Supporting learners in a crisis context with smart self-assessment

Zacharoula Papamitsiou, Martin Lunde, Jakob Westermoen and Michail N. Giannakos.

Abstract Sustaining learning in the times of a global crisis is complex. It's beyond the previously employed online learning and teaching approaches: this new setting brings on the same page the challenge to support learners not to lose their motivation and interest in learning, and the opportunity for new learning and teaching formats to emerge and be developed. It also makes space for reconsidering autonomous learning as a choice for maintaining the inherent need for self-determination. This study presents how students' motivation is affecting the usage of an online self-assessment service – enhanced with analytics – and the different perspectives of more vs. less motivated learners to continue their learning, despite the contextual shift due to the covid-19 crisis.

1 Introduction

The Covid-19 pandemic is an ongoing crisis that has caused the disruption of normal educational processes and has raised new challenges about the way we learn. In such conditions, it is important to keep in mind that education is linked to crises in three main stages [35]: in the prevention of a crisis (e.g., familiarize with alternative teaching formats), during a crisis (e.g., sustaining online teaching/learning formats) and in post crisis (e.g., safe reopening of schools). Previous research on educational technologies had already developed and practically demonstrated methods and tools for carrying out and supporting online teaching and learning formats (e.g., 2012 was announced as the "year of the Massive Open Online Course (MOOC)"[31]). However, in practice, it is shown during the Covid-19 crisis, that the "real" world was not prepared to switch to fully online learning and teaching; as emphasized by UNESCO's Assistant Director-General for Education [16, p. 4]: "We need to

Zacharoula Papamitsiou, Martin Lunde, Jakob Westermoen and Michail N. Giannakos.
Norwegian University of Science and Technology, Trondheim, Norway,
e-mail: zacharoula.papamitsiou@ntnu.no

come together not only to address the immediate educational consequences of this unprecedented crisis, but to build up the longer-term resilience of education systems." In a sense, learning and teaching in times of crises require seizing opportunities for creating spaces to rethink, pay attention and reflect on the sustainability of alternative educational approaches (i.e., "come together not only to address the immediate educational consequences"), as well as for novel methods and tools to be coupled with the use of modern technologies (i.e., "build up the longer-term resilience of education systems"). In other words, ensuring that we learn from and fully exploit existing learning and teaching practices shifts from a *nice-to-have* to a *must-have*.

When disruptions happen – like the one we experienced due to the Covid-19 crisis – we often speak of restoring a sense of educational "normality" for everybody. It is necessary to devise special methods of education delivery *for all* who have come through the crisis; there is an opportunity here, so new technologies can play a pivotal role in continuing educating everybody in the aftermath of a global crisis, by investing on individuals' motivation to learn and their autonomous learning capabilities.

We present the results from a study that was conducted during the Covid-19 pandemic in Norway – the first part of the study was completed the day the lock-down was announced, and the second part was conducted during the lock-down. The study aimed to explore students' motivation to use an online self-assessment service – enhanced with analytics – and the different perspectives of more vs. less motivated learners to continue their learning, despite the contextual shift due to the crisis.

2 Background and Related Work

2.1 Learning in a crisis context: an autonomous learning perspective

The outbreak of Covid-19 led the governments worldwide to take drastic measures and impose restrictions to deal with and control the spread of the virus; 107 countries had implemented national schools and universities closures by March 18, 2020, and switched to online teaching and learning [38]. However, it's ground truth that, in online conditions, strong self-regulated learning (SRL) skills are required in order the learning to be efficient (e.g., [4, 40]). SRL is conceptualized as an "active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features in the environment" [32, p. 453]. Self-regulated learners are guided by their motivation to learn, are aware of their learning processes and adjust their behaviour to keep themselves on track towards their desired outcomes [10, 22, 32, 33]. Researchers indicated that students with strong SRL skills were more likely to be successful in online learning [4]. These studies and the development of adequate tools for self-assessment of learners have become necessary to guarantee good performance in e-learning environments [26, 37] especially in crisis times such as the Covid-19 pandemic. Particularly, empirical research conducted during the Covid-19 outbreak, showed that students' learning motivation was highly correlated with academic achievements and autonomous learning [41].

Indeed, practising SRL has been proposed to develop learning autonomy [28]. In learning contexts, autonomy is experienced as an implicit need "to take charge of one's learning [...] the responsibility for all the decisions concerning all aspects of this learning [...]" [14, p. 3]. This definition assumes accepting responsibility over all spectrum of the learning process, regardless of the learning context or the specifications of the learning environment; autonomy places the learners at the outset of the learning tasks – it is always the learners who choose and control when, what, where, how and how much to learn. As such, learners' responsible self-initiative is prioritized, allowing them to determine what will be learned and to critically reflect on the selected learning tasks [6]: the autonomous learners are able to unravel their own learning issues and *to define what needs to be learned and how*. It has been argued that the capacity to control learning (i.e., autonomy) embraces learners' desire (motivation), ability (knowledge and skills to plan, monitor, evaluate learning), and freedom (permission to control) to do so [15]. As core dimension of learning autonomy, learners' motivation and it's regulation need to be further strengthen in times of crisis and educational disruptions, in order to be sustained, and the learners need to be supported so that they do not procrastinate or lose their interest in learning.

2.2 On continuous adaptive self-assessment

As already mentioned, autonomous and self-regulated learning are closely related to learners' motivation and require that learners periodically engage themselves in self-assessment. Iterative self-assessment leads students to a greater awareness, by training them to self-regulate their motivation and behaviour, as well as by fostering reflection on their own progress in knowledge or skills, and finally, to understanding themselves as learners [23]. Students who take practice self-assessment often outperform students in non-assessment conditions such as restudying, practice, or filler activities [1]. It is ground truth that retrieval practice (i.e., calling information to mind rather than rereading it or hearing it, in order to trigger "an effort from within" to induce better retention) is better at reinforcing knowledge than restudying information, and that testing is a good way to activate this retrieval process [8]. Research has provided evidence that multiple-choice testing had the power to stabilize access to marginal knowledge, highlighting how relatively simple it is to reactivate and consolidate knowledge [7], and at the same time, a growing number of studies on this topic have reported robust benefits of testing on transfer of learning [9].

To further support individuals to maintain their learning motivation, the activities need to be tailored to fit learners' mastery levels, and feedback can support their personalized needs, i.e., introduce adaptivity in the learning and assessment settings. According to the 2019 NMC Horizon Report, adaptive learning is a "breakthrough teaching model of the future" that needs to be scaled to its potential [2]. In adaptive contexts the underlying learners' motivation is amplified and encouraged via adaptation: the connection between motivation and on-task engagement is catalyzed by the personalization of the learning experience and the provided feedback. The adaptive activities aim to encourage learners to stay motivated, i.e., to feed their autonomous desire to learn. By deriving suitable adaptation mechanisms, the learning process is controlled in a way that meets learners' motivation, whereby motivation is often considered an impetus for engagement in learning [21]. Indeed, motivation and theories of goal orientation can help to explain the reasons for students' engagement in a task [32]. Goals are considered a facet of motivation given that they provide a purpose or focus for the task and thus, influence students' learning behaviors [13].

2.3 Motivation of the research and research question

Synopsizing the above, it becomes apparent that the educational disruption due to crisis like the Covid-19 pandemic may result in demotivating learners with overall unwanted consequences for their learning. It has been found that students with strong SRL skills, i.e., students capable of regulating their motivation (among others), can continue learning in an efficient way, even when the set-up is fully online. Maintaining and supporting learners' motivation is critical. To achieve that objective, periodically practising adaptive self-assessment test is a promising approach. This study is motivated exactly by that idea, and the research question (RQ) is as follows:

RQ: *How learners' motivation is affecting their dispositions to use an online adaptive self-assessment service, and continue learning during a crisis?*

For addressing the RQ, we conducted a study at a Norwegian University, using an adaptive self-assessment service (briefly described in next section). The service is enhanced with analytics that can utilize students' data and automatically build visual representations of students' learning progress and provide feedback in different learning analytics reports [3, 29]. The study was conducted in two phases: the first one – i.e., the pilot usage of the service – was completed on the day that the lock-down was announced, and the second one – i.e., the participants' interviews about their motivation to use the service – was conducted during the lock-down.

3 Brief introduction to SmartU

Students' self-assessment data were collected with SmartU (**S**elf-assessment **M**easured with **A**nalytics on **R**un-**T**ime for YO**U**), an online dedicated service for adaptive self-assessment that consists of (a) a dashboard interface, (b) an adaptation mechanism, (c) a tracker that logs interaction data, and (d) a database storing information about the students and questions. SmartU is a revised version of a previous service [30].

The main interface consists of a dashboard that displays the available activities (i.e., the different self-assessment tests for the different courses) (Fig. 1); the learner can select one of those activities, and perform it multiple times ("attempts"). As shown in Fig. 2, the learner can see descriptive information about her progress in the activity, the results from the previous attempts, the progression on the overall response-times, additional analytics about her self-regulation (e.g., average time-spent, overall effort) for building her self-awareness, and relevant announcements. The learner can also see descriptive information in comparison to her peers (Fig. 3).

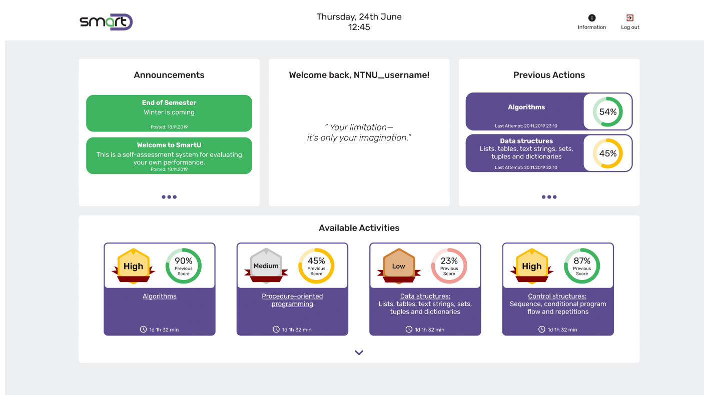

Fig. 1 smartU – Main dashboard

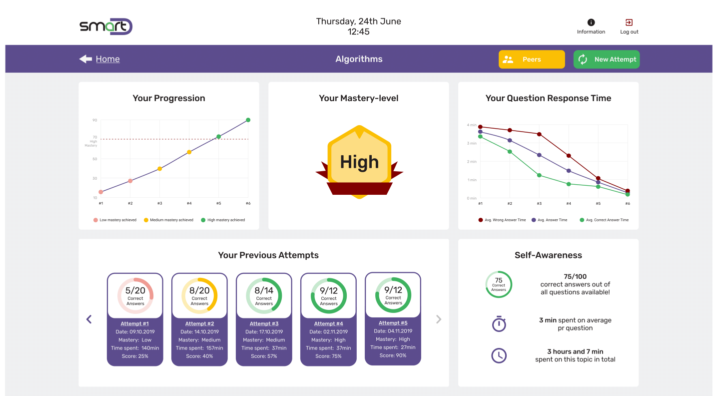

Fig. 2 smartU – Selected activity

When a learner selects to take a self-assessment test, the questions are delivered to her one-by-one according to the underlying adaptation mechanism: the next most appropriate question to deliver is selected according to the correctness of the student's response to the previous question and the discrimination ability of the remaining questions, so that the student's mastery level can be estimated by administering the minimum number of questions. The questions have up-to four possible answers, but only one is the correct. Every time the student submits an answer to a question, her mastery class is revised accordingly, and the next question is delivered to her. The selection of the next question is based on entropy, a maximum information gain strategy from Information Theory; the goal is to select the question that has the greatest expected reduction in entropy, i.e., that better fits the learner's mastery

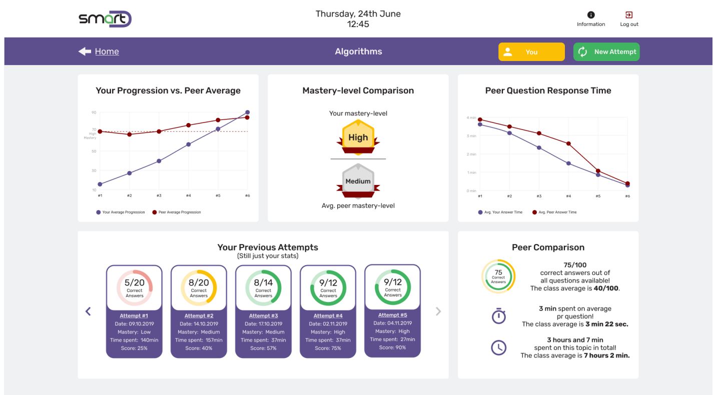

Fig. 3 smartU – Peer comparison

class, based on the answers she provided on the previous questions. For adapting the self-assessment, the Measurement Decision Theory (MDT) [36] was employed.

Example of a question is illustrated in Fig. 4. Here, the question has been answered wrongly by the learner, and the correct answer is shown upon request.

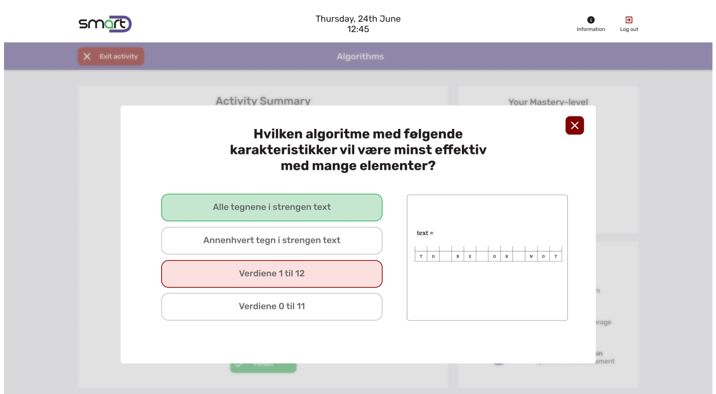

Fig. 4 smartU – Question with correct answer

The service also delivers task-related visual analytics per question, based on the logged interactions data, as explained in [29]. The analytics are shown in Fig. 5. The task-related information provided to the learner was determined so as this knowledge to activate learner's monitoring, reflection and judgment (i.e., metacognition) about the questions, with an ultimate goal to help the learner to meet the requirements of each question, i.e., the actual difficulty, the actual effort needed to deal with each question, and the time required to allocate on each question. Using properly this information is expected to support the learner to efficiently regulate herself and her motivation, i.e., to improve her effort allocation, time-management and help-seeking skills, and metacognitive inference-making based on her own learning goals [20]. Previous research has shown that visualization of aggregated temporal indexes increases the teachers' awareness on students' progress and helps them revise their considerations about the actual requirements of the assessment tasks [27].

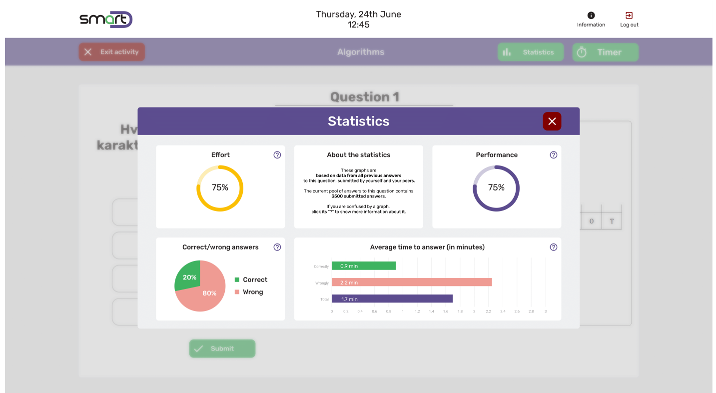

Fig. 5 smartU – Question with correct answer

4 Methods

4.1 Participants and study design

Thirty-five participants volunteered initially to take the self-assessment tests and were scheduled, but due to the outbreak of the Covid-19, finally 27 could be conducted prior to the lock-down of the University (55% males, 41% females and 4% non-binary gender, aged 19-27 years-old [M=22.4, SD=2.0]). The sample consisted of students spread across different years of study (M=3.4 years, SD=1.5, min=1, max=5), with most students being enrolled in a programming related study-program (77%), and the rest (23%) being enrolled in other programs (e.g., chemistry or biology). For the needs of the study, the educational material from the Introduction to Programming course was utilized. The item-bank consisted of 120 multiple-choice questions.

The study followed an experimental strategy, using a static group comparison design [12]. The participants were split randomly into two groups, with 12 being in the group with access to task-related visual analytics (the experimental group) and 15 not having access to the task-related visual analytics (the control group). None of the participants had any previous experience with the SmartU service. A maximum of two participants were taking a self-assessment test at the same time, with both participants belonging to the same group, to ensure a controlled setting. Each session lasted for a maximum of 45 minutes, and due to the Covid-19 outbreak, all devices used in each session were cleansed thoroughly with antibacterial wipes.

The study was conducted in two phases. The first phase took place in lab conditions, in three steps and was completed on the day that the lock-down was announced. The second phase was conducted during the lock-down and included one step.

Step 1: Briefing

All participants signed an informed consent form prior to their participation, explaining them the procedure and giving the right to researchers to use the data collected for research purposes. Next, the participants were briefly introduced to SmartU and its adaptivity through a printout providing step-by-step explanations. Participants in the experimental group were additionally introduced to the task-related visual analytics and how to use it. The printouts were available throughout the procedure.

Step 2: Test-taking procedure

After getting familiar with the service, the participants were prompted to imagine using the service from home to practice their knowledge and programming skills on the introductory course, aiming to create the feeling of studying fully online from home. It was also clarified that their achieved self-assessment results would have no participation to the their final course grade (i.e., no rewards as external motivation).

The participants were asked to complete two or three self-assessment activities, depending on how much time was spent on each attempt. Each attempt contained approximately 10 to 12 questions, with two to four possible answers, but only one answer was correct. The questions were delivered to the students according to the adaptation mechanism described in the previous section. Only one answer could be submitted to a question, and the students could not return to a previous question or change an answer once submitted, due to the adaptation mechanism.

Step 3: Debriefing

At the last step of the first phase, the participants had to fill-in a post-assessment questionnaire that measures their motivation to use the service and their opinions about its usability and usefulness. They responded to the questionnaire individually, on the computer they used for the self-assessment, right after completing the tests.

Upon finishing the questionnaire, all participants received a gift card as a reward of voluntarily contributing their time and data to the research.

Step 4: Interviews

The final step – i.e., the second phase – included a follow-up interview and was conducted three to four days after the first phase. Ten participants were randomly selected from the two groups (5 participants from each group). Due to the Covid-19 outbreak and the university closing the campus, the interviews had to be conducted online, using the free online conference room software Whereby, and using as backup Discord, another free online conference room software. Open Broadcaster Software, a free and open source software for recording and live streaming, was used to record the audio from the interviews. Although web-cameras were used during the interviews, the video was not recorded. All interviews were conducted in Norwegian, as all interviewees were more comfortable with Norwegian than English.

4.2 Data Collection

4.2.1 Questionnaires

The questionnaire captures the participants' opinions on the usability of the service and the different elements (e.g., visualizations), as well as their attitudes and motivation to use the service. For this purpose, the items utilized were questions on a 5-point Likert-like Scale (1="Strongly Disagree", 5="Strongly Agree") [25, p. 113].

Category	**Acronym**	**Constructs**
System Usability Score	SUS	Usability score
Overall evaluation of the usability of SmartU	OEUS	General Usability
Attitude towards Graphs and Visualizations	AGV	Usage Motivation Usefulness Positive and Negative emotions Intent for further use
Usability of Graphs and Visualizations	UGV	Usability of features Intuitiveness Usefulness
Graphs and visualizations in an assessment	GVA	Usability of features Usage Usefulness Intuitiveness Usefulness Motivation Intent for further use Positive and Negative emotions

Table 1 Overall categories in the questionnaire and their constructs

As shown in Table 1, the first two categories, i.e., *System Usability Score (SUS)* and *Overall evaluation of the usability of SmartU (OEUS)*, contained questions about the usability of the service and were related to the design and creation strategy, establishing whether or not the user-interface of SmartU met the requirements of usability posed by the participants of the study. SUS is a standardized method of measuring usability [5]; OEUS weas created to assess whether or not the system met the definition of usability as described in ISO 9241-210:2019: *extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*.

The two next categories in Table 1, i.e., *Attitude towards Graphs and Visualizations (AGV)* and *Usability of Graphs and Visualizations (UGV)*, included questions about the participant's attitudes towards the graphs and visualization dashboards implemented in the service, and the perceived usability of these elements. The questions and constructs in these categories were adapted from previous relevant studies and altered to fit the context and the service used in this study [34, 11, 18].

The last category in Table 1, *Graphs and Visualizations in an Assessment (GVA)*, contained questions about the task-related visual analytics available during the self-assessment. This category covered the participants' perceived usage and usefulness, and their attitude towards such visualized statistics. The questions and constructs used in this category was also adapted from previous relevant studies [34, 11, 18].

4.2.2 Interviews

During the second phase of the study, to gain additional understanding of the participants' motivation to use (or not) the service, semi-structured interviews [39] were conducted, according to a guide created beforehand. This guide contained a set of situations requiring extra attention by the interviewer and a set of prepared questions, created based on interesting overall trends found in the answers from the questionnaire. However these questions was just used as conversation-starters to be able to let the interviewee tell about her experiences, feelings and thoughts and thus allowing the interviewer to probe with appropriate follow-up questions [24, p. 188]. Examples of the prepared questions can be found in the bullet-point list below.

- Do you think that the service would improve your motivation of studying? If you do, why would it do that?
- Would you use the system again? If so, what encourages you to do so?
- Could you mention some features which were easy or hard to understand?
- Did you feel like the statistics helped you understand the scope of the question? How did the statistics help you?

4.3 Data analysis

4.3.1 Quantitative analysis

To explore the differences in participants' attitudes and motivation to use the service, descriptive statistics, correlation analysis and independent samples t-test were applied between the control and the experimental groups, using IBM's SPSS.

4.3.2 Qualitative analysis

To analyze the interviews, they first had to be transcribed. While transcribing the interviews the broad and recurring themes were noted for later use in the analysis.

For the analysis itself Nvivo, a qualitative data analysis software, was used to code the interviews into categories. The categories used for the analysis were initially based of an deductive approach, by using theories based on information from the literature review [24, p. 269]. However, as the coding was iterative, the notes taken when transcribing and a repeatedly reading through the interviews, quickly formed new categories. Thus, the analysis also followed an inductive approach [24, p. 269].

Clusters of similar or connected content were split into two, more granular, categories, having their original category as their parent category. Similarly, smaller categories were merged with other categories. Thus, a “tree-like” structure was established, providing a detailed separation of the content of the interviews.

In regard to working with qualitative data, it was important that the the actual interviewees’ quotes was intact and not altered. Thus, when an answer to a question or a quote was ambiguous, the correct context for the quote was added, followed by "red. anm.", an abbreviation for *redaksjonell anmerkning*, meaning *editorial/writers remarks* in English. These remarks were clearly marked with parentheses.

5 Results

5.1 Questionnaire Mean Variables

Table 2 demonstrates the descriptive statistics for the questionnaire constructs. Overall, there is an above average (i.e., positive) attitude towards SmartU’s use of visualizations and graphs for both AGV and UGV indexes. However there is a lower minimum score towards OEUS, indicating that the overall usability is fluctuating from user to user. There are also broad opinions and attitudes towards the usefulness of GVA, considering the standard deviation.

	N	Min	Max	Mean	Std Dev
Overall evaluation of the usability of SmartU	27	2.67	4.89	4.26	.49
Attitude towards Graphs and Visualizations	27	3.29	5.00	4.39	.43
Usability of Graphs and Visualizations	27	3.20	5.00	4.33	.51
Graphs and Visualizations in an assessment	12	2.53	4.47	3.50	.55

Table 2 Questionnaire Mean Variables of categories

Furthermore, when looking into the descriptives of each group separately in Ttable 3, there are some key differences that should be noted. Compared to the experimental group, the control group’s OEUS Mean and Minimum value suggests that the overall usability is better than when task-related visual analytics are introduced. This indicates that those metacognitive statistics might introduce a new level of difficulty to the service, increasing the required effort to understand the visualizations.

5.2 Correlations

Table 4 gives an overview of Pearson Correlations between the questionnaire categories. The correlation is significant at the 0.05 and 0.01 level (2-tailed). Variables compared in the table: OEUS, AGV, UGV and GVA.

	Group 1 (control), N=15				Group 2 (experimental), N=12			
	Min	Max	Mean	Std dev	Min	Max	Mean	Std Dev
Overall evaluation of the usability of SmartU	3.56	4.89	4.33	.38	2.67	4.89	4.17	.60
Attitude towards Graphs and Visualizations	3.29	5.00	4.45	.46	3.59	4.88	4.32	.41
Usability of Graphs and Visualizations	3.20	5.00	4.37	.55	3.60	5.00	4.27	.46
Graphs and Visualizations in an assessment					2.53	4.47	3.50	.55

Table 3 Questionnaire Mean Variables between groups

	Mean (std dev)		OEUS	AGV	UGV	GVA
Overall evaluation of the usability of SmartU (OEUS) (N=27)	4.259 (0.485)	Pearson Corr. Sig. (2-tailed)	1			
Attitude towards Graphs and Visualizations (AGV) (N=27)	4.392 (0.433)	Pearson Corr. Sig. (2-tailed)	**.746**** .000	1		
Usability of Graphs and Visualizations (UGV) (N=27)	4.326 (0.505)	Pearson Corr. Sig. (2-tailed)	**.732**** .000	**.758**** .000	1	
Graphs and Visualizations in an assessment (GVA) (N=12)	3.499 (0.546)	Pearson Corr. Sig. (2-tailed)	-.133 .727	-.161 .617	.239 .455	1

*. *Correlation is significant at the 0.05 level (2-tailed).*
**. *Correlation is significant at the 0.01 level (2-tailed).*

Table 4 Correlation table for questionnaire

Correlation analysis revealed some strong positive relations between AGV and OEUS (r = .746, n = 27, p = .000), UGV and OEUS (r = .732, n = 27, p = .000) and UGV and AGV (r = .758, n = 27, p = .000). These correlations indicate that there is a continuous level of motivation and attitude towards visualizations, graphs and usefulness throughout the whole SmartU service.

5.3 Independent Samples T-test

As seen in Table 5, there are no statistically significant differences on the perceptions regarding the usability of and motivation to use SmartU between the two groups.

Equal Variances Assumed	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Overall evaluation of the usability of SmartU	-.882	25	.386	-.117	.189
Attitude towards Graphs and Visualizations	-.785	25	.440	-.133	.169
Usability of Graphs and Visualizations	-.538	25	.595	-.107	.198

$*p < 0.05$

Table 5 Independent Samples T-test for the differences of perceived use and motivation to use with and without statistics.

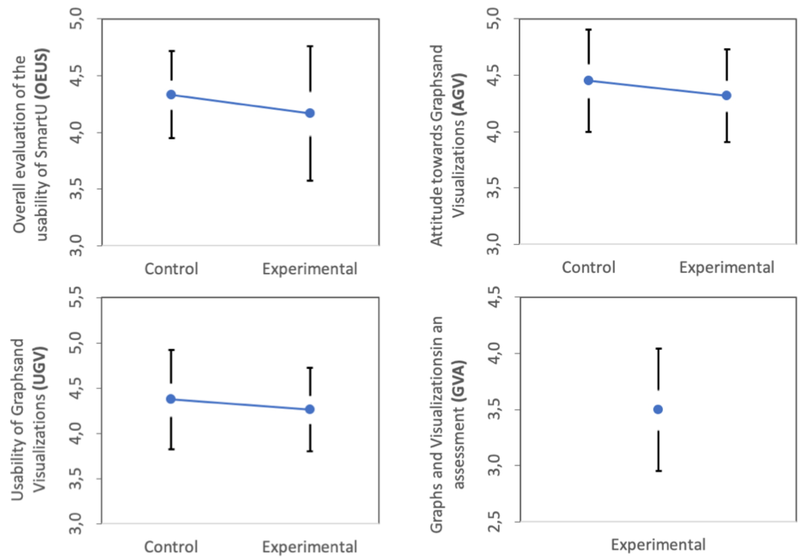

Fig. 6 Graphical representation of Independent Samples T-tests between the control and experimental group. The dots represent the mean value while the black bars represent the standard deviation. Statistical significance is not marked due to lack of significant results.

5.4 System Usability Score

Table 6 shows the average SUS-score for each of the groups in the study, and their average score for all questions in the SUS-schema. The best score between the two groups for each question is highlighted in bold. The total average SUS-score for the control group was 83, the total average SUS-score for the experimental group was 84,375, while the total average SUS-Score for both groups combined was 83.61. Furthermore, from all the SUS-scores, the *lowest score* was 67.5, the *highest score* was 95, while the *median score* was 85.

5.5 Results from the interviews

The results presented in this section are extracted from the 10 conducted interviews. The original statements by the participants of the study, expressing their opinions and thoughts, were in Norwegian, but here we will present an English translation.

The tree-like hierarchy created as a result of categorizing the statements of the participants during the analysis, contained seven categories (hereby referred to as nodes) at the top layer. Four of the seven nodes (i.e., Features, Learning, Motivation

Question	**Control group** (N=15)	**Experimental group** (N=12)
Q1 - Frequent Use	**4,27**	3,92
Q2 - Unnecessarily complex	2,27	**2,25**
Q3 - Easy to use	4,33	**4,42**
Q4 - Would need support	1,33	**1,25**
Q5 - Well integrated functions	4,27	**4,42**
Q6 - Too much inconsistency	1,80	**1,42**
Q7 - Learn it quickly	4,53	**4,58**
Q8 - Slow or complicated	**1,20**	1,50
Q9 - Felt confident	4,00	**4,25**
Q10 - Required training	1,60	**1,42**
Average SUS-score for group	83	84,375

Table 6 Average SUS-score of both groups for the variables in the SUS-schema

and Remarks) only contained other nodes and no direct content in the form of statements. The three remaining nodes (i.e., Clarifications, Colors and Gamification) contained only statements and no other nodes.

5.5.1 Participants insights towards motivation of use

The first noticeable group emerged within the parent category containing nodes with statements regarding motivation. The use of the word motivation in these nodes, revolves around the participants feeling more or less motivated to use the service for educational purposes and whether or not there was an increase in the motivation to study in general. This could indicate participants stating either being motivated or demotivated, or simply stating that there was no change in their motivation. There was also a separate node containing all references to the statistics, overlapping with the other motivation categories mentioned. The nodes and their number of references are shown in table 7.

Name	**Files referenced**	**Statements in total**
Negative Motivation	7	14
Positive Motivation	10	47
No Change	5	10
Statistics	8	21

Table 7 Nodes making up the Motivation-group, with number of files referenced and number of statements.

As can be seen in table 7, all 10 interviewees had statements regarding their motivation being positively impacted while using the SmartU service.

Most of the answers in both the Positive Motivation and Negative motivation revolved around the topic of comparing results, either comparing against themselves or against their peers. One statement in particular, in the Positive Motivation, highlighted this:

> "... when you receive information about how others are doing, it becomes very natural that there will be some kind of competition present. Which allows you to quickly set goals like: "I will be this good compared to everyone else or in comparison to yourself". So you can compete both against yourself and against others, and that helps my motivation, instead of... yeah... just doing everything on your own."

The element of competition and comparison to others, being the topic most frequently talked about in the interviews, was further reflected on with a participant stating:

> "What motivates me is, again, that i can see myself and my level of knowledge, and especially that you can see the rest of the users of the system and their level."

The comparison with ones own previous performance was also emphasized with one participant stating that:

> "... if I were to use this it would give me more motivation if I see a steady growth (in the graph illustrating performance, red. anm.), or if I see a decline I might think «ah man, I haven't focused for long enough, or not focused enough» if it goes on for a longer period. So it becomes easier to be self-critical and analytical when you look at your own work..."

And another participant stating:

> "It's more or less unchanged as I feel I have quite high motivation myself in different courses, so I don't necessarily need to see the statistics of others to compete with them. I am more interested to compete against my own results."

Despite all participants having some positive feedback towards the system, there was also statements regarding their motivation being negatively influenced. The possibility of comparing ones results to ones previous performances and ones peers was a key example of this, proving to be a double-edge sword as it also received feedback of negatively influencing the motivation. One of the participants summarized this in this statement:

> "I might have taken it a little personally if I had been placed very low and I saw that others were placed very high and that the curriculum was easy for them. So if this was voluntary and had no impact on the grade, I might have just studied by myself and thought it probably was sufficient."

Other participants shared this opinion, as another participant added to this by stating:

> "If you are far behind towards the end of the semester, you'd think «why bother, there is not enough time». I would probably be demotivated if I was that far behind."

5.5.2 Participants insights towards usefulness and intention to use

When asking the participants if SmartU was an intersting system, and what made it so, some of replies from the participants revolved around the aspects of digitized learning. Two of the participants found the system interesting because it had game-like features. The first of the participants stated that:

> "I think it was interesting because it sort of resembles a game, which is a very different approach than the system i have been using in relation to ITGK."

The second participant was a little more specific pointing at the mastery-level feature, stating that:

> "... the little «ranking-system» in the middle of the page, where it says «High» or «Medium» or creating a separation of sorts, and the fact that you get a little medal, in my eyes, creates a cooler... yeah ..."

When asking the participants whether or not they would use the service again, and if so why they would use it, the response was mostly focused around the way the SmartU service could be used as a tool for enhanced learning. One of the participants stated that:

> I am a big fan of using things like this instead of for instance just sitting in reading a book. So this makes it a little fun. After all, it is a slightly different way of learning, to just be served questions and answers, rather than just sitting and reading. It's a slightly more effective way to do exam sets, perhaps."

A second participant took a broader and more summarizing approach in its answer, stating that:

> "I would definitely use the system, if possible, in most subjects I could use it for really. It gave a, umh... Very concrete right / wrong and progression and... You could look at aspects you couldn't aggregate yourself, like response time. You can interpret from the mastery level, not only if you were right, but also if you were right on difficult questions. Or if you just have a superficial understanding, in that you are right on a lot of questions, but there have been easy questions so you are fooling yourself into thinking you know more than you really do. And it had a pretty nice user interface with visualizations that was easy to understand by just looking at them for two seconds, and then just moving on."

6 Discussion and Conclusions

6.1 Factors affecting attitude and motivation

Sustaining learners' motivation in online learning settings is a challenging task. Previous work has associated motivation with learners' need to experience autonomy [28, 33], which in turn, has been related to practising SRL skills. It has also been proposed in existing literature that self-assessment tests can be an efficient way for learners to strengthen their self-reflection, and improve their self-regulation capabilities [23, 29]. Even more when the self-assessment tests are adaptive, the tailored content can further promote learners' motivation to remain engaged in their learning [21]. However, in times of global crisis, such as the one we are experiencing due to the Covid-19 outbreak, and given the closure of the schools and universities and the subsequent fully online teaching and learning, the need to support learners' motivation in their learning is more urgent than ever before [16]. The RQ that guided this study was: *"How learners' motivation is affecting their dispositions to use an online adaptive self-assessment service, and continue learning during a crisis?"*.

The results from the SUS-schema shown that the experimental group had a higher SUS-score than the control group, and that 8 out of the 10 questions that are related to the usability of the service had a higher score in the experimental group. This finding indicates that the experimental group found adaptive self-assessment service more usable than the control group. However, due to the low sample-size, the results are easily affected by personality types and outliers in the responses [19].

Furthermore, from the interviews it became explicit that the participants had more statements regarding potential positive influences in their motivation than negative (Table 7). Most of these statements concerned participants comparing results, either comparing against their own previous results or against their peers. The peer comparison feature particularly was deemed the most useful feature in the system, as mentioned in section 5.5.1. This feature was also particularly interesting regarding the attitude and motivation of the participants and can be connected to the correlation between the AGV and UGV-variables. From the interviews, statements claiming that by receiving information of how the average of other participants were doing and ones own previous performances, a competition emerges naturally. The creation of goals of being better than the average participants or beating ones own high score is a natural reaction to this. Thus, the results from the interviews could indicate that the graphs and visualizations displaying ones score and the peer comparison feature were contributing to the positive influence in motivation.

However, as explained in [17], peer comparison should be used cautiously, as different types of learners perceive it differently. Their research found a connection between academic performance and the perception of social comparisons. Based on the interviews conducted in this study, one can see similarities to the findings of [17] as the participants states that their motivation would change based on their positioning in the peer comparison. Some stated that being ahead of the average curve in the peer comparison would influence their motivation positively, as it was perceived as a confidence boost. Others however, claimed their motivation would be bigger if position slightly behind the average curve, as it was a way of improving your own skills more than the others. What recurred in most interviews was that most participants would feel demotivated if positioned too far behind the average curve. This was especially the case in specific contexts, like when using the system for practicing for a course in a school context and the exam date is approaching.

6.2 Research limitations: Covid-19 and its implications

While conducting the study, the outbreak of Covid-19 reached Norway and ultimately forced the university to shut down the campus. This had an impact on various aspects of this study and how it was conducted.

6.2.1 Final sample-size of participants

The most notable limitation was the fact that students and staff were not allowed to stay on campus. This happened during the second day of the study and implied that the remainder of the tests had to be canceled. Furthermore, the days before the study started, some of the participants canceled their scheduled session due to fear of the high risk of infection Covid-19 has. In total Covid-19 had a big impact on the sample-size of the participants as the total number of participants went from 35, as originally planned, to 27. The low sample-size means that the statistical power behind the statistical analysis is low. The results from the Pearson correlations and the Independent Samples T-tests, should therefore be carefully considered against the sample-size as there is a possibility that the correlations do not reflect a true effect. There is also a possibility that other true effects within the results are not discovered.

6.2.2 Physical attendance needed for testing

As the study was conducted in a controlled environment, requiring physical attendance for taking the self-assessment tests, another implication of Covid-19 was the strict demands for disinfection and general hygiene. As stated in section 4.1, all laptops and other equipment used throughout the study were thoroughly cleansed with antibacterial wipes between every conducted test. Furthermore, the table and chairs used for were also cleaned between tests. This was a time consuming routine and caused delays in the study as the day progressed, due to a tight schedule, not made with disinfection of equipment in mind.

Mostly, as the number of infected population increased rapidly during the study, the uncertainty about what was going to happen was also a factor that had side-effects on the study. Although it was not possible to be measured, the mood and atmosphere during the study was a bit pressed, most likely due to the Covid-19. From the observation notes, all participants seemed concentrated and were quiet, however some expressed concern for being infected at campus during small talk after completing the self-assessment tests.

6.2.3 Interviews conducted over video

As explained in section 4.2.2, semi-structured interviews were conducted with selected participants. However, due to the Covid-19 pandemic, and the consequent closure of the university, it was not possible to conduct these interviews in a face-to-face setting. Thus, Whereby and Discord was used to conduct the interviews. On the one hand, this allowed for the easier recording of the conversation. However, some of the aspects of a face-to-face interview, such as the ability to observe the body-language and the fine-tuned facial mimics, were not possible to capture. Furthermore, as all participants had to stay in their homes while being interviewed due to the campus being closed, the setting and ambiance of the interview were also changed. In a sense, the mindset and concentration of the participants was most likely also influenced by the situation and the implications it had for the participants on a personal level.

6.3 Conclusions

As extensively explained, the need to support and sustain learners' motivation in their learning during global crises is a priority. Furthermore, investing on developing intuitive and efficient adaptive self-assessment services to facilitate this objective, has been considered a meaningful and promising step. Although the sample-size of the present study could not satisfy statistical power for the quantitative analysis, the qualitative analysis of the interviews revealed the potential that the SmartU service has to promote autonomous and self-regulated learning in fully online settings. From the results, it becomes apparent that dedicated services such as SmartU, have the capacity to motivate learners to remain engaged in regularly practising their knowledge through self-assessment quizzes. The core characteristic of such services is that they tailor the content to the personalized learning mastery of each individual. Furthermore, the intuitive dashboard offers a variety of learning analytics and gamification elements that turn the learning experience into a playful and fruitful learning gain.

References

1. Adesope, O.O., Trevisan, D.A., Sundararajan, N.: Rethinking the use of tests: A meta-analysis of practice testing. Rev. Educ. Res. **87**(3), 659—701 (2017)
2. Alexander, B., Ashford-Rowe, K., Barajas-Murph, N., Dobbin, G., Knott, J., McCormack, M., Pomerantz, J., Seilhamer, R., Weber, N.: Educause horizon report 2019 higher education edition. Tech. rep., EDU19 (2019)
3. Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., Verbert, K.: Open learner models and learning analytics dashboards: a systematic review. In: Proceedings of the 8th international conference on learning analytics and knowledge, pp. 41–50 (2018)
4. Broadbent, J.: Comparing online and blended learner's self-regulated learning strategies and academic performance. The Internet and Higher Education **33**, 24–32 (2017)
5. Brooke, J., et al.: Sus-a quick and dirty usability scale. Usability evaluation in industry **189**(194), 4–7 (1996)
6. Candy, P.C.: Self-Direction for Lifelong Learning. A Comprehensive Guide to Theory and Practice. San Francisco: Jossey-Bass (1991)
7. Cantor, A.D., Eslick, A.N., Marsh, E.J., Bjork, R.A., Bjork, E.L.: Multiple-choice tests stabilize access to marginal knowledge. Memory & Cognition **43**(2), 193—205 (2015)

8. Carpenter, S.K.: Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition **35**(6), 1563 (2009)
9. Carpenter, S.K.: Testing enhances the transfer of learning. Current directions in psychological science **21**(5), 279—283 (2012)
10. Carver, C., Scheier, M.: Self-regulation of affect and action. In: Handbook of self-regulation: Research, theory, and applications, pp. 3–21. Guilford Press New York, NY (2011)
11. Chiu, C.M., Wang, E.T.: Understanding web-based learning continuance intention: The role of subjective task value. Information & Management **45**(3), 194–201 (2008)
12. Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., Schauble, L.: Design Experiments in Educational Research. Educational Researcher **32**(1), 9–13 (2003). DOI 10.3102/0013189X032001009
13. Elliot, A.J., Murayama, K.: On the measurement of achievement goals: Critique, illustration, and application. J. Educ. Psychol. **100**(3), 613—628 (2008)
14. Holec, H.: Autonomy and foreign language learning. Oxford: Pergamon. (First published 1979, Strasbourg: Council of Europe) (1981)
15. Huang, J.P., Benson, P.: Autonomy, agency and identity in foreign and second language education. Chinese Journal of Applied Linguistics **36**(1), 7–28 (2013)
16. Huang, R., Liu, D., Tlili, A., Yang, J., Wang, H., et al.: Handbook on facilitating flexible learning during educational disruption: The chinese experience in maintaining undisrupted learning in covid-19 outbreak. Beijing: Smart Learning Institute of Beijing Normal University (2020)
17. Jivet, I., Scheffel, M., Specht, M., Drachsler, H.: License to evaluate: preparing learning analytics dashboards for educational practice. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge, pp. 31–40. ACM (2018)
18. Lee, M.C.: Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model. Computers & Education **54**(2), 506–516 (2010)
19. Lewis, J.R.: The system usability scale: past, present, and future. International Journal of Human–Computer Interaction **34**(7), 577–590 (2018)
20. Lonn, S., Aguilar, S.J., Teasley, S.D.: Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. Computers in Human Behavior **47**, 90–97 (2015)
21. Martin, A.J., Ginns, P., Papworth, B.: Motivation and engagement: Same or different? does it matter? Learning and Individual Differences **55**, 150 — 162 (2017)
22. McCardle, L., Hadwin, A.F.: Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. Metacognition and Learning **10**(1), 43–75 (2015)
23. Nicol, D.J., Macfarlane-Dick, D.: Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in higher education **31**(2), 199–218 (2006)
24. Oates, B.J.: Researching information systems and computing. Sage (2005)
25. Oppenheim, A.N.: Questionnaire design, interviewing and attitude measurement. Bloomsbury Publishing (2000)
26. Özyurt, Ö., Özyurt, H.: Learning style based individualized adaptive e-learning environments: Content analysis of the articles published from 2005 to 2014. Computers in Human Behavior **52**, 349–358 (2015)
27. Papamitsiou, Z., Economides, A.A.: Temporal learning analytics visualizations for increasing awareness during assessment. International Journal of Educational Technology in Higher Education **12**(3), 129–147 (2015)
28. Papamitsiou, Z., Economides, A.A.: Exploring autonomous learning capacity from a self-regulated learning perspective using learning analytics. British Journal of Educational Technology **50**(6), 3138–3155 (2019)
29. Papamitsiou, Z., Economides, A.A., Giannakos, M.N.: Fostering learners' performance with on-demand metacognitive feedback. In: European Conference on Technology Enhanced Learning, pp. 423–435. Springer (2019)

30. Papamitsiou, Z.K., Economides, A.A.: Towards the alignment of computer-based assessment outcome with learning goals: the laers architecture. In: 2013 IEEE Conference on e-Learning, e-Management and e-Services, pp. 13–17. IEEE (2013)
31. Pappano, L.: The year of the mooc. The New York Times **2**(12), 2012 (2012)
32. Pintrich, P.R.: The role of goal orientation in self-regulated learning. In: Handbook of self-regulation, pp. 451–502. Elsevier (2000)
33. Reeve, J., Ryan, R., Deci, E.L., Jang, H.: Understanding and promoting autonomous self-regulation: A self-determination theory perspective. Motivation and self-regulated learning: Theory, research, and applications pp. 223–244 (2008)
34. Roca, J.C., Chiu, C.M., Martínez, F.J.: Understanding e-learning continuance intention: An extension of the technology acceptance model. International Journal of human-computer studies **64**(8), 683–696 (2006)
35. Ruben, B.D., De Lisi, R., Gigliotti, R.A.: A guide for leaders in higher education: Core concepts, competencies, and tools. Stylus Publishing, LLC (2016)
36. Rudner, L.: The classification accuracy of measurement decision theory. In: annual meeting of the National Council on Measurement in Education, Chicago, vol. 2325. Citeseer (2003)
37. Sharma, K., Papamitsiou, Z., Giannakos, M.: Building pipelines for educational data using ai and multimodal analytics: A "grey-box" approach. British Journal of Educational Technology **50**(6), 3004–3031 (2019)
38. Viner, R.M., Russell, S.J., Croker, H., Packer, J., Ward, J., Stansfield, C., Mytton, O., Bonell, C., Booy, R.: School closure and management practices during coronavirus outbreaks including covid-19: a rapid systematic review. The Lancet Child & Adolescent Health (2020)
39. Wilson, C.: Chapter 2 - semi-structured interviews. In: C. Wilson (ed.) Interview Techniques for UX Practitioners, pp. 23 – 41. Morgan Kaufmann, Boston (2014). DOI https://doi.org/10.1016/B978-0-12-410393-1.00002-8. URL http://www.sciencedirect.com/science/article/pii/B9780124103931000028
40. Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G.J., Paas, F.: Supporting self-regulated learning in online learning environments and moocs: A systematic review. International Journal of Human–Computer Interaction **35**(4-5), 356–373 (2019)
41. Zhou, L., Li, C.: Can student self-directed learning improve their academic performance? experimental evidence from the instruction of protocol-guided learning in china's elementary and middle schools. Experimental Evidence from the Instruction of Protocol-Guided Learning in China's Elementary and Middle Schools (2020)