A Lightweight Method for Modelling Technology-Enhanced Assessment Processes

Michael Striewe[a]
University of Duisburg-Essen, Essen, Germany

Keywords: Technology-enhanced Assessment Processes, Modelling Method, Evaluation, ESSENCE Standard.

Abstract: Conducting assessments is one of the core processes in educational institutions. It needs careful planning that can be supported by appropriate process models. Many existing assessment process models take a technical perspective and are not necessarily suitable for communication among educators or other people concerned with assessment organization. The paper reports on an alternative and more lightweight modelling approach and provides two sample process models for technology-enhanced assessments to illustrate its usage. Positive results from an evaluation of the modelling method in a workshop with eight participants demonstrate its suitability from the educator's perspective.

1 INTRODUCTION

Conducting educational assessments is surely one of the core processes in universities, schools and similar institutions. Even the simplest informal assessments involve at least two actors (examiner and examinee) and some conceptual objects (assessment items, feedback) that may materialize physically (i. e. as a piece of paper) or just exist virtually (i. e. verbally). In more advanced scenarios, administrative staff may join as an additional actor and more physical objects (rooms, lab equipment) may appear that need to be prepared as part of the assessment process. As a particular aspect of technology-enhanced assessments, an assessment systems takes over crucial duties in the process and thus many additional task related to that system appear. Finally, the general process for educational assessments may need to be aligned with other educational processes throughout a term and also with administrative processes of the institution.

Consequently, there are many recommendations to plan assessments carefully (Reynolds et al., 2009; Johnson and Johnson, 2002; Banta and Palomba, 2015; Dick et al., 2014), which essentially requires to define a process. Indeed there are several approaches and results on modelling actual processes at single institutions (Danson et al., 2001; Wölfert, 2015) as well as more generic or generalized assessment processes (Lu et al., 2013; Hajjej et al., 2016) using standard techniques for process modelling. All of these examples are motivated from the technical perspective of requirements engineering for technology-enhanced assessment systems. They are thus not necessarily suitable for communication among educators who are planning assessments without focus on the technical details. They also may miss process elements that are not related to the assessment system, but that are nevertheless part of the assessment process. This can also be seen when comparing these processes with the results from the FREMA framework (Millard et al., 2006; Wills et al., 2007) that collected assessment activities based on interviews with educators. However, the FREMA framework does not provide a facility to model such processes.

In previous work (Striewe, 2019), the author proposed a methodology for modelling educational assessment processes. That work raised two research questions: (1) Can that modelling method be used to model technology-enhanced assessment processes in a lightweight way independent of a particular assessment system in use? (2) Is the modelling method intuitive to use by practitioneers and usable for communication among them?

The current paper provides answers to these questions by two contributions: It provides examples for modelling two technology-enhanced assessment processes to answer the first research question. It also provides results from an evaluation of the modelling methodology in a workshop with staff from several different universities to answer the second research

[a] https://orcid.org/0000-0001-8866-6971

Striewe, M.
A Lightweight Method for Modelling Technology-Enhanced Assessment Processes.
DOI: 10.5220/0010879600003182
In *Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022) - Volume 1*, pages 149-156
ISBN: 978-989-758-562-3; ISSN: 2184-5026

question. The paper is organized as follows: Section 2 provides a short overview on the basic concepts used throughout the paper. Section 3 provides two sample process models for a summative and a formative assessment process and briefly discusses their similarities and differences. Section 4 reports on the evaluation of the modelling methodology that was carried out in a workshop with university staff.

2 BASIC CONCEPTS

2.1 Abstract Model Elements

From existing process descriptions a list of concrete elements and element types can be compiled which occur in at least some of the descriptions. These elements and types are candidates for elements to be included in a generalized and universal process model for educational assessment.

Table 1 lists three different actors that commonly appear in literature. They are listed with synonyms for their names that can be found in various publications. Some sources in literature make stronger differences between more actors which are aligned here to a more general set. The most important distinction is the one between people developing tests and people using tests, that can be found for example in (Reynolds et al., 2009) and (IMS-QTI, 2016). The idea is that domain experts create assessment items and compose meaningful tests from them, while teachers may use these tests to assess their students. Although there are surely many assessments conducted this way, there are also many in which teachers themselves author assessment items or at least amend items they picked from an item pool. Moreover, there are also assessments in which teachers use items or complete assessments they authored years ago. In these cases it is nearly impossible to draw a sharp border between people developing tests and people using tests. Consequently, it seems to be sufficient to have one actor who prepares and conducts the assessment and who may or may not be the author of the test items as well.

Other actors than the ones listed in table 1 also appear in the literature. However, they seem to be passive and only acting on demand of one of the actors named above. This includes staff like proctors or invigilators monitoring assessments, tutors helping students to review their results or technical staff helping to set up the assessment environment.

Table 2 lists five different concepts or objects that commonly take part in formalized assessment processes in literature. They are listed with synonyms for their names that can be found in various publications. There are more concepts that can be found in literature, but they do not appear to be common for assessment processes. This applies in particular to concepts related to physical objects such as exam sheets, which do neither appear in electronic assessments nor in oral assessments.

2.2 Modelling Language

The ESSENCE standard (Essence, 2015) defines a modelling language for process descriptions that is based on naming the key objects or concepts relevant in a process as well as their states they may be in. These key elements are called “alphas” in the standard and jointly form a so-called “kernel”. These are supposed to be relevant in any project. They form some kind of basic building blocks that allow to start with defining process right ahead without thinking to much (and possibly missing) about actors or objects to be included. Although the ESSENCE standard originally is about software engineering, there is neither a technical need to stick to the kernel defined by the standard, nor to apply the modelling language only to software engineering processes.

Each alpha defines an ordered set of states with checklists that allow to track project progress. Simple process descriptions can be created by grouping states across alphas and thus defining phases or milestones. As a means of graphical representation, the ESSENCE standard introduces the notion of alpha state cards. They are concise representations of an alpha state and its checklist items that can actually be used in form of small physical cards.

2.3 Assessment Kernel

The proposed kernel consists of eight alphas from which two are optional. Table 3 provides and overview on these alphas and their states. The most important aspects for each alpha can be summarized as follows:

A *test item* is the smallest consistent unit within an assessment that allows candidates to demonstrate their competencies. A test item contains a task description and candidates are expected to respond to it in some way. The alpha states reflect that test items have some formal properties (such as an item type or language) which are defined in the first state, while their functional properties (such as a task description and a sample solution) are defined in the second state. The third state handles verification and double-checking. The fourth state reflects the didactic practice to review the outcomes of a test with respect to

Table 1: Different actors found in formal assessment process descriptions in literature.

Actor Name and Synonyms	Short Description	References in Literature
Student (also: Candidate, Learner, Test-taker)	A person who is supposed to sit the exam and to answer questions in the test.	(Johnson and Johnson, 2002) (Banta and Palomba, 2015) (Sindre and Vegendla, 2015) (Sclater and Howie, 2003) (Kiy et al., 2016) (Wills et al., 2007) (Hajjej et al., 2016) (Cholez et al., 2010) (Reynolds et al., 2009) (Gusev et al., 2013) (Küppers et al., 2017) (Lu et al., 2013) (Kaiiali et al., 2016) (Pardo, 2002) (IMS-QTI, 2016) (Tremblay et al., 2008) (Moccozet et al., 2017) (Lu et al., 2014)
Teacher (also: Author, Examiner, Faculty, Instructor, Professor)	A person who prepares and conducts assessments and decides about grades and feedback. May also be the one who creates assessment items and designs tests.	(Johnson and Johnson, 2002) (Banta and Palomba, 2015) (Sindre and Vegendla, 2015) (Sclater and Howie, 2003) (Kiy et al., 2016) (Wills et al., 2007) (Hajjej et al., 2016) (Reynolds et al., 2009) (Gusev et al., 2013) (Lu et al., 2013) (Kaiiali et al., 2016) (Pardo, 2002) (IMS-QTI, 2016) (Tremblay et al., 2008) (Moccozet et al., 2017) (Lu et al., 2014)
Exam Authorities (also: Exam Office, Departmental Secretary)	An institution responsible for formal or organizational aspects of assessments.	(Sindre and Vegendla, 2015) (Sclater and Howie, 2003) (Kiy et al., 2016) (Wills et al., 2007)

Table 2: Different objects or concepts found in formal assessment process descriptions in literature.

Concept Name and Synonyms	Short Description	References in Literature
Question (also: Assessment Item, Test Item, Assignment)	A single item within an exam which can be answered by a student independently of other items.	(Johnson and Johnson, 2002) (Banta and Palomba, 2015) (Sclater and Howie, 2003) (Wills et al., 2007) (Cholez et al., 2010) (Reynolds et al., 2009) (Lu et al., 2013) (Kaiiali et al., 2016) (IMS-QTI, 2016) (Tremblay et al., 2008) (Moccozet et al., 2017) (Lu et al., 2014)
Exam (also: Test, Question Set, Assessment, Quiz, Test paper)	A collection of questions that is delivered to the students.	(Johnson and Johnson, 2002) (Banta and Palomba, 2015) (Sindre and Vegendla, 2015) (Sclater and Howie, 2003) (Kiy et al., 2016) (Wills et al., 2007) (Hajjej et al., 2016) (Reynolds et al., 2009) (Küppers et al., 2017) (Lu et al., 2013) (Kaiiali et al., 2016) (IMS-QTI, 2016) (Moccozet et al., 2017) (Lu et al., 2014)
E-Assessment System (also: Digital Environment, Tool, Exam Server)	An electronic system used in the assessment process mainly for delivering exams, collecting responses or creating feedback.	(Anderson et al., 2005) (Wills et al., 2007) (Cholez et al., 2010) (Kaiiali et al., 2016) (Pardo, 2002) (IMS-QTI, 2016) (Moccozet et al., 2017) (also in (Hajjej et al., 2016) and (Küppers et al., 2017) as actor)
Room (also: Physical Environment)	The location where students are supposed to be while sitting the exam.	(Sindre and Vegendla, 2015) (Wills et al., 2007) (Wölfert, 2015) (Lu et al., 2013) (Kaiiali et al., 2016)
Feedback (also: Grade, Score, Results)	The pieces of information produced to describe and inform about the exam results.	(Johnson and Johnson, 2002) (Banta and Palomba, 2015) (Sindre and Vegendla, 2015) (Wills et al., 2007) (Wölfert, 2015) (Hajjej et al., 2016) (Cholez et al., 2010) (Kaiiali et al., 2016) (Pardo, 2002) (IMS-QTI, 2016) (Tremblay et al., 2008) (Moccozet et al., 2017)

Table 3: Overview on the eight kernel alphas and their states.

Alpha “Test Item” 1. Scoped 2. Designed 3. Verified 4. Outcome reviewed	Alpha “Test” 1. Goals clarified 2. Designed 3. Generated 4. Conducted 5. Evaluated
Alpha “Grades and Feedback” 1. Granularity decided 2. Prepared 3. Generated 4. Published	Alpha “Organizers” 1. Identified 2. Working 3. Satisfied for start 4. Satisfied for closing
Alpha “Candidates” 1. Scoped 2. Selected 3. Invited 4. Present 5. Dismissed 6. Informed 7. Satisfied	Alpha “Authorities” 1. Identified 2. Involved 3. Satisfied for start 4. Satisfied for closing
Alpha “Location” 1. Defined 2. Selected 3. Reserved 4. Prepared 5. In use 6. Left	Alpha “System” 1. Defined 2. Selected 3. Available 4. Ready for start 5. In use 6. Ready for closing

test item performance in order to identify test items with unexpected results.

A *test* is a collection of test items that is delivered to the candidates of the assessment. The alpha refers to the test as an abstract construct and does not ask whether the test is a static composition of test items or generated adaptively. The first and second state correspond to the first two states of the alpha for test items, as also the whole test needs both a definition of its formal and functional properties. The third state is fulfilled when an actual instance of the test is created for each candidate. The fourth state is fulfilled when all candidates have completed their tests. The fifth state represents the fact that a test needs to be evaluated and also includes the retrospective analysis of test item performance as above.

As the outcome of test evaluation can be very different depending on the purpose and context of an assessment, *grades and feedback* form a separate alpha. Each response to a test item contributes to the test result which may consist of marks, credit points, texts or anything else which is used to inform the candidates about their performance. Again, the first two are concerned with preparations: The first state reflects the fact that there are many ways of giving feedback and that the purpose of the assessment determines the choice. The second state refers to the creation of appropriate marking schemes or alike as well as organizational set-up of grading sessions or configuration of an automated assessment system. The third state is fulfilled if all grades and feedback are created. The final state is fulfilled when grades and feedback are available to the candidates.

For each assessment there is at least one person responsible for organizing it. For larger assessments the group of *organizers* may include more people like test item authors, assessors and technical staff. The first state represents the fact that it may require some work to find out who needs to be involved into the assessment for which tasks. The second state is fulfilled when all responsible persons have picked up their duties. Once they have done everything that is required to start the actual assessment, the third state is reached. Similarly, the final state is reached when all evaluation and post-processing is done and the organizers have no more open duties.

The largest group of people concerned with an assessment are usually the *candidates*. They are involved personally in the assessment process for a relatively short period of time. The first two states refer to the part of the process in which it is first defined who is allowed to take part in the assessment and secondly the actual persons are identified. The third state is fulfilled when candidates know how to prepare themselves for the assessment. The following two states refer to the physical presence of the candidate at the location where the assessment takes place. The sixth and seventh state reflect the fact that candidates need explicitly to be informed about their results and get some time to place complaints before the grades formally count as accepted.

In some scenarios, an official party may be formally responsible for legal issues related to conducting the assessment. As this may introduce additional process steps or dependencies between states, *authorities* are introduced as an additional optional alpha in the kernel. The states are almost similar to the ones of the organizers with a subtle difference in the naming of the second state.

Each assessment needs some physical *location* where candidates will be located while taking part in the assessment, even if they are not all in the same place. Quite similar to the states for candidates, the first two states for the location refer to the fact that first some abstract requirements are formulated towards the properties of the assessment location and then an actual room or set of rooms is selected. As rooms are physical resources, they may cause conflicts with other assessments happening at the same time. Hence the third state is explicitly introduced to cover the necessary communication. If all set-up is done, the alpha reaches the fourth state. The final two states correspond to some extent to state five and six for the candidates but also cover the fact that the location needs to be restored after the assessment.

If a computer-aided assessment *system* is used, it can be represented by an additional alpha. It covers all possible duties of the system such as administering the tests or performing grade and feedback generation automatically. Similar to the previous alpha, the first two states reflect the fact that (at least in an ideal scenario) one would first define some abstract requirements towards the assessment system and then select an actual system. In reality, organizers sometimes have no choice and must use the system provided by their institution. In that case, these two states are fulfilled by default. The third state refers to the fact that the selected system also needs to be accessible to continue preparation in state four. The fifth state models the period of time in which candidates interact with the system. This is also the period of time in which it performs tasks like automated grading on its own. The final state makes no assumptions on whether the whole system will actually be closed or whether it is just the assessment that is closed and archived.

2.4 State and Phase based Models

Processes can be described by chaining alpha states in the order they have to be reached. One way of doing so is to define process phases and group all alpha states belonging into the same phase. One phase in such models can cover more than one state of a single alpha, but there may also be alphas that do not contribute one of their states for a particular phase. The idea of using phases as a means of structuring a process model is a common concept in process modelling and has also been used in several papers on assessment processes (e. g. (Wölfert, 2015; Lu et al., 2013; Moccozet et al., 2017)).

For the cases studies presented in the next section, up to five different phases are used: (1) “Planning” for the conceptual and theoretical preparations of an assessment, (2) “Construction” for the practical preparations of an assessment, (3) “Conduction” for the phase where participants work on the assessment, (4) “Evaluation” for the phase in which grades and feedback are produced, and (5) “Review” for any remaining things steps. Neither of them has to be considered mandatory for assessment process descriptions. Similarly, a process description may also add an additional phase if necessary. The names of each phase may change, if phases are removed or added.

3 SAMPLE PROCESS MODELS

This section provides two process models for educational assessments to demonstrate how processes can be modelled by arranging alpha state into different phases and by skipping single states or complete alphas.

3.1 Case 1: A Summative Electronic Assessment

This case study considers an computer-aided exam or alike. It assumes that candidates come to the exam hall which is equipped with appropriate systems for the purpose of publishing the test and collecting submission. It also assumes that there is no need to provide direct feedback to the candidates while they are present in the exam hall. Grading of the solutions can thus happen asynchronously (Striewe, 2021). This case has thus the following characteristics: First, all alphas including the optional ones must used, as we employ an electronic system and involve the exam authorities. Second, we can use all five phases suggested above, as we can clearly separate the conduction phase from the evaluation phase. The resulting process description is depicted in figure 1.

The first phase contains almost only the first state for each of the alphas, as it is concerned with planning but not with practical preparations. Some of the checkpoints of these states may be fulfilled right from the beginning, such as the language used in the assessment or the organizers that are involved. Depending on the habits in a particular institution it may also happen that some checkpoints or even states from the next phase are also fulfilled right from the beginning (e. g. for alpha “Test Items”, if the assessment is based on a pre-defined item pool). However, there is no immediate need to shift the respective states to the first phase for that reason.

The construction phase also has contributions from all alphas. When all states in this phase are fulfilled, everything is ready to start the actual assessment. Notably, state “Generated” from alpha “Test” is located in the construction phase, since we assume in this case study that the test is not created dynamically for each individual candidate. To handle adaptive e-assessments, the state needs to be moved to the conduction phase as case study 2 will show.

The conduction phase has only contributions from four alphas. This is not surprising, as test items, or-

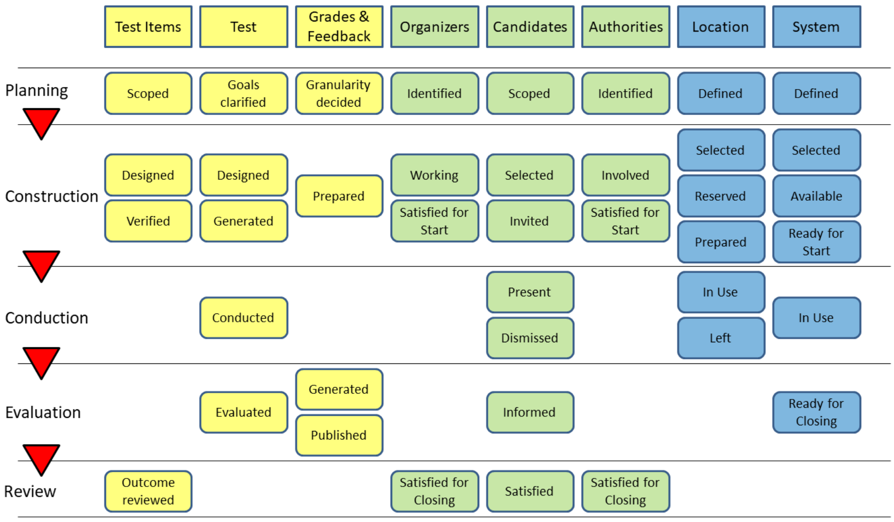

Figure 1: Overview on the assessment process for a summative e-assessment using five phases. The process assumes the application of asynchronous grading, so evaluation happens in a separate phase after conduction.

ganizers and authorities are not supposed to change their state while the assessment is conducted. Alpha "Location" already reaches its final state, as the location is not supposed to be involved in asynchronous grading or review. Consequently, the evaluation phase also has contributions from just four alphas. Three of them also reach their final state in this phase. One of them is the assessment system, as we assume that it is not needed for the review of results. For scenarios in which this assumption is not true, the final state can be shifted to the review phase.

Notably, we can skip the alpha "System" from the process and retain a process that represents a traditional written exam which is graded manually after conduction.

3.2 Case 2: A Distributed Formative Electronic Assessment

The second case study looks at an e-assessment where participants can work from at home on a formative assessment (like homework assignments). We assume that candidates are allowed to make submissions, receive immediate feedback from the system and can improve their previous answers or proceed with subsequent tasks. We also assume that the content of the exercises is to some extend generated dynamically, e. g. by randomization of variables. Notably, this scenario has been sketched before the Covid19 pandemic, but only needs small amendments to cover summative assessments conducted from at home.

Again we can identify specific characteristics of this scenario: As this scenario does not represent a formal exam, we do not need to include the alpha "Authorities". As we use direct feedback and questions that are generated dynamically, it is suitable to join conduction and evaluation phase. Notably, we do not need to pay special attention to alpha "Location", although the scenario does not define a single physical location in which the candidates will meet. Instead, we make use of the fact that "Location" is defined abstract enough to represent a physically distributed location which virtually consists of the private workplaces from where the candidates take part in the assessment. The resulting process description is depicted in figure 2.

The planning phase is the same as it was in the previous case. The construction phase shows two differences: State "Generated" for alpha "Test" has been moved to the next phase. As already discussed above, this reflects the fact that contents of the test are generated individually for each participant. The second difference concerns state "Prepared" of alpha "Location". The fact that this scenario considers a physically distributed location is the reason for placing this state in the conduction phase. In a distributed formative assessment, there is no possibility to ensure that all candidates have completed to set up their personal workspace before the assessment starts. Hence individual workspace preparations may happen while other candidates are already submitting solutions or even have finished the test.

The reasoning for joining conduction and evaluation into one phase was already discussed above. From the four alphas contributing to this phase, three also reach their final state in this phase. This also matches the expectation that test, grades and feedback, and location will not change their state after the

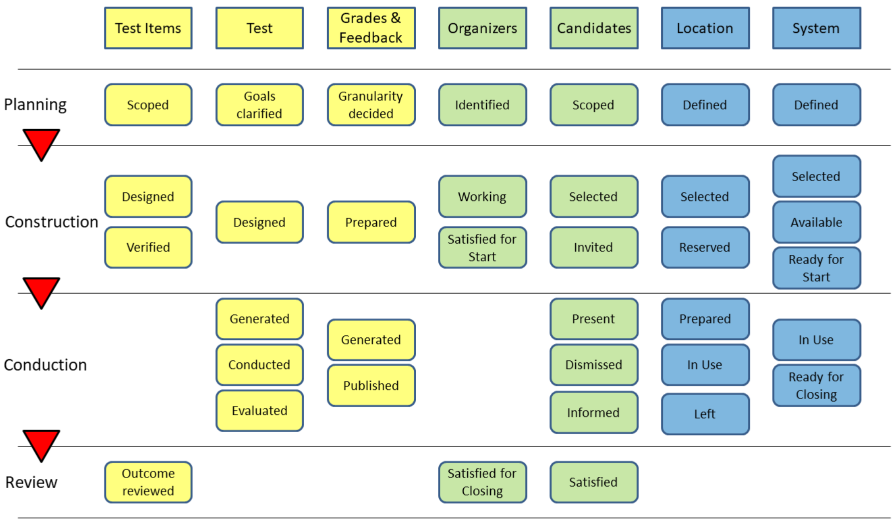

Figure 2: Overview on the assessment process for a distributed formative a-assessment using four phases. The formative setting allows to skip the alpha “Authorities” from the process description. Alpha “Location” is included, although candidates are not required to show up at the same physical location.

assessment has been conducted and evaluated. Hence the remaining review phase only contains the final states of the remaining alphas.

4 EVALUATION WORKSHOP

The method for modelling assessment processes has been evaluated at a workshop at a symposium on electronic assessment. Eight persons from eight different academic institutions took part in the evaluation. All participants were academic staff with experience in organizing and conducting assessments.

The workshop took 90 minutes. Within the first 35 minutes, participants received an introductory presentation about the modelling language, the assessment kernel, and the concept of state and phase based models. One sample process was included in the presentation. The participants were free to ask questions on any aspect. A total amount of approx. 10 minutes throughout the presentation was used for that. For the next 45 minutes, participants split into four pairs working independently. Each team received a set of physical cards for the alpha states and was asked to model at least one assessment process from one of their institutions. Finally, 10 minutes were used to share experiences among all participants. Since eight persons is a quite low sample size only qualitative results were recorded.

All participants perceived the modelling language and method as useful and usable. Participants liked the idea of using state and phase based models as a monitoring tool for actual assessments as well as a documentation aid on the abstract level. Moreover, the participants understood the models as a kind of multi-dimensional checklist that is easy to grasp.

Working in small teams with physical cards on the table was perceived as a good way to start structuring assessment processes. One participant explicitly stated to think about using the same method for an internal workshop on assessment planning with colleagues at their institution. Notably, none of the four sets of physical cards was returned to the workshop organizer, but participants took all with them to continue using them at their institutions. However, one participant stated that it might be necessary to translate the initial models into some other modelling language later on to monitor running processes.

One team created additional kernel elements during the working phase to come as close as possible to their local situation. They added an additional state in which organizers are on stand-by during the conduction of a written exam, since at their institution written exams are usually proctored by non-academic staff. They also added their assessment support unit as another alpha with several states throughout the process. Finally, they also added an additional phase to their process used for consultations between the support unit and academic staff.

5 CONCLUSIONS AND FUTURE WORK

The paper presented a recap of an existing modelling method for educational processes. The modelling method was evaluated to be beneficial at a workshop. It can be concluded that the modelling method is suitable to model one of the core processes in universities, schools and similar institutions. While this is a positive result, a more detailed analysis of the understandability of such process models is surely possible. Since understandability of conceptual models can be measured in several dimensions, empirical studies with quantitative results can be used to enrich the existing qualitative results.

Another area of future work is the value of the process models for quality assurance. On the one hand, the process models may help to eliminate weaknesses within the processes. On the other hand, process models can help to evaluate tools by the degree of process coverage they offer. That can help to find aspect that are not covered by any tool, but also conflicts when two tools are used within the same process with overlapping duties.

REFERENCES

Anderson, H. M., Anaya, G., Bird, E., and Moore, D. L. (2005). A Review of Educational Assessment. *American Journal of Pharmaceutical Education*, 69(1).

Banta, T. W. and Palomba, C. A. (2015). *Assessment Essentials*. John Wiley & Sons Inc, 2nd edition.

Cholez, H., Mayer, N., and Latour, T. (2010). Information Security Risk Management in Computer-Assisted Assessment Systems: First Step in Addressing Contextual Diversity. In *Proceedings of the 13th Computer-Assisted Assessment Conference (CAA 2010)*.

Danson, M., Dawson, B., and Baseley, T. (2001). Large Scale Implementation of Question Mark Perception (V2.5) – Experiences at Loughborough University. In *Proceedings of the 5th Computer-Assisted Assessment Conference (CAA)*.

Dick, W., Carey, L., and Carey, J. O. (2014). *The Systematic Design of Instruction*. Pearson Education, 8th edition.

Essence (2015). Essence - Kernel and Language for Software Engineering Methods. http://www.omg.org/spec/Essence/1.1.

Gusev, M., Ristov, S., Armenski, G., Velkoski, G., and Bozinoski, K. (2013). E-Assessment Cloud Solution: Architecture, Organization and Cost Model. *iJET*, 8(Special Issue 2):55–64.

Hajjej, F., Hlaoui, Y. B., and Ayed, L. J. B. (2016). A Generic E-Assessment Process Development Based on Reverse Engineering and Cloud Services. In *29th International Conference on Software Engineering Education and Training (CSEET)*, pages 157–165.

IMS-QTI (2016). IMS Question & Test Interoperability Specification. http://www.imsglobal.org/question/.

Johnson, D. H. and Johnson, R. T. (2002). *Meaningful Assessment: A Manageable and Cooperative Process*. Pearson.

Kaiiali, M., Ozkaya, A., Altun, H., Haddad, H., and Alier, M. (2016). Designing a Secure Exam Management System (SEMS) for M-Learning Environments. *TLT*, 9(3):258–271.

Kiy, A., Wölfert, V., and Lucke, U. (2016). Technische Unterstützung zur Durchführung von Massenklausuren. In *Die 14. E-Learning Fachtagung Informatik (DeLFI 2016)*.

Küppers, B., Politze, M., and Schroeder, U. (2017). Reliable e-Assessment with GIT - Practical Considerations and Implementation. In *EUNIS 23rd Annual Congress*.

Lu, R., Liu, H., and Liu, B. (2014). Research and implementation of general online examination system. *Advanced Materials Research*, 926-930:2374–2377.

Lu, Y., Yang, Y., Chang, P., and Yang, C. (2013). The design and implementation of intelligent assessment management system. In *IEEE Global Engineering Education Conference, EDUCON 2013, Berlin, Germany, March 13-15, 2013*, pages 451–457.

Millard, D. E., Bailey, C., Davis, H. C., Gilbert, L., Howard, Y., and Wills, G. (2006). The e-Learning Assessment Landscape. In *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pages 964–966.

Moccozet, L., Benkacem, O., and Burgi, P.-Y. (2017). Towards a Technology-Enhanced Assessment Service in Higher Education. In *Interactive Collaborative Learning*, pages 453–467. Springer International Publishing.

Pardo, A. (2002). A Multi-agent Platform for Automatic Assignment Management. In *Proceedings of the 7th Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '02, pages 60–64. ACM.

Reynolds, C., Livingston, R., and Willson, V. (2009). *Measurement and Assessment in Education*. Alternative eText Formats Series. Pearson.

Sclater, N. and Howie, K. (2003). User Requirements of the "Ultimate" Online Assessment Engine. *Comput. Educ.*, 40(3):285–306.

Sindre, G. and Vegendla, A. (2015). E-exams and exam process improvement. In *Proceedings of the UDIT / NIK 2015 conference*.

Striewe, M. (2019). Lean and Agile Assessment Workflows. In *Agile and Lean Concepts for Teaching and Learning: Bringing Methodologies from Industry to the Classroom*, pages 187–204. Springer Singapore.

Striewe, M. (2021). Design Patterns for Submission Evaluation within E-Assessment Systems. In *26th European Conference on Pattern Languages of Programs*, EuroPLoP'21, pages 32:1–32:10.

Tremblay, G., Guérin, F., Pons, A., and Salah, A. (2008). Oto, a generic and extensible tool for marking programming assignments. *Software: Practice and Experience*, 38(3):307–333.

Wills, G. B., Bailey, C. P., Davis, H. C., Gilbert, L., Howard, Y., Jeyes, S., Millard, D. E., Price, J., Sclater, N., Sherratt, R., Tulloch, I., and Young, R. (2007). An e-Learning Framework for Assessment (FREMA). In *Proceedings of the 11th Computer-Assisted Assessment Conference (CAA)*.

Wölfert, V. (2015). Technische unterstützung zur durchführung von massenklausuren. Master's thesis, Universität Potsdam.