s content, method, sequencing, and sociology, a team of experts was assembled, including assessment and evaluation experts (n = 2), NE developers (n = 5), computer-based training experts (n = 2) and a digital platform consultant (n = 1). The training was placed in the Examination Information System (EIS¹), a digital platform familiar to the prospective trainees as the medium where computer-based national standardised tests and admission tests are prepared, administered and assessed. The platform allows for the use of a variety of task types and is relatively easy to manage both for the training developer and the participants. As the training was the first of its kind to be delivered in the EIS environment, a basic linear, fixed-form test format was picked to further sustain ease of access.

The next step was to decide the overall format and structure of the training, the topics and their scope to be covered as well as the timeframe. The training was built as a course to be individually covered by a trainee where they would first familiarise themselves with the expectations set in the Education and Youth Board (2021) guidelines for examiner and rater behaviour during NE interviews. They would then be steered, through focused tasks, to reflect on the main features of both roles during the interview. Thus, the main structural units of the training – interviewer training and rater training – were set. The training could not be confined to either interviewer or assessor training, as all English teachers employed as NE examiners are expected to function as both, changing roles on different examination days.

Both the interviewer and the rater have key skills that have to be mastered for the speaking test to be managed in a standardised way, and for the student performances to be rated consistently, contributing to the reliability and the overall validity of the speaking test results. Identifying the key skills for the training program constituted the next phase. The key interviewer skills to be developed were set as follows: general guidelines to interviewer behaviour (their role, general demeanor, rate of speaking, managing instructions and questions, etc.), managing the speaking test tasks (giving instructions to its introduction, Task 1, Task 2, managing transitions from instruction to preparation, from preparation to task performance, managing follow-up questions to both tasks), managing frequent interviewer dilemmas (e.g. responding to student questions, managing situations where the student either says

¹ <https://eis.ekk.edu.ee/eis>

very little or is excessively talkative, etc.). The key rater skills comprised applying the general guidelines of rater behaviour during the interview, noticing and applying salient features in the marking scale for speaking (highlighted in the scale) and the actual practice of marking student performance. Next, the skills were grouped into clearly structured manageable segments (e.g. overall structure of the interview, managing the follow-up questions) to be covered in very short video segments by a trainer. This would constitute the training input.

Overall, the training was designed to follow the ‘modeling, coaching, and scaffolding’ (Boling et al. 2012) sequence involving the participants in a variety of problem-solving activities, all in an online format. Thus, steps were designed to start combining the trainees’ content knowledge with their technological knowledge (see the TPACK framework above). They first familiarise themselves with the interviewer/rater guidelines, viewing the short explanatory video presentations where the trainer highlighted the key skills at a particular stage of the interview or the salient features of the marking scale or those of the student performance (coaching and modeling). After that, the trainees completed video-based comprehension check and problem-solving tasks and finally had to reflect on the impact of the training session by completing a trainee feedback form. These scaffolding activities were planned so that the trainee would move from fairly straightforward reading, viewing and note-taking tasks to more demanding comprehension checks (true/false statements, multiple choice items, open questions, matching). As the training was the first of its kind, a linear approach was adopted as stated above; this means that all the trainees moved along the same training track, solving the tasks in exactly the same order as other trainees. Skipping training tasks and changing their order was not possible. The test developers’ decision to employ multiple choice tasks (with a varying number of distractors), open questions, matching and true and false statements as item types was taken for two reasons: the chosen platform (EIS) provided favourable conditions for creating and developing such tasks; the item types are familiar to the trainees and would thus not add to the challenge of completing the training online. After completing each task, the trainees received instant feedback on their success rate and could review their own solution to the task.

A considerable amount of time was spent working on sequencing the input and tasks. The first part of the training – the interviewer training sequence – was designed to follow the NE interview structure: trainer input followed by video-based interactive tasks focusing on managing the interview introduction, picture description (Task 1) and monologue (Task 2), as well as transition from one part of the interview to the next. The section has eight videos, each lasting between one and three minutes. The trainee then tests their own interviewer management skills by responding to five true/ false statements on the basis of a pre-recorded NE mock interview. Once all the individual parts of the interview had been discussed, the trainee was guided to complete reinforcement tasks – a multiple choice task to reflect on the reasons for particular interviewer responses and a matching task as a summary review. A further segment that was deemed necessary had emerged during earlier face-to-face training, namely non-standard student behaviour during an interview. This was built in as a reflection task and took the form of multiple choice questions. The content of the task also provided a logical transition to the second part of the training.

The second part focuses on the acquisition and honing of rating the students' speaking ability in the NE context. The approach to rater training was similar to that proposed in part one for the interviewers. The trainees rely on independent reading and video input to study the role and requirements set to raters. They then listen to the trainer discuss the intended interpretation of the marking scale for speaking, which serves as an introduction to the standardisation of marking. The second step in the standardisation procedure is watching a pre-assessed actual NE interview and analysing the reasons given for the marks awarded. For this activity to be possible during training, certain ethical hurdles needed to be cleared: consent was obtained from the interview participants with regard to the use of their interview for the training purposes. Similar consent was also needed for the tasks connected with the final part of the training. Here the trainee had to watch a complete NE interview and rate the student performance, accounting for the marks given.

The training course ended with feedback from the trainees. Because the training had already been quite lengthy, the feedback was designed so that it could be completed with minimum effort, yet allow genuine expression of opinion with concrete suggestions for improvement. Thus, background information was sought with the help of multiple choice and gap-filling tasks, e.g. teaching/NE examiner experience and the type of school they teach at. The trainees' response to the new training medium was obtained through 3 yes/no questions (use, ease of completion, logicity) and 9 open-ended sentences requesting their opinion as to whether the training met their expectations, what they found appealing, useful or otherwise and how the training could be improved.

It was estimated that completing both interviewer and rater training would take between 2 to 3 hours altogether. The time was calculated based on the number of aspects that needed consideration, the method of input delivery and the time required to reflect on the topics and complete the tasks. There was an attempt to make the course long enough to cover all the relevant features of both the interviewer's and the rater's work during the speaking test but not to make it overly long so the trainees would lose their motivation to complete the training. To allow for the trainees to adapt the training to their own training preferences, the course could be completed either in one session, or broken up into several shorter sessions. EIS saved all the answers, so if the trainee logged out and later in again, they could continue from where they had left off.

The process of item-writing within the task development underwent a number of moderation sessions: adjusting the length of the tasks, clarifying the wording of stems and distractors, editing the input videos, etc. Once a complete version of the program was decided on, it was pre-tested by two experts involved in NE development. Relying on their feedback, changes were made to the wording of task instructions, the order of items in the tasks and the feedback format. The designing process showed quite clearly that during the process, the training program designers had to rely on their assessment literacy content knowledge and combine it with their technological expertise and pedagogical know-how (see TPACK framework above) in order to generate training tasks to meet the overall training outcomes. The tasks had to be considered from both the didactic and the technological viewpoint in that while having particular instructional goals in mind, the tasks needed to allow learner independence and autonomy within the technological constraints of the program.

5. Implementation and feedback

The training program was launched in March 2021 and completed by 194 English teachers whose NE examiner and rater qualification needed either to be certified or renewed. 42 teachers, i.e. 22%, participated in examiner or rater training for the first time, 152 teachers, i.e. 78% had prior experience of being trained in a large group by a trainer in a face-to-face training session. The training program was open to trainees for 14 days, during which they could choose a training time and tempo which best suited their schedule. EIS allowed the trainees to complete the course in parts: the responses given to tasks were saved by the program and if the trainee decided to leave the program, they could resume their work from where they left off at a later date. All those who started the training also finished it, there were no incomplete training instances. The average time spent to complete the training was 2 hours and 42 minutes, which broadly corresponded to the time spent on face-to-face training earlier (three academic hours). The shortest time spent on completing the program was 36 minutes, with all the tasks completed without an error. This suggests that the trainee did not study the trainee input but went straight to task completion. Solving all the tasks without errors may indicate that the trainee was either a very proficient examiner/rater, or that the correct responses were obtained from a colleague. In the former case, there is clearly a need for more challenging training tasks for experienced examiners and raters, which would motivate them to refresh their examiner practices and, more importantly, hone their assessment skills. If very proficient, such examiners and raters could be involved in developing future training tasks or as mentors for the new trainees, embarking on the NE assessment career.

Although most trainees seemed to have completed the training within one session, there were 42 teachers who spent more than 24 hours to complete the training. There is one instance where the program remained unfinished for 9 days. Thus, the vast majority of trainees preferred to assign a specific time for the training and complete it within a single instance, thus maintaining their focus on the topic at hand. Because the two parts of the training – interviewing and assessment – are closely intertwined, this may have seemed a rational decision to keep the momentum and save time.

Trainee feedback to the training program was overwhelmingly very positive (e.g. ‘I simply enjoyed going through the training’, ‘I am very grateful that such a training opportunity was made available’, ‘it was such a good alternative to contact training’). 192 participants (of 194) thought that the training was very useful, all 194 remarked that the structure of the course was logical and 168 said that it was easy to complete. Some of the reasons given for the positive opinion echo the reasons given in earlier studies: an online solution as a comfortable way of obtaining a qualification in a tempo and at a time convenient for the trainee (see Elder et al. 2007, Knoch et al. 2016). Besides those, other reasons were listed: content and appropriateness of trainer commentary, clarity of the training program, the input provided by the videos, a good sequence of tasks and the nature of the tasks set.

There were participants, however, whose response to the training was less favourable. Two participants considered the training unnecessary, failing to see the need for examiner skills being periodically refreshed (‘I know it already’). Here,

the most fundamental understanding of necessary examiner characteristics seems to be missing: how the test reliability and validity is secured through maintaining examiner reliability and how the latter is achieved. In addition to those who could not see the need for training, there were 26 trainees (13%) who noted that the training was somewhat difficult. The difficulty was not content-related but was invariably connected to struggling with the medium of training, i.e. inadequate computer skills. This was a find that mirrored other similar earlier studies (see Hsu 2011, Knoch et al. 2016). Although none of the trainees left the training because they could not cope with the online medium, their training experience must have been considerably affected by it.

Trainee feedback provided a number of suggestions for examiner and rater training development. The most frequently quoted request was for a more enhanced rater training, which would include more commented student performances of different levels for the raters to develop a better idea of the standard. They also wished to see instances of how to respond to non-standard student responses as raters (e.g. long periods of silence, off-topic responses etc.) Not only commented samples were requested but also additional student performances for the trainees to rate were sought. Some trainees wished for an opportunity to discuss student performances with either the trainer or other colleagues as part of their training. This bespeaks one of the challenges of online training, voiced also in earlier studies (see Knoch et al. 2016): the relative loneliness of the trainee during such training and their longing for opportunities of peer support. Still other suggestions involved particular technical solutions: ability to follow the examiner script on screen while watching the training videos, viewing NE related forms and how to fill them, etc.

6. Summary and implications

The current action research stemmed from an urgent need to introduce a change in the training of NE examiners – from the constraints constituted by the required scope of training, the time available, the number of people needing (re)qualification, and the Covid restrictions – and to replace face-to-face examiner training by a respective online training.

The first research question – the nature of the process needed to develop an online examiners' training program – was answered relying on the findings of previous research combined with action research built on that. Research into earlier respective studies prompted a 'shift from a teaching to a learning paradigm' (Canals and Al-Rawashdeh 2019), where the responsibility for completing the training is predominantly with the trainee. Collins's Cognitive Apprenticeship Model (CAM) combined with the guidelines of the TPACK framework (2006) served as the basis for a self-study online program to certify NE examiners and raters. The four dimensions of the model – content, method, sequencing, and sociology – guided the elements as well as the tasks included in the training program. The key criteria guiding the development of training activities were modelling the expected behaviour, commenting on the performance, and offering supporting material to achieve the expected examiner and rater standard. Meaningful problem-solving tasks were created and set so that the trainees could complete them in their own time and

tempo, thus empowering them to take control of their own professional development. The development of the program went through a number of clear stages and was iterative in nature, forcing the developers to revisit the stages periodically.

The second research question focused on the trainee response to the training obtained with the help of the online program. The overwhelmingly positive trainee reaction justified the decision to switch from face-to-face to the online format. The trainee expectations of the content and the examiner knowledge provided and technological knowledge expected (see the TPACK framework above) thus seemed to have been mostly very appropriately estimated by the program developers. The participants seemed to develop the necessary technological content knowledge to successfully complete the course. Moreover, participating in the online training hopefully enhanced new combinations of knowledge proposed by the TPACK framework (ibid). The extent of the said growth should be determined in further studies.

The final focus of the current study was to explore the implications of the above trainee response. It was quite clear that online training comes with its own challenges. Although the course content was easily managed by the trainees and most trainees seemed to have adequate technological skills, there were aspects that would need attention while planning any future respective training: honing trainee general competence of working online so they would not divert attention away from the content of training, providing opportunities of trainee collaboration during training, perhaps having access to the trainer during training, increasing the input database for a better selection of training materials, etc. In the current training context, a need seemed to have arisen for developing a separate training program for examiners and raters, so the trainees could focus on their individual training needs more precisely. Here, too, considering the uneven development of the trainees' TPACK framework aspects – content, pedagogical and technological knowledge – the different training needs of the participants could perhaps be more advantageously met. All the lessons learned would be valuable considerations during the current training program development process and while designing any future similar online training program.

References

- Anderson, Hope M. 2018. *Blended Basic Language Courses: Design, Pedagogy and Implementation*. London–New York: Routledge. <https://doi.org/10.4324/9780203702468>
- Bluemel, Eberhard; Hintze, Axel; Schulz, Torsten; Schumann, Marco; Stuering, Stefan 2003. Virtual environments for the training of maintenance and service tasks. – *Proceedings of the 2003 Winter Simulation Conference*. Vol. 2. IEEE, 2001–2007. <https://doi.org/10.1109/WSC.2003.1261664>
- Boling, Erica C.; Hough, Mary; Krinsky, Hindi; Saleem, Hafiz; Stevens, Maggie 2012. Cutting the distance in distance education: Perspectives on what promotes positive, online learning experiences. – *The Internet and Higher Education*, 15 (2), 118–126. <https://doi.org/10.1016/j.iheduc.2011.11.006>
- Canals, Laia; Al-Rawashdeh, Amenah 2019. Teacher training and teachers' attitudes towards educational technology in the deployment of online English language courses in Jordan. – *Computer Assisted Language Learning*, 32 (7), 639–664. <https://doi.org/10.1080/09588221.2018.1531033>

- Caulfield, Jay 2011. *How to Design and Teach a Hybrid Course: Achieving Student-Centered Learning through Blended Classroom, Online and Experiential Activities*. Sterling, VA: Stylus Publishing.
- Coe, Robert; Waring, Michael; Hedges, Larry V.; Arthur, James (Eds.) 2017. *Research Methods and Methodologies in Education*. London: SAGE Publications.
- Cohen, Louis; Manion, Lawrence; Morrison, Keith 2007. *Research Methods in Education*. London–New York: Routledge. <https://doi.org/10.4324/9780203029053>
- Collins, Allan M. 2006. Cognitive apprenticeship. – R. Keith Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge: Cambridge University Press, 47–60. <https://doi.org/10.1017/CBO9780511816833.005>
- Education and Youth Board 2021. *Guidelines for the Oral Part of the Examination*. <https://tinyurl.com/4bb677n6> (17.10.2021).
- Elder, Cathie; Barkhuizen, Gary; Knoch, Ute; von Randow, Janet 2007. Evaluating rater responses to an online rater training program. – *Language Testing*, 24 (1), 37–64. <https://doi.org/10.1177/0265532207071511>
- Examination Information System (EIS). <https://eis.ekk.edu.ee/eis> (15.8.2021).
- Gustavsson, Bertil; Ohlsson, Per; Ullmark, Peter 2003. Educational methods for teaching vocational knowledge in a virtual environment. – *IFAC Proceedings*, 36 (22), 2003, 175–180. [https://doi.org/10.1016/S1474-6670\(17\)37713-3](https://doi.org/10.1016/S1474-6670(17)37713-3)
- Hsu, Pi-Sui 2012. Examining the impact of educational technology courses on pre-service teachers' development of technological pedagogical content knowledge. – *Teaching Education*, 23 (2), 195–213. <https://doi.org/10.1080/10476210.2011.622041>
- Kadel, Robert 2005. How teacher attitudes affect technology integration. – *Learning & Leading with Technology*, 32 (5), 34.
- Knoch, Ute; Fairbairn, Judith; Huisman, Annemiek 2016. An evaluation of an online rater training program for the speaking and writing sub-tests of the Aptis test. – *Papers in Language Testing and Assessment*, 5 (1), 90–106.
- Knoch, Ute; Read, John; von Randow, Janet 2007. Re-training raters online: How does it compare with face-to-face training? – *Assessing Writing*, 12 (1), 26–43. <https://doi.org/10.1016/j.asw.2007.04.001>
- Koehler, Matthew J.; Mishra, Punyashloke; Kereluik, Kristen; Shin, Tae Seob; Graham, Charles R. 2014. *The technological pedagogical content knowledge framework*. – *Handbook of Research on Educational Communications and Technology*. 4th ed. New York: Springer, 101–111. https://doi.org/10.1007/978-1-4614-3185-5_9
- Laherand, Meri-Liis 2008. *Kvalitatiivne uurimisviis [‘Qualitative Research’]*. Tallinn: OÜ Infotrikk.
- Malmköld, Lennart; Örtengren, Roland; Carlson, Blair E.; Nylén, Per 2007. Instructor based training versus computer based training: A comparative study. – *Educational Technology Systems*, 35 (4), 457–478. <https://doi.org/10.2190/L56G-6R11-0317-5047>
- Savage, Pamela A.; Kemp, David G.; Millen, David R.; Roberts, Linda A. 1990. The effectiveness of traditional versus computer-based training techniques. – *Proceedings of the Human Factors Society Annual Meeting*, 34 (3), 234–238. <https://doi.org/10.1177/154193129003400312>
- Yin, Robert K. 2009. *Case Study Research: Design and Methods*. CA: SAGE Publications.
- Zulaiha, Siti; Mulyono, Herri 2020. Exploring junior high school EFL teachers' training needs of assessment literacy. – *Cogent Education*, 7 (1), a1772943. <https://doi.org/10.1080/2331186X.2020.1772943>

RIIGIEKSAMI EKSAMINEERIJATE VEEBIPÕHISE KOOLITUSPROGRAMMI KOOSTAMISPROTSESS, RAKENDAMINE JA JÄRELDUSED

Ene Alas¹, Kristel Kriisa²

Tallinna Ülikool¹, Haridus- ja Noorteamet²

Artiklis kirjeldatakse kvalitatiivse uurimusena juhtumiuuringu vormis läbiviidud tegevusuuringut, milles analüüsitakse inglise keele riigieksami intervjuerijate ja hindajatate koolitamiseks loodud veebipõhise koolituse arendamist ja rakendamist Eestis. Tegemist on uuritud kontekstis esimese omataolise veebipõhise koolitusega. Koolituse raamistiku ja sisu väljatöötamisel, selle rakendamisel ja osaliste tagasiside kogumisel toetuti Collins'i Cognitive Apprenticeship mudelile ning Koehleri jt TPACK raamistiku elementidele. Protsessi analüüs tõi välja selles selgesti eristatavad etapid ning protsessi tsüklilise iseloomu. Tekkinud veebipõhine koolitusprogramm, mille sisu moodustasid videotel põhinev koolitussisend ning interaktiivsed eksamiintervjuu juhtimist ja õpilaste keeleoskuse hindamist puudutavad ülesanded, oli ühelt poolt märkimisväärselt edukas – ilmnas osavõtjate rahulolu koolituse kvaliteediga ja õppijate autonoomsuse hea tase – kuid osalejalt saadud tagasiside osutas ka mitmesugustele võimalustele koolitusprogrammi edasiseks täiendamiseks.

Võtmesõnad: hindamine, koolituse sisu, ülesannete koostamine, ülesannete järjestus, õppija autonoomsus, koolituse sotsioloogia, õppija tagasiside

Ene Alas (Tallinna Ülikool) on erialalt inglise filoloog, doktorikraad rakenduslingvistika alal. On uurinud akadeemilise kirjutamise hindamist, suulise keeleoskuse mõõtmist, üldist hindamisalast kirjaoskust ning lõimitud aine- ja keeleõpet. Narva mnt 25, 10120 Tallinn, Estonia
enealas@tlu.ee

Kristel Kriisa (Haridus- ja Noorteamet) on erialalt inglise filoloog, haridustehnoloog, magistrakraadid kirjaliku tõlke ja haridustehnoloogia alal. On uurinud keeleoskuse mõõtmist, õpetamismeetodeid ja arvutipõhist hindamist. Lõõtsa 4, 11415 Tallinn, Estonia
kristel.kriisa@harno.ee