Eindhoven University of Technology

MASTER

Design and Evaluation of a Student-Facing Learning Dashboard Using Self-Regulated Learning Theory

Peters, B.A.

Award date:
2023

Eindhoven University of Technology
Department of Industrial Engineering & Innovation Sciences

Design and Evaluation of a Student-Facing Learning Dashboard Using Self-Regulated Learning Theory

B.A. Peters

in partial fulfillment of the requirements for the degree of
Master of Science in Human-Technology Interaction

Supervisors:
Dr. U. Matzat
Dr.ir. M.A. Conijn

Eindhoven, March 2023

Acknowledgments

This Master Thesis report was the final step in getting my Master's degree in Human-Technology Interaction at TU/e. I didn't do it all alone, and I would therefore like to take a moment to thank everyone that helped me along the way.

First of all, I would like to give a special thanks to my supervisors Uwe Matzat and Rianne Conijn. I truly valued your continuous guidance throughout my master thesis project, and I certainly feel that your expertise and constructive criticism has made me grow during the last few months. I acknowledge (and fully understand) that my email-novels must have become challenging to respond to over time. Nonetheless, you have continued to provide me with insightful feedback again and again.

Secondly, I would like to thank the IMS-team at TU/e, in particular Jeroen Bovend'eerdt, Tom Vogten and Isa Vogten. I would not have been able to develop the learning dashboards as they are right now without your help, and I see having the chance of working with all of you as a valuable experience.

Finally, I would like to thank all of my friends and family who have supported me until the moment I submitted this master thesis. Even with some small bumps in the road, you have cheered me on until I reached the finish line. Thank you for that.

Bram Peters
March, 2023

Abstract

This study focuses on how the design and evaluation of student-facing learning dashboards can be grounded in self-regulated learning theory, and to what extent the (in)consideration of the theory's phases yields different cognitive and affective learning outcomes. Despite learning dashboards being a promising development within the learning analytics field, their overall use and long-term impact on students' learning outcomes remain low (Gašević, Dawson & Siemens, 2015). The current approach to designing, developing, and evaluating learning dashboards seems to rely heavily on the data that is made available by the learning management systems onto which they are built, rather than a solid theoretical foundation that is grounded in learning sciences (Matcha, Uzir, Gasevic & Pardo, 2020). Despite many studies on learning dashboards suggest their designs to support self-regulated learning, many dashboard do not truly cover all phases of this learning process. Following a theory-driven, user-centered design approach, a qualitative interview study was performed – guided by insights from the literature - to understand the extent to which students' personal experiences and needs for specific features in a learning dashboard can be realized within a design that is guided by self-regulated learning theory. Results from the interview study were used as guidance in constructing design features that support each phase of self-regulated learning in a way that is contextually appropriate. In an experimental study, differences in dashboard evaluations and cognitive and affective learning outcomes were examined between a group of students that had access to a learning dashboard that supported all three phases of self-regulated learning, and a group of students that had access to a learning dashboard that only supported the performance phase. The experimental findings highlight that the differences between the two learning dashboards was moderated by several factors. More support for self-regulated learning increased perceived usefulness and motivation, and this effect was different depending on the type of course for which the dashboard was enrolled. The findings highlight the need for future research to explore the relationship between different factors that influence learning outcomes, including dashboard design, perceived usefulness and support for self-regulated learning, perceived autonomy, and metacognitive awareness.

Table of Contents

1. Introduction

Over the years, online learning has become a viable solution in education to reduce costs and increase flexibility in offering courses (Allen & Seaman, 2014). The large-scale deployment of online learning platforms has enabled educational institutions to gather vast amounts of data generated from students' use of these systems. This data contains information on how students navigate the university's learning management systems and how they progress throughout their academic studies (Verbert, Govaerts, Duval, Santos & Klerkx, 2014). The advancements of online learning platforms open up a fruitful ground for new understandings on how to analyse and make use of this data. The field of learning analytics aims to exploit these opportunities by measuring, collecting, and analysing online student-data, with the aim of uncovering behavioural patterns that could help understand how to optimize students' learning processes more effectively.

At the forefront of these developments, many educational institutions have started to deploy a variety of learning analytics systems. Whereas during the early days of learning analytics the focus was on providing teachers with tools to offer timely interventions to students if needed (i.e., by identifying and monitoring students that are falling behind), recently the focus has shifted towards reporting this data directly to students themselves (Baker, 2016). One tool that illustrates this trend are so-called student-facing learning analytics dashboards, or online learning dashboards. As described by Schwendimann et al. (2017) online learning dashboards are "single displays that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualizations". By presenting students with visual indicators of their study behaviour and progress, the fundamental aim of learning dashboards is to trigger a process of self-reflection and strategy evaluation (Bodily & Verbert, 2017).

Despite learning dashboards being a promising development within the learning analytics field, their overall use and long-term impact on students' learning outcomes remain low (Gašević et al., 2015). A series of systematic reviews of the literature on learning dashboards has exposed a notable pattern of shortcomings in the design and evaluation of dashboards that might explain their low use and long-term impact (Jivet, Scheffel, Specht & Drachsler, 2018; Matcha, Uzir et al., 2020). The current approach to designing, developing, and evaluating learning dashboards seems to rely heavily on the data that is made available by the learning management systems onto which they are built, rather than a solid theoretical foundation that is grounded in learning sciences (Matcha, Uzir et al., 2020). This is seen as a critical shortcoming, because truly understanding how best to facilitate online learning creates opportunities to construct and provide more valuable feedback to students on how to steer their study strategies. This might ultimately increase a dashboard's long-term impact on learning outcomes (Matcha, Uzir et al., 2020). An inadequate implementation of learning science frameworks during the design and evaluation of learning dashboards could possibly result in tools with little value that will quickly be dismissed (Jivet, Scheffel, Drachsler & Specht, 2017).

Online learning dashboards that state their designs to be grounded in learning sciences commonly aim to do so by implementing a theoretical framework called self-regulated learning theory (Jivet et al., 2018). Zimmerman and Schunk (2001) present this framework as a cyclical process whereby learning goes through several phases, including goal-setting and task-planning in a forethought phase, monitoring performance on those tasks in a performance phase, and reflecting on learning outcomes and adapting to new strategies in a self-reflection phase. Many learning dashboard studies that argue their designs to support self-regulated learning usually only support the performance phase (Jivet et al., 2018; Matcha, Gašević et al., 2020). There is a scarcity of conclusive findings on the extent to which supporting all three phases of self-regulated learning in the design of learning dashboards would yield different cognitive and affective learning outcomes compared to supporting only the performance phase (i.e., which has been a common design approach). By addressing this research gap, the study contributes to a deeper understanding of the impact of dashboard design on student learning outcomes. Therefore, the main research question is defined as follows:

Research Question: *How - and to what extent - does providing support for all three phases of self-regulated learning in a learning dashboard impact cognitive and affective learning outcomes differently compared to only supporting the performance phase?*

This study focuses on how the design and evaluation of student-facing learning dashboards can be grounded in self-regulated learning theory, and to what extent the (in)consideration of the theory's phases yields different cognitive and affective learning outcomes. While doing so, the underlying goal resides in providing targeted design recommendations that could possibly resolve some of the shortcomings in current learning dashboards. The relevance of this study for the field of Human-Technology Tnteraction lies in its focus on designing and evaluating learning dashboards with the ultimate goal of optimizing students' online learning experiences. With the growing importance of online education, it is becoming increasingly important to design technologies that can facilitate effective learning experiences. Learning dashboards can play a crucial role in this regard, as they provide students with access to key information about their progress, goals, and achievements, and help them stay on track with their learning objectives. By exploring how students use and evaluate learning dashboards, this study aims to contribute to the development of more effective and user-friendly dashboards, and therefore help to inform the design of learning technologies that are better aligned with students' needs and preferences and have a sustained long-term impact.

Overview: The next chapter discusses the literature underlying self-regulated learning theory, its application in current learning dashboards as well as empirical findings from previous dashboard studies. Chapter 3 discusses an interview study that was performed with students to understand how the empirical findings from previous dashboard studies correspond to personal experiences, and to what extent students' self-reported needs for specific features can be realized within a dashboard design that is guided by self-regulated learning theory. Findings from the literature and the interview study guided the design of two new learning dashboards, which are demonstrated in Chapter 4. Chapter 5 and Chapter 6 describe the methods and the results of an experimental study in which the hypotheses were tested by evaluating and comparing the two learning dashboards. These results will be discussed with respect to the research question and findings from the literature and interview study in Chapter 7.

2. Theoretical Background

This literature review starts with introducing two models of self-regulated learning theory that are relevant to the design and evaluation of learning dashboards. After that, several shortcomings in the design and evaluation of current learning dashboards - as brought forward by previous systematic literature reviews - will be discussed in relation to both models of self-regulated learning theory. Since there is a scarcity of experimental studies in which these claims are empirically tested, the remainder of the chapter is devoted to formulating hypotheses to the research question by examining the extent to which suggested shortcomings are reflected in empirical findings from previous dashboard studies.

2.1 Self-Regulated Learning Theory

Self-regulated learning refers to one's ability to monitor, direct and regulate their actions in order to reach some pre-defined goal (e.g., academic achievement or skill-improvement). Zimmerman and Schunk (2001) describe self-regulated learning as a cyclical process whereby a learner goes through successive phases of learning. During the forethought phase, learners set goals for themselves and plan for the tasks needed to reach those goals. These goals are guided by their knowledge of these tasks as well as metacognitive knowledge about their own strategies. During the performance phase, learners monitor their performance on achieving the goals and completing the tasks they set for themselves in the forethought phase. This is where actual learning happens, because the effectiveness of one's strategies and tactics is closely monitored in this phase. Eventually, learners reflect on the outcome of their performance in the self-reflection phase (e.g., in terms of a success/effort evaluation). Based on the evaluation of their learning in the self-reflection phase, learners formulate new goals for themselves and navigate through a set of new tasks they will need to complete to reach those goals. These tasks are ideally guided by new and more effective learning strategies, such that the cycle moves from the self-reflection phase back to the forethought phase (Panadero, 2017). In essence, self-regulated learning theory captures the way in which learners can understand and dynamically adapt to their learning environment through self-regulated learning skills. These skills include goal-setting, self-monitoring, self-assessment, self-instruction, and self-reinforcement (Schunk, 2008). Differences in academic achievements and success have been predominantly attributed to a learner's self-regulation skills (Schmitz, 2001).

Even though there are many different analytical models of self-regulated learning, all are situationally dependent and approach self-regulating learning in ways that are appropriate to a

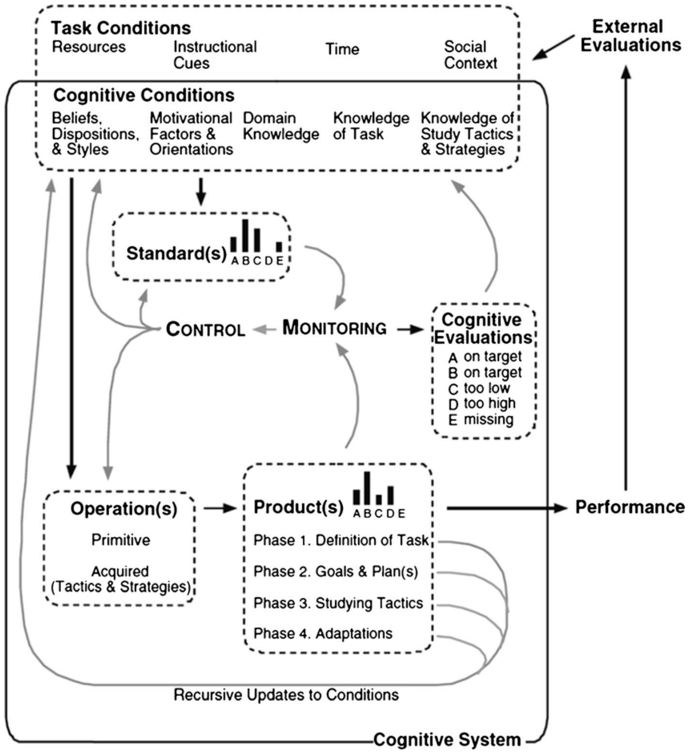


Figure 1: Schematic representation of the Winne and Hadwin (1998) COPES-model of self-regulated learning (Ranalli, 2012)

given learning context. One model that is widely used in the design and evaluation of computer-based learning environments is the Winne Hadwin-model of self-regulated learning (Winne & Hadwin, 1998). This model reconsiders Zimmerman's conventional cyclical model of self-regulated learning as a recursive process where five components work together dynamically within different stages of learning. These components include conditions, operations, products, evaluation, and standards, which together form the COPES-model. As shown in Figure 1 (Ranalli, 2012), learners formulate goals and devise the steps needed to reach them by considering internal and external constraints in the first stage of learning (Winne & Hadwin, 1998). The model refers to this initial stage of learning as 'task definition'. The internal constraints that delineate what a learner considers realistic for themselves - in terms of what they think they are capable of - are referred to as cognitive conditions. These are defined by a learner's knowledge of the task (e.g., learning material), their motivational factors and orientations (e.g., how much effort a learner is willing to offer), their beliefs and dispositions (e.g., their own interpretation of what a task demands from them), as well as knowledge of their learning tactics (i.e., metacognitive awareness of the strategies that they can apply). On the other hand, external constraints outline what a learner believes is realistic given the learning environment and are referred to as task conditions. As shown in Figure 1, these external constraints (i.e., task conditions) include time constraints, resource restrictions, the instructions learners are given, as well as their social environment. In the next stage, learners define expectations of how long certain tasks will take, and what quality and quantity can be achieved given the cognitive and task conditions (i.e, the internal and external constraints). When learning begins, learners operate and monitor their learning by using strategies and tactics that enable them to create so-called 'products' of learning (e.g., memorization, completion of assignments). Eventually, learners evaluate the products of their learning using the standards they have set for themselves in the first phase of learning, referred to as 'cognitive evaluations' in Figure 1. These products of learning can also be evaluated by an external entity (e.g., a teacher or a learning dashboard). The outcomes of the cognitive and/or external evaluations can result in the adaptation of new strategies, a reconsideration of earlier task definitions, or a revision of the goals and standards altogether.

2.2 Support for SRL in Current Learning Dashboards

Many studies on student-facing learning dashboards that state their designs to be grounded in learning sciences commonly argue to do so by implementing self-regulated learning theory (Jivet et al., 2017). However, the majority of these learning dashboards provide substantial support for only a small fragment of self-regulated learning as a whole (Bodily & Verbert, 2017; Matcha, Uzir et al., 2020). Very often, awareness and reflection are mentioned as the primary goal of learning dashboards (Jivet et al., 2018; Verbert et al., 2014). Within the Zimmerman and Schunk (2001) model of self-regulated learning, awareness and reflection only reside in the performance phase and – to some degree – in the self-reflection phase. Gaining awareness about the effectiveness of one's learning strategies is facilitated by presenting a learner with feedback on how they are performing, and thus helping them to monitor their progress in the performance phase. Reflection is facilitated during the self-reflection phase, but only to a limited extent. In the self-reflection phase, the aim of a learning dashboard is to facilitate a learner in generating new knowledge how effective previous learning strategies were. A learning dashboard ideally facilitates this by providing actionable feedback that learners can use to evaluate their learning and reformulate more appropriate and realistic goals for themselves. In that way, a learner can use these newly gained experiences and evaluations to steer course and plan for new tasks needed to reach more appropriate goals. In a sense, a learning dashboard should thus guide learners in closing the cycle of self-regulated learning by helping them to adopt more effective learning strategies to reach their learning goals, all of which are based on a performance evaluation that is facilitated by the learning dashboard. Even though many studies report their dashboard to be able to do so, the feedback their learning dashboards present are often not clear or actionable enough for learners to take the appropriate next steps (Schumacher & Ifenthaler, 2018). In terms of the Zimmerman and Schunk (2001) model of self-regulated learning, goal setting and planning in the forethought phase as well as post-evaluative strategy adaption in-between the self-reflection phase and the forethought phase are rarely facilitated in dashboard designs up until now (Jivet et al., 2018; Bodily & Verbert, 2017).

Boud and Molloy (2012) (p. 6) defined feedback as "a process whereby learners obtain information about their work in order to appreciate the similarities and differences between the appropriate standards for any given work - and the qualities of the work itself - in order to generate improved work". In essence, learning dashboards are a form of external feedback that learners use to evaluate their learning, their progress towards reaching their learning goals, and to revise the effectiveness of their study strategies. Assisting students support with evaluating their performance through feedback is an essential capability for learning dashboards because studies have shown that learners are often inaccurate at judging their own performance (Winne & Hadwin, 1998). However, the lack of support in the forethought phase and self-reflection phase could arguably limit the potential of learning dashboards to do so, as many studies have assumed so far that making a learner aware of their study progress (i.e., by including some visual indicator of how their learning thrives or fails) would automatically lead learners to take the appropriate next steps. In that way, awareness and self-reflection are assumed to trigger and guarantee action in the form of strategy adaptation. This assumption has led learning dashboard designs to rarely assist students in seeking a fitting course of action, leaving the responsibility to learners to interpret dashboard data appropriately and to revise and adapt their study strategies accordingly (Matcha, Uzir et al., 2020). One could argue that awareness and reflection are not supposed to be an end in itself for learning dashboards, but rather a means to be able to complete the cycle of self-regulated learning as a whole.

In terms of the Winne and Hadwin (1998) model of self-regulated learning, systematic reviews have argued that task conditions (i.e., external constraints to learning) are only minimally represented in most learning dashboard evaluation studies (Matcha, Uzir et al., 2020). For example, many dashboards disregard any estimation of the time needed to complete a task or the resources that are available to the learner. This consolidates earlier findings that dashboards rarely provide adequate support in the forethought phase of self-regulated learning (Schumacher

& Ifenthaler, 2018). Similarly, internal constraints such as a learner's knowledge of their study tactics and motivational orientation (i.e., cognitive conditions in Figure 1 (Ranalli, 2012)), are rarely addressed properly in learning dashboards. In addition, (Matcha, Uzir et al., 2020) argue that there seems to be a significant lack in dashboards that have shown to meaningfully incorporate the nature of the tasks that a learner is working on, impairing the extent to which they are able to provide contextually appropriate and meaningful feedback. If learning dashboards are not seamlessly integrated into a learner's online environment - and are thus not fully able to provide feedback that is contextually appropriate - they are arguably incapable of providing proper support in the forethought and self-reflection phase of self-regulated learning. Current educational psychology research stresses the importance of providing effective feedback on learning strategies and tactics to advance a learner's further personal development, as learners are often unaware of the most effective study strategies (e.g., spaced learning and repeated self-testing). If learning dashboards are simply able to remind learners of these study strategies and tactics, chances are already significantly increased that these students will actually adopt them in their learning (McCabe, 2011). Only when learning dashboards actually trigger further action (i.e., through actionable recommendations) do they cover the full cycle of self-regulated learning. In summary, whereas most learning dashboards provide only substantial support during the performance phase of self-regulated learning, only few seem to provide proper support in the forethought phase (i.e., assistance in setting realistic goals and tasks) or self-reflection phase (i.e., assistance in steering course to reach those goals) (Jivet et al., 2018).

2.3 Empirical Findings from Previous Dashboard Studies

Naturally, the question remains how substantial these differences in the implementation of self-regulated learning theory truly are. If the effects that a learning dashboard's (in)consideration of the theory's phases could have on online learning are that strong, then that should somehow translate into empirical findings from previous dashboard studies. The following section will discuss patterns in the results from several dashboard studies and how these relate to varying implementations of self-regulated learning theory. To navigate the vast body of literature on learning dashboards more efficiently, the following sections will formulate relevant hypotheses to three sub research questions:

Research Question 1*: How is a learning dashboard that supports all three phases of self-regulated learning evaluated differently compared to a learning dashboard that only supports* *the performance phase?*

Research Question 2: *What cognitive learning outcomes are impacted differently when having access to a learning dashboard providing support for all three phases of self-regulated learning compared to a learning dashboard only supporting the performance phase?*

Research Question 3: *What affective learning outcomes are impacted differently when having access to a learning dashboard providing support for all three phases of self-regulated learning compared to a learning dashboard only supporting the performance phase?*

2.3.1 Perceived Usefulness and Frequency of Use

Kim, Jo and Park (2015) conducted an experimental study in which students in a course were randomly allocated to the treatment-group that received access to a learning dashboard after a mid-term exam (i.e., through the online course page), or to the control-group that was not. The learning dashboard contained visualizations of the student's log-in frequency in comparison to others, as well as how they navigated the course pages differently (i.e., views in a certain course module or repository). Clickstream data from the course page as well as students' average grades were measured for all students. Those that were assigned to the treatment-group were given an additional questionnaire on the usability of the dashboard. The results from Kim et al. (2015) show that students in the treatment-group received higher final course grades than those in the control-group. Kim et al. (2015) explained the effect of the dashboard on final course grades as being mediated by increased self-regulated learning behaviour. The authors proposed that the learning dashboard provided students with information about their own learning process, which facilitated their ability to self-monitor their progress and adjust their learning strategies accordingly. However, the authors did not directly measure these factors, so their impact is more speculative. It's also possible that the dashboard increased course grades simply by increasing motivation and engagement (i.e., which the authors argued increased). These changes may have ultimately led to higher grades for students in the treatment group. Even though students in the treatment-group evaluated the dashboard to be user-friendly, the frequency of use decreased progressively during the timespan that students had access.

In another study by Corrin and Barba (2014) that followed a mixed-method design, data on students' academic achievements and interaction with the learning management system were used to populate a dashboard that displayed both individual performance data as well as class averages (e.g., scores on tests, log-in frequency, modules completed). Before students received access to the dashboard, they were asked to complete a survey on their motivation and goal orientation. The dashboard was then evaluated by students in an interview using a think-aloud procedure. Even though students found the dashboard to provide impressively detailed information, they couldn't actually describe how the dashboard would be able to assist them with their learning. As a result, having access to the learning dashboard did not truly lead to behavioural change, an improvement in learning outcomes or an increase in motivation. As Corrin and Barba (2014) conclude themselves, students were not able to connect the dashboard information to a needed behavioural change. Similar to Kim et al. (2015), use of the dashboard decreased progressively over time.

Even though the dashboards in the studies by Kim et al. (2015) and Corrin and Barba (2014) were perceived to be usable and their overall acceptance was high, students in neither of the studies used the dashboard again when they were not prompted to do so. In addition, a majority of the students in both studies were not able to describe how either of the dashboards would be able to assist them with their learning in the long run. While the study by Kim et al. (2015) found that the learning dashboard increased students' final course grades, the authors could not directly measure the impact of the dashboard on self-regulated learning behavior, and it is possible that the increased grades were due to increased motivation and engagement. Both studies – like many other dashboard studies (Bodily & Verbert, 2017) - were conducted in a controlled setting where students willingly participated and agreed to be prompted to use the learning dashboard. If the overall use of their dashboards remained low even amongst these favourable conditions, it is arguably risky to assume that in a 'real' situation - where use is completely voluntary – frequency of use would not be even lower, let alone guaranteed. Learning dashboards in reality cannot support the improvement of a student's success if they are not being used outside the experimental setting. Even though frequency of use is not believed to be a strong predictor of learning outcomes (Kim et al., 2015), a dashboard's perceived usefulness has – in fact - been found to be a positively correlated with learning outcomes (Matcha, Uzir et al., 2020). Even though Kim et al. (2015) and Corrin and Barba (2014) have provided valuable insights on how to improve the usability and acceptance of learning dashboards, it is important to understand what aspects of the learning dashboards' way of providing feedback can explain their low perceived usefulness and frequency of use.

2.3.2 Data-Driven, User-Centered and Theory-Driven Design Approaches

There are some similarities to be observed in the dashboards from Kim et al. (2015) and Corrin and Barba (2014). Both dashboard designs seem to rely heavily on visualizing course grades and interaction with the learning management system (e.g., log-in frequency, clickstream variability, completed modules) rather than on indicators related to the course material specifically (e.g., achieved skills, knowledge gained). Similar to the dashboards from Kim et al. (2015) and Corrin and Barba (2014), the designs of current learning dashboards commonly follow a data-driven approach by relying on quantifiable platform-interaction data (e.g. number of resources consulted, log-in length and frequency). Several systematic reviews of dashboard studies have found this to be a trend within dashboard design. Instead of using theoretical frameworks outlining how to assist students during their learning processes effectively, many dashboard designs seem to be guided by the data that is generated by the learning management system onto which they are built (Bodily & Verbert, 2017; Verbert et al., 2014). This is not necessarily surprising. Visualizing observable online behaviour is cheap, quick and relatively easy considering that modern learning management systems automatically capture vast amounts of clickstream data. However, it is unclear to what extent this 'observable' behaviour accurately represents 'learning' behaviour – and by extension – to what extent a learning dashboard can truly assist learners if its feedback is solely based on this type of 'overt' data (Jivet et al., 2017, 2018). In fact, studies have shown that presenting quantitative metrics like these rarely offer any actionable guidance to a learner's study strategies (Kovanovic, Gašević, Dawson, Joksimovic & Baker, 2016). A considerable contribution to this shortcoming is that limited research has been done on what data types are useful for visual representation in learning dashboards (Sansom, Bodily, Bates & Leary, 2020).

While the previous two studies relied heavily on visualizing course grades and interaction with the learning management system, the design of the learning dashboard StepUp! by Santos, Verbert, Govaerts and duval (2013) followed a user-centered design approach. During the design process, the researchers sat together with students in brainstorming sessions to identify learning issues and formulate needs for a new learning dashboard. Visualizations were developed that incorporated an indication of how much time a student was expected to spend on a task, a list of resources the students could consult to finish an assignment, and a bar chart depicting how other students spent their efforts differently within the course (e.g., discussion posts, use of software). Participating students were asked to provide an assessment of how much the learning dashboard supported them in parts of their learning that they felt they were struggling with. Students found StepUp! to be usable and acceptable, and they also reported better understanding of their study habits and how others studied differently. In addition, students consistently used the dashboard outside brainstorming and evaluation sessions. These findings suggest that a user-centered design approach can lead to a more effective dashboard that students actually want to use, which stands in contrast to the data-driven approach often seen in other dashboard designs. StepUp! did not significantly increase motivation nor course grades, although the authors argue that this was because motivation and course grades were already above average.

Learning dashboard Mastery Grids (Loboda, Guerra, Hosseini & Brusilovsky, 2014) was designed using a theory-driven, user-centered design approach. Mastery Grids uses an innovative visualization technique to display the resources a learner could access in a course (e.g., questions, examples, readings, lecture recordings). These resources are represented as vectors across two dimensions: an intensity level that shows how much the learner has mastered a certain skill using that resource, and an intensity level that shows how much other learners have consulted that resource. The skill mastery level was measured using assessments (e.g., quizzes or exams) that were designed to test the learner's understanding of specific concepts or skills. The learner's performance on these assessments was used to determine their level of mastery for each skill. In Mastery Grids, this level of mastery is represented as an intensity level on one axis of the visualization. The higher the level of mastery, the more intense the color of the corresponding vector. Follow-up iterations of the dashboard were developed using self-reported needs from students. The dashboard was deployed for all participating students in three courses. Participants were asked to complete a survey on the dashboard's perceived usefulness, satisfaction and usability. In addition, usage patterns of the dashboard were tracked with clickstream data. The results of Loboda et al. (2014) show that students who used Mastery Grids performed better on quizzes, evaluated the dashboard to be useful, satisfying to use and usable. Similar to Santos et al. (2013), frequency of use was sustained outside the evaluation sessions. In addition, students indicated that the dashboard was able to guide them to resources that were more suited for their level of understanding of the material. Besides making the student aware of their progress and achievements, the dashboard arguably provided actionable feedback that student could use to steer course and plan for more appropriate tasks need to reach their goals. Mastery Grids borrowed theoretical frameworks from cognitive sciences and learning theory, and Loboda et al. (2014) argue its design to be guided by established instructional design principles. Using the cognitive sciences and learning theory as a theoretical foundation, Mastery Grids emphasizes the importance of active, learnercentered approaches to instruction that promote deep learning and mastery. Instructional design principles such as breaking learning into manageable chunks, providing opportunities for practice and feedback, and scaffolding support for learners are also incorporated into the design of Mastery Grids to optimize learning outcomes. As illustrated by Jivet et al. (2018), current educational psychology research stresses the importance of providing effective feedback on learning strategies and tactics to advance a learner's further personal development, as learners are often unaware of the most effective study strategies.

2.3.3 Supporting SRL with Course-Integrated Dashboards

The dashboards by Kim et al. (2015) and Corrin and Barba (2014) only minimally incorporated course-specific indicators of a task to be completed (i.e., external constraints) or representations of a learner's study tactics and motivational goals (i.e., internal constraints). On the contrary, the dashboards developed by Santos et al. (2013) and Loboda et al. (2014) were grounded in the course material significantly more and refrained from visualizing clickstream data only. In line with the findings from Matcha, Uzir et al. (2020), these dashboards seemed to properly represent both external constraints (e.g., visualizations of course-based tasks to be completed including indications of how long they are supposed to take) and internal constraints (e.g., visualizations that represent the learner's current study tactics). The dashboards developed by Santos et al. (2013) and Loboda et al. (2014) are more closely aligned with the actual course material and tasks, and therefore provide more relevant information to the learner in terms of what they need to do to succeed in the course. This, in turn, can help learners to plan their tasks and set appropriate goals, which are key aspects of the forethought phase of self-regulated learning. By contrast, the dashboards developed by Kim et al. (2015) and Corrin and Barba (2014) were found to be less grounded in the course material and provided only minimal information about the tasks to be completed or the learner's study tactics and goals, which may limit their effectiveness in supporting task-planning and goal-setting. While these studies do not directly compare the dashboards, they do suggest that dashboards that are more closely aligned with course material are able to provide more meaningful assistance in goal-setting and task planning (i.e., providing more support in the forethought phase of self-regulated learning), and may therefore be perceived as more useful. These findings serve as a foundation for the first hypothesis to RQ1:

Hypothesis 1.1: *Perceived usefulness will be higher for a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the* *performance phase.*

In addition, the dashboards by Kim et al. (2015) and Corrin and Barba (2014) did not seem to provide any clear and actionable feedback on how the student could possibly improve their learning. Arguably, the dashboards' capability to do so was already restricted by the type of information that it provides. Since most visualizations were based on clickstream data, neither of the dashboards seem to be able to provide meaningful feedback to the student on what to do next within the course itself. On the contrary, the dashboards by Santos et al. (2013) and Loboda et al. (2014) provide actionable recommendations by highlighting resources that other students have used differently or more frequently. This seems to be reflected in higher levels of metacognitive awareness, as reported by both studies. By including features that support all three phases of self-regulated learning, a dashboard can provide a comprehensive and integrated view of the learning process, which can improve the user's understanding and awareness of their learning progress. For example, the inclusion of features that facilitate planning and goal-setting can help learners to better understand what they need to do to achieve their learning goals. Similarly, the inclusion of features that support monitoring and reflection can help learners to identify areas where they need to improve and adjust their learning strategies accordingly. These findings serve as a foundation for the second hypotheses (i.e., to RQ1):

Hypothesis 2.1*: Metacognitive awareness will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the performance phase.*

The inclusion of all three phases of self-regulated learning can improve trustworthiness in a few ways. First, a dashboard that includes all three phases can provide a more comprehensive overview of the learning process, which can increase the transparency and credibility of the dashboard. By including information on how the learner plans their learning, how they monitor their progress, and how they reflect on their learning, the dashboard can give a fuller picture of the learner's efforts and progress. While there is limited empirical evidence directly examining the relationship between the inclusion of all three phases of self-regulated learning and trustworthiness, there are several studies that indirectly suggest a positive relationship between the two. Research has shown that perceived usefulness and perceived ease of use are two key factors in determining the perceived trustworthiness of technology (Venkatesh & Bala, 2008). As mentioned earlier, dashboards that include all three phases of self-regulated learning may be perceived as more useful due to their ability to provide a more comprehensive view of a learner's progress and goals. Similarly, user-centered design principles can enhance ease of use, which can contribute to perceptions of trustworthiness. These findings serve as a foundation for the second hypothesis to RQ1:

Hypothesis 1.2: *Perceived trust will be higher for a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the performance phase.*

2.3.4 Improving Motivation for Different Goal Orientations

For a learning dashboard to serve as a valuable device for feedback, it needs to provide a representative frame of reference that students can use to make sense of the dashboard's visualizations (Wise, Zhao & Hausknecht, 2014). Two common types of reference frames in learning dashboards are individual reference frames and social comparison reference frames (Jivet et al., 2018). An individual reference frame considers the students' current state (i.e., performance, motivation, online behaviour) in relation to a previous state (i.e., to illustrate progression or lack thereof). A social comparison reference frame considers the students' current state (i.e., in terms of similar metrics) in relation to other students at that point in time. The majority of learning dashboards use a social comparison frame of reference in their designs by benchmarking a student's performance in relation to others (Jivet et al., 2018; Matcha, Uzir et al., 2020). Social comparison theory argues that the ability to compare oneself to others - when there is no other means of comparison – increases motivation and the drive to perform better (Festinger, 1954). In addition, studies in educational psychology have demonstrated that an increase in motivation is positively correlated with an improvement in learning outcomes (Ryan & Deci, 2000). Therefore, it would only seem intuitive to expect that the use of a social comparison reference frame in learning dashboards would ultimately – through higher levels of motivation – result in better learning outcomes than a dashboard that does not provide a social comparison reference frame.

However, previous learning dashboard evaluation studies seem to have widely different conclusions on how appropriate the use of a social comparison reference frame is to increase motivation and improve learning outcomes (Jivet et al., 2018). Kim et al. (2015) found that the effects on motivation by having access to a learning dashboard that used social comparison as a frame of reference to evaluate performance was different for every student, depending on their academic achievement level. Their results showed that low-achievers' level of motivation generally increased more from being able to compare oneself to others than that of high-achievers. The study suggests that high-achievers are generally already highly-motivated students so that the marginal effect that access to the learning dashboard has on motivation is rather low. This is in contrast with findings from Tan, Yang, Koh and Jonathan (2016). In an experimental study, differences in final grades, clickstreams and motivation were evaluated within a course between a group of students that had access to a learning dashboard that displayed visualizations of performance in relation to others (i.e., course grades and online activity) and a group that did not. Their results indicated that high-achievers – in contrast to Kim et al. (2015) – reported much higher levels of motivation than low-achievers did. As an explanation, Tan et al. (2016) suggest that this might be due to 'healthy peer pressure' and 'informal competition'.

The stark contrast between the findings from Kim et al. (2015) and Tan et al. (2016) seems surprising at first. However, a more elaborate comparison with other dashboard studies that evaluated motivation seems to provide a potential explanation for these discrepancies. In the previously mentioned mixed-method study by Corrin and Barba (2014), students reported that being able to evaluate their own performance in comparison to others increased their motivation. Nonetheless, more than half of those students also indicated that the ability to compare themselves to other students distracted them from the goals they had set for themselves at the beginning of the course. Once students saw that they performed better than the average student (i.e., in terms of the given metrics), they tended to feel demotivated as the additional gain of working harder was perceived to be only minimal. Similarly, students that performed worse than the 'average' student reported feeling motivated through comparison with others only if the average performance of other students was the level that they were trying to achieve. If their target grade was lower than then the average grade of other students to begin with, social comparison did not have a negative nor positive effect on motivation. In a qualitative interview study on student-facing learning dashboard PeerLA (Konert, Bohr, Bellhauser & Rensing, 2016) - which was evaluated using self-reported subjective measurements like motivation and engagement - a majority of students reported feeling more motivated after using the dashboard. In contrast to Corrin and Barba (2014), students in this study did not indicate that they experienced any distraction from their individual goals, even after having the ability to see the results from other students. The biggest difference between the dashboards developed by Konert et al. (2016) and Corrin and Barba (2014) is that PeerLA required students to indicate their grade goal, and subsequently only showed the results of other students with similar goal orientations; Blackboard by Corrin and Barba (2014) did not.

This suggests that for social comparison to be an effective frame of reference, learning dashboards need to account for different goal orientations. The findings from Konert et al. (2016) – whose design yielded a steady increase in motivation – arguably provide a valuable explanation for widely varying fluctuations in measurements of motivation by Kim et al. (2017) and Tan et al. (2016). Neither of their dashboard designs supported any functionality that allowed students to set goals for themselves (e.g., target grade), and thus provide feedback based only on students with similar goal orientations. This argumentation is consolidated by an experimental study by Fleur, van den Bos and Bredeweg (2020) which examined the effects of having access to a learning dashboard on students' motivation within a course. Their learning dashboard used a social-comparison reference frame in combination with goal orientation; students were able to see their average grade in comparison to the average grade of 9 anonymous peers with a similar grade goal orientation. Their results show that the majority of those who had access to the dashboard reported higher levels of extrinsic motivation and ultimately obtained higher course grades than those who did not. However, motivation initially increased when students first received access to their learning dashboard while it decreased progressively after students had access for a while. Fleur et al. (2020) argue that this might have been due to a novelty effect wearing off (i.e., a dashboard's initial effect on learning outcomes might diminish after prolonged use) or simply due to course-related changes in motivation (e.g., students might become less motivated further into the course due to stress). It is possible that course-related changes could also influence the control group and not just the group with access to the learning dashboard. However, the study by Fleur et al. (2020) was designed as a pre-post experimental study where participants were randomly assigned to either the treatment group (with access to the learning dashboard) or the control group (without access to the learning dashboard). This design helps to control for some of the potential confounding factors that may influence the study outcomes, including course-related changes in motivation.

In a sense, the question remains whether these findings imply that a learning dashboard that supports the forethought phase – through goal-setting features that account for goal orientation – could ultimately result in higher (i.e., and more sustained) levels of motivation than a learning dashboard that only visualizes performance. By extension – through considering goal orientations in the forethought phase – learning dashboards can arguably provide more valuable feedback in the self-reflection phase by giving study strategy recommendations that are based on other students with similar goal orientations. In contrast with Kim et al. (2015), Corrin and Barba (2014), and Tan et al. (2016) (i.e., dashboards not accounting for differences in goal orientation) - but in accordance with Konert et al. (2016) (i.e., dashboard accounting for differences in goal orientation) - it seems that Fleur et al. (2020) demonstrated that accounting for differences in goal orientation reduces the fluctuations in motivation in learning dashboards that use social-comparison reference frames. As such, it seems that using social comparison as a frame of reference for learning dashboards can only support improvements in motivation under the condition that the design accounts for goal orientation. Goal orientation functionality is – by being a 'goal-setting' feature – in essence a core element within the forethought phase of self-regulated learning theory. Therefore, a learning dashboard that supports all three phases of self-regulated learning can provide students with the tools they need to set their goals, plan their strategies, monitor their progress, reflect on their learning, and make adjustments as needed. This, in turn, can help them feel more motivated under the strict condition that their goal orientation is taken into account properly. Considering that accounting for goal orientation is a defining feature in the forethought phase, a learning dashboard that only supports the performance phase of self-regulated learning will be restricted in meeting these conditions. These findings serve as a foundation for the first hypothesis of RQ3:

Hypothesis 3.1*: Motivation will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

There is a lack of empirical evidence that directly compares the differences in perceived autonomy resulting from different implementations of self-regulated learning (Matcha, Uzir et al., 2020). Providing support for all three phases of self-regulated learning could possibly contribute to a greater sense of perceived autonomy because it can help students to better understand their own learning processes and make more informed decisions about their learning (Kuo, Walker, Schroder & Belland, 2014). By providing information on goal-setting, monitoring progress, and reflecting on learning outcomes, these dashboards can give students a greater sense of control over their own learning, which can contribute to a greater sense of autonomy (Hardebolle, Jermann, Pinto & Tormey, 2019). In contrast, dashboards that only support the performance phase of self-regulated learning may provide limited information about the student's learning process and progress, which can limit their ability to make informed decisions about their learning. Despite the lack of empirical findings in previous dashboard studies that can support this reasoning, the current study aims to investigate the potential differences in perceived autonomy among varying implementations of self-regulated learning in learning dashboards by testing the following hypothesis for RQ3:

Hypothesis 3.2: *Perceived autonomy will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

2.3.5 Tying the Knots: How Can Learning Dashboards Improve Course Grades?

Previously mentioned studies that have implemented self-regulated learning principles in their learning dashboards have reported positive effects on course grades (Kim et al., 2015; Fleur et al., 2020; Corrin & Barba, 2014; Tan et al., 2016). While many other learning dashboard studies also report an increase in course grades (Jivet et al., 2018), it is important to consider whether these effects are really due to the implementation of self-regulated learning theory. While many learning dashboard studies claim to support self-regulated learning, they often fail to measure differences in self-regulated learning effectively, raising questions about whether the positive effects on course grades are actually due to improved self-regulation or to other factors, such as higher motivation, as suggested by Kim et al. (2015). AB-testing helps to isolate the effects of specific design features, such as supporting all three phases of self-regulated learning, from other potential confounding variables. By using AB-testing, researchers can be more confident that the positive effects of a learning dashboard on learning outcomes are truly due to the specific dashboard design being tested, rather than to other factors that might be present in the study. However, previous learning dashboard studies have rarely used AB-testing to measure the effects of supporting all three phases of self-regulated learning on learning outcomes. As a result, it is unclear to what extent an increase in course grades can be attributed to the way self-regulated learning is implemented in these dashboards. Furthermore, given the heterogeneity of students across different dashboard studies, it is difficult to compare the effectiveness of different implementations of self-regulated learning, leaving open the question of how learning dashboards that support self-regulated learning truly improve course grades.

The final hypothesis will be reasoned through a combination of empirical findings. As argued in the previous section, previous dashboard studies have shown that learning dashboards that provide support for all three phases of self-regulated learning (i.e., forethought, performance, and self-reflection) have been associated with higher levels of motivation (Kim et al., 2015; Fleur et al., 2020; Jivet et al., 2018; Matcha, Uzir et al., 2020). Motivation has been shown to positively impact learning outcomes (Ryan & Deci, 2000). Secondly, studies have suggested that learning dashboards that incorporate social-comparison feedback and goal orientation support in the forethought phase are associated with increased motivation and ultimately higher course grades (Fleur et al., 2020; Konert et al., 2016). This finding is further supported by research indicating that goal setting and monitoring are key components of self-regulated learning, and have been linked to improved learning outcomes (Zimmerman & Schunk, 2001; Schunk, 2008). Finally, learning dashboards that provide feedback and support for the self-reflection phase have been shown to improve metacognitive awareness and study strategy adjustments (Bodily & Verbert, 2017; Jivet et al., 2018; Santos et al., 2013; Loboda et al., 2014). These factors are important for optimizing learning and improving performance. Taken together, it is hypothesized that a learning dashboard that supports all three phases of self-regulated learning would be associated with higher course grades compared to a learning dashboard that only supports the performance phase. This is because the former would provide students with the necessary support and feedback to set and achieve goals, monitor progress, and adjust study strategies accordingly:

Hypothesis 2.2: *Course grades will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

2.4 Current Study

This study examined the differences in students' cognitive and affective learning outcomes between a learning dashboard that supports all three phases of self-regulated learning (i.e., forethought phase, performance phase and self-reflection phase) and a learning dashboard that only supports the performance phase. By doing so, the study aimed to formulate targeted design recommendations that could resolve some of the shortcomings in current learning dashboards. Even though it is arguably safe to expect that - at least to some degree - most learning dashboard studies considered the needs of their target users (i.e., students), only a minority of these studies reported performing a formal needs assessment (Bodily & Verbert, 2017). The underlying motive of learning dashboards, regardless of the design approach taken, is to improve a student's cognitive and affective learning outcomes. Therefore, it is only sensible to go beyond theoretical expectations of the most effective way to design a learning dashboard by also consulting students themselves. This way, it becomes easier to understand how the empirical findings in previous dashboard studies correspond to personal experiences, and to what extent students' self-reported needs for specific features can be realized within a design that is guided by self-regulated learning theory. A qualitative interview study was performed with students at the Eindhoven University of Technology – guided by insights from the literature - to understand and formalize their needs for a hypothetical new learning dashboard for Canvas (i.e., the university's learning management system). Through this process, the current study followed a theory-driven, user-centered design approach in which the results from the interview - together with findings from the literature review – were used as guidance for constructing design features that support each phase of self-regulated learning in a way that is appropriate to the given context of use. In an experimental study, the two learning dashboards were evaluated and compared using measurements that are relevant to the theory being discussed.

3. Interview Study

3.1 Methods

3.1.1 Participants and Design

An exploratory, qualitative interview study was performed through individual semi-structured interviews. Participants were recruited through the research-studies participant database of the Human-Technology Interaction-group at Eindhoven University of Technology. Participants were required to speak English, do full-time studies in the first semester of the academic year 2022/2023, and be familiar with using the university's learning management system Canvas. Previous experience with learning dashboards was not required. After excluding the interview data from one participant due to technical problems, the interview data of 5 participants were used for further analysis.

3.1.2 Materials

Before the interview session started, participants were shown three examples of existing student-facing learning dashboards to avoid any potential misunderstandings with regards to the focus of the study. These examples were selected based on their relevance to the interview questions, the scope of the study, and their applicability to Canvas. An interview guide was constructed consisting of eight questions covering several different topics (Appendix B). The first two questions were introductory and addressed the participant's past experiences with learning dashboards and their first thoughts on what kinds of features or visual information the participant would find valuable in a learning dashboard for Canvas. The next three questions separately discussed the three phases of self-regulated learning theory by addressing how - and to what extent - the participant thinks a learning dashboard would be able to support that phase. The questions were formulated in a way that refrained from using theoretical or technical terminology while still addressing key aspects for each of the phases. The question that addressed the forethought phase, for example, addressed whether and how (i.e., in terms of visual information or features) a learning dashboard could help them set goals and plan tasks for themselves. The sixth question addressed potential effects that the use of social comparison could have on affective learning, whereas the seventh question addressed how much the participant would trust the information and feedback that the learning dashboard would give to them. The last question addressed whether there was any type of information or feature that the participant would not want to have in a learning dashboard.

3.1.3 Procedure

Participants were invited via email to partake in an interview study on the design of a new learning dashboard. The interviews were scheduled to take 20 minutes and conducted online through communication platform Microsoft Teams. At the start of the session, each participant was asked to read and sign an informed consent form (Appendix A), stating the general purpose of the study and their willingness to participate. After the interviewee signed the informed consent form, the interviewer continued the session by introducing the idea of using a new learning dashboard for Canvas. The interviewer highlighted that this dashboard would be for one single course, specifically designed for the participant alone, and would not be accessible to other students or teachers. Afterwards, participants were shown three screenshots of currently existing learning dashboards that are in line with the scope and aim of the current study. The interviewer started the recording after confirming that the interviewee understood the general concepts and continued with the interview questions. After completion of the interview questions, the participants were debriefed, thanked for participation and rewarded a compensation.

3.1.4 Analysis

The audio recordings of the interviews were transcribed using Microsoft Teams' built-in meeting transcription tool. A deductive thematic analysis was performed on the data to find patterns of meaning across the interviewees' responses, delineated by the components of self-regulated learning theory. Parts of the transcripts that exhibited critical statements, feelings or thoughts in relation to the topic being discussed were coded and grouped together to form overall attitudes to specific learning dashboard elements. These attitudes were then categorized according to their position within the different phases of self-regulated learning.

3.2 Results

Forethought phase: respondents were divided on how a learning dashboard could truly help them with goal-setting and task-planning. Participants that were receptive to the idea expressed needs for to-do list type of features that facilitate in planning for reasonable study goals and future tasks. Two participants noted persistently that a learning dashboard should provide a comprehensive overview of the different contents that need to be addressed for completing a course successfully, ideally through calendar-type features containing a chronological display of deadlines and exams:

"It can be very daunting for students to get a grip of where to start [...] it would help with time management if you can streamline that a little more with a calendar [...] it could be a priority time schedule, like this deadline is coming up and you haven't downloaded the assignment yet."

Participants that responded negatively towards a learning dashboard assisting them in goal-setting and task-planning were largely sceptical towards its ability to truly consider differences between the ways that courses are organized. That is, participants argued that a learning dashboard should only assist in goal-setting for students if it reflects an understanding of the different ways in which that goal can be achieved. As the following statement illustrates, a learning dashboard should only assist in setting goals if it considers the nature of a given task appropriately:

"For theoretical courses, yes it could work. But it's hard to achieve for some courses that require you to program something [...] you just keep debugging for five hours and not making any progress so the dashboard won't be able to see that [...]"

Performance phase: participants addressed two interpretations of performance; as a proportional measure of progression through a course (e.g., number of completed modules) or as a measure of content-based achievements at a specific moment in time (e.g., grades in a quiz). Different examples of desired features were given to illustrate this consideration of performance, of which a 'progress bar'-type feature was mentioned by a majority of the participants:

"What I think I struggle most with during courses, is to get an understanding of how far I'm at [...] I think a progress bar of things I've completed would be beneficial to see how I'm doing."

Besides a visual representation of their overall progression through a course (e.g., a progress bar), participants generally considered course-grades as the only valuable, quantifiable measure of performance. Notably, all except one participant addressed comparing their course grades to other students even before the concept of social comparison as a frame of reference was introduced in an interview question. Most participants indicated that being able to compare course grades with peers would be a powerful benchmark to evaluate performance:

"For most of the people, comparing to other students will probably work [...] nobody wants

to be lagging behind, right? Everybody wants to be up front, I mean. Obviously I would kind of expect people to rather work harder that way."

Participants provided various explanations as to why social comparison would be an effective evaluation tool for them, including that it would motivate them to keep up with or even outperform the 'average student'. However, when asked to reflect on whether there would be situations where comparing their performance to that of others would actually demotivate them, participants expressed multiple concerns. For example, a recurrent theme was that the extent to which social comparison works beneficially could actually be dependent on their current level of performance to begin with. That is, participants frequently mentioned that comparing themselves to others would only have positive effects on their learning and motivation if their own performance was not too far off from the average performance of their peers:

"I would say if you're like slightly below the average, it's probably going to work in your favour, but if we're more below the average, it might actually backfire and be even more demotivating [...] if you're like not lagging behind but you see that other people are almost doing better than you, you can find probably some motivation to crank it up and stay ahead of the curve, [...]"

This might suggest that if a participant's performance would be much lower than that of peers, it would have a demotivating effect on their learning. Similarly, if their performance would be too far above average, participants felt like there would be no need to improve their performance even further. Generally, participants concluded that being able to compare themselves to others would have the most positive effects on motivation only if they would be performing slightly better or slightly worse than the average student.

"It depends on how close you are with your peers [...] I can just approach them very easily for materials or for help [...] but in a not so supportive community, I'm just ashamed of my performance and then I'm not really going to reach out for help."

Self-reflection phase: participants formulated several approaches for a learning dashboard to assist them in taking an appropriate further course of action, mostly mentioning a dashboard's use of targeted strategy recommendations. Some students mentioned that a learning dashboard's feedback should ideally address which parts of the course material the student has to focus on based on their previous course grades:

"Maybe just for each module a quiz [. . .] could help so that I know my knowledge on this topic [is] already good enough or not. And then like maybe it can connect to certain parts of the readings that are about that."

Also, participants indicated that an appropriate type of recommendation on how to revise their study strategy further could possibly demonstrate how strategies of peers differ from theirs. Being able to see how other students use an online learning platform to study was believed to prompt one to try out new strategies if their performance is below average:

"Yeah, maybe like which materials people access the most [. . .] would help me identify where to look for or like what to put more emphasis on [. . .] or if I am comparing to another year, we're doing significantly worse or not."

Even though participants were driven to come up with various ways in which a learning dashboard could provide feedback on revising their study strategy, they were notably reluctant to determine whether they would actually trust that feedback. Generally, participants indicated that they would only value a learning dashboard's feedback if it clearly reflected a performance evaluation that aligns with their own. Reasons included that trust needs to be built over time, and that a learning dashboard would first have to understand how the participant studied to be able to provide such feedback meaningfully. With regards to a learning dashboard's design, most participants stressed the need for simplicity. Examples of distracting features and redundant information included animations, advertisements, irrelevant graphs without explanation, a cluttered user interface, and too many required steps to complete basic tasks. Participants noted that they would rather want a learning dashboard to follow a simple structure that reflects the course material logically:

"I would like to see a clear hierarchy of information and maybe also information grouped along certain dimensions and certain aspects [. . .] Orders don't really work without any explanations. Like for me [it] should make some sense why some information is presented next to another piece of information [. . .] and the UI should not be cluttered on the surface like that."

3.3 Discussion

The aim of the interview study was to understand to what extent students' personal experiences and self-reported needs for specific features in a learning dashboard can be realized within a design that is guided by self-regulated learning theory. Overall, participants expressed a need for assistance in phases of self-regulated learning that many previous dashboard designs have failed to account for. The key features that participants consistently mentioned include:

- **Forethought phase:** a to-do list of course elements that are yet to be completed (e.g., assignment deadlines) and a calendar displaying the course's chronological events (e.g., exams, lectures). Participants expressed that both of these features should ideally incorporate a clear representation of how much time and effort tasks are supposed to take.
- **Performance phase:** a progress-bar type of feature that shows progression through the course's elements (e.g., completed modules) and visualizations showing academic progression over time (e.g., course grades). Participants expressed that an element of social comparison should be used cautiously; it is believed to only have beneficial effects on (affective) learning if participants are not too far off from class averages, and if they are only compared to other students that are like them or have the grade that they are aiming for.
- **Self-reflection phase:** recommendations on what resources the participant should consult (e.g., course material, Canvas-modules, additional readings). Participants indicated that these recommendations could be based on their own test results (e.g., 'you should rewatch this lecture') or based on how other students studied differently (e.g., 'students that performed better consulted this literature more frequently').

The type of forethought phase assistance that students desired (i.e., where goal-setting and task-planning are facilitated using calendars and to-do list features) has been seen to be lacking in previous dashboard designs (Jivet et al., 2018). Even though participants were explicit about these desires, they were questioning a learning dashboard's ability to do so effectively and meaningfully. Arguments usually revolved around the extent to which a dashboard can truly consider differences between the organization and goals of different courses. Arguably, a course's learning objectives need to be properly embedded in the way that learning dashboards present performance data. This way, situations where students deviate from the course structure (e.g., spending more time on assignments or parts of the material than expected) can be considered appropriately without misinterpreting it as a decline in learning. As such, for a learning dashboard to provide actionable feedback on how a student could increase their performance, it needs to show a clear consideration of the different ways in which the course's learning objectives can be approached and achieved. This finding is in line with the Winne and Hadwin (1998) model of self-regulated learning, which stresses the importance of a learning dashboard design incorporating some representation of external task conditions.

Whereas a majority of the participants initially indicated that being able to see the grades of other people would motivate them to perform better, many also expressed a concern about whether this would work in all cases. A positive effect on motivation would only be achieved if the participant's own performance was not too far off from the average of other students within that course. Performing far below the average or far above the average would demotivate them. This is in line with (Kim et al., 2015) where showing high-achieving students their performance in comparison to others increased participants' satisfaction with their academic achievement to a point that demotivated them to improve. Similarly, (Tan et al., 2016) found extremely low-achieving students to feel stressed and demoralised after being shown the performance of others. Whenever students lay within one of those extremes in terms of their performance in comparison to others, they tend to prefer a learning dashboard that contextualizes their performance in a more criterion-based, self-referenced way (i.e., in comparison to their previous performance). These findings suggest that a learning dashboard that adopts a social comparison reference frame should do so by displaying information that matches other students that are similar to the current user, or that matches other students that have the grades that the user is aiming for.

In terms of the self-reflection phase, many suggested approaches on how a learning dashboard could assist students in revising their study strategy were based on providing actionable recommendations. In a self-referenced way, those recommendations were largely based on how the dashboard could relate its measurement of performance back to the course material by providing recommendations on what course material the student needs to focus on based on their previous course results. This further stresses the need for a learning dashboard to be properly grounded in the way that the course is structured to be able to provide valuable guidance (Corrin, 2022). Similar to forethought-phase findings, a learning dashboard should be able to reflect a consideration of task conditions in order to provide students course-specific strategy recommendations for improving their performance (e.g., specifying which materials the student should access) (Winne & Hadwin, 1998). When comparing themselves to other students, participants addressed a need for the dashboard to provide recommendations based on how higher-achieving students studied differently. For both reference frames, the participants' level of trust in a recommendation was largely dependent on the extent to which the dashboard's evaluation of performance aligned with their own evaluation. This reflects the findings of the systematic literature review by Matcha, Uzir et al. (2020) that evaluated different learning dashboard design studies using the Winne Hadwin-model of self-regulated learning. There, the ability of a learning dashboard to consider different self-assessment standards to evaluate performance was demonstrated to be a necessity to increase actionability of feedback (Winne & Hadwin, 1998).

3.4 Preliminary Design Recommendations

For the sake of clarity and conciseness, the learning dashboard that supports all three phases of self-regulated learning (i.e., forethought phase, performance phase, and self-reflection phase) will from now on be referred to as *SRLdashboard*, while the learning dashboard that supports only the performance phase will be referred to as *PRFdashboard*.

With respect to forethought phase features, the design of SRLdashboard will include a calendar feature and a to-do list feature to represent external constraints. The underlying design objective will revolve around embedding them sufficiently in the course material so that they clearly represent external constraints (e.g, time constraints and assignment requirements). With respect to performance phase features, the design of SRLdashboard and PRFdashboard will include a progress-bar that shows progression through the course's elements (e.g., completed modules), a visualization showing academic progression over time (e.g., course grades), and visualizations of online activity within Canvas. The kind of online activity that will be visualized will be based on the technical capabilities of the software used to develop the dashboards as well as the clickstream data that offered by Canvas. With respect to self-reflection phase features, the design of SRLdashboard will include recommendations on what resources the user should consult (e.g., course material, Canvas-modules, additional readings).

As participants expressed that an element of social comparison should be used cautiously, an individual reference frame will be used in performance phase features. This decision was made to keep PRFdashboard as neutral as possible. Participants believed social comparison to only have beneficial effects on (affective) learning if participants are not too far off from class averages, and if they are only compared to other students that are like them or have the grade that they are aiming for. Therefore, SRLdashboard will use a 'mild' social comparison reference frame where results used to populate the self-reflection phase features will be based on the goal-setting functionality implemented in the forethought phase features. 'Mild' in this context means that the recommendations will not follow a strict, directive and comparative phrasing (e.g., 'you completed 2 modules this week while other students completed 5 modules this week'). Instead, recommendations will draw no direct comparison with the user and will be based on a subset of other students that have the grade that the user is aiming for (e.g., 'other students regularly consulted this module'). The following chapter will discuss how each of these features was implemented in more detail.

4. Designing a New Learning Dashboard

4.1 Design Objectives

In collaboration with the Information Management Services-team at TU/e, two learning dashboards were developed in Microsoft PowerBI using real-time data of students' interaction with Canvas. The dashboard features that support each of the three phases are based on a conceptualization of how students' self-reported needs for specific features in the interview study can be realized within the phases of self-regulated learning theory (section 3.4). The features that support the different phases of self-regulated learning are listed in Table 1 and will be demonstrated more elaborately in this chapter. A schematic diagram of the data model underlying the visualizations in SRLdashboard and PRFdashboard is included in Appendix C.

Table 1: Design features for the three phases of self-regulated learning

SRLdashboard	SRLdashboard | PRFdashboard	SRLdashboard
Forethought phase	**Performance phase**	**Self-reflection phase**
Grade goal slider (user)	Course grades (user)	Most-viewed items (course)
Course calendar (course)	Course progress (user)	Least-viewed items (user)
To-do list (course)	Submission timeline (user)	Online activity (course)
	Online activity graph (user)	
	Types of activity chart (user)	
	Most-viewed items table (user)	

Note. The features that support the different phases of self-regulated learning are listed are discussed and demonstrated more elaborately in section 4.3.

4.2 Data Sources

By default, Canvas clickstream data is automatically collected for students who are enrolled in the 0HV100 and 0HV30-courses at TU/e. The IMS-team at TU/e shared this data from students who provided consent in a pseudonymized format with the researcher through a secure Azure Databricks connection in Microsoft PowerBI. This enabled the researcher to fetch Canvas-indicators for these students (e.g., clickstreams, grades, submissions) as described by the Canvas Data warehouse dictionary (Canvas Data Portal, n.d.). To be able to visualize this data efficiently in a PowerBI dashboard, several aggregate tables were created using Power Query. These include a simplified 'online activity'-table derived from thousands of clickstream entries (i.e., representing information on how frequent a student was online and what items they accessed while they were online). Furthermore, a 'course results'-table was created from aggregated data

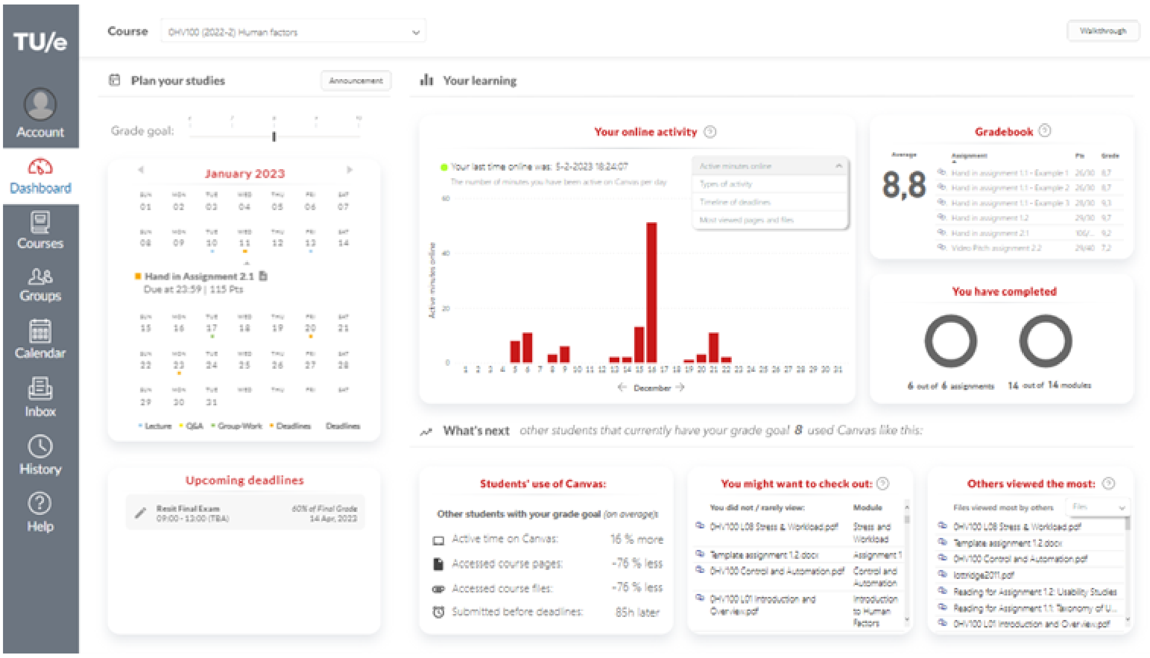


Figure 2: SRLdashboard included features supporting the forethought phase, performance phase and self-reflection phase of self-regulated

of students' timestamped submissions, grades and course-page progression. Finally, a 'student comparison'-table was created that evaluated course-wide averages of data from the 'online activity' and 'course results'-tables to that of a specific pseudo-user-id (see Appendix C for data model). For each student participating in the study, a so-called 'role'-profile was created within PowerBI using row-level security. The use of roles allowed for targeted filtering of the aggregated tables on that specific student as soon as they logged into their Power BI-dashboard, regardless of whether they had access to SRLdashboard or PRFdashboard. Both learning dashboards were published to the PowerBI webhosting service and were accessible to students through single sign-on with their TU/e account. Since each participating student was assigned a unique role within PowerBI, they were only able to access their own learning dashboard after logging in. Students could therefore not share the link to the dashboard with non-participating peers, as no role was created for their TU/e-account. The researcher was able to review a learning dashboard of a specific pseudonymized user-id, but in no way was it possible to infer which learning dashboard belonged to which TU/e-student.

4.3 Design Implementation

The two learning dashboards were developed using Microsoft PowerBI-desktop. Figure 2 displays a screenshot of SRLdashboard, Figure 3 displays a screenshot of PRFdashboard. The screenshots included in this report are that of real pseudonymized students participating in the experimental study. Feature implementations in the user-interface of both dashboards will be discussed according to the phases of self-regulated learning theory (Table 1). Both user-interfaces were built onto a static wireframe of the Canvas-homepage to emulate more realistically what a TU/e learning dashboard for Canvas could look like. When the student first logged into either of two dashboards, they were prompted to complete a small walkthrough tutorial that demonstrated and explained the features of the learning dashboard. Explanations of the different features were also accessible by clicking the (?)-icon next to each visualization. The calendar and to-do list were coded in HTML, CSS and JavaScript code and imported using an HTML-viewer widget in PowerBI. Any assignment, deadline or lecture presented in either of these features was hyperlinked to the corresponding Canvas webpage. A more elaborate demonstration of these secondary features and a link to a video demonstration are included in Appendix C.

4.3.1 Forethought Phase Features

The forethought phase features that were implemented in SRLdashboard are included in the far left-vertical panel of the user-interface under the heading ‘Plan your studies’ (Figure 2). Upon first log-in, students were prompted to indicate their grade goal using the slider. The slider-input was used to filter the self-reflection phase functionality under the ‘What’s next’ heading accordingly (section 4.3.3). This implementation was realised by only showing information based on other students that already have the grade that the user is aiming for. Additionally, the forethought phase features allow the student to view recent announcements posted to the course’s Canvas page, the course’s calendar (e.g., planned lectures, exams) as well as upcoming deadlines in the form of a to-do list (see Appendix C for more detailed screenshots).

4.3.2 Performance Phase Features

The performance phase features are implemented in the ‘Your learning’-panel of the user-interface and include visualizations of the student’s online activity, their gradebook as well as course progression pointers (Figure 3). Performance phase features were accessible to both SRL-

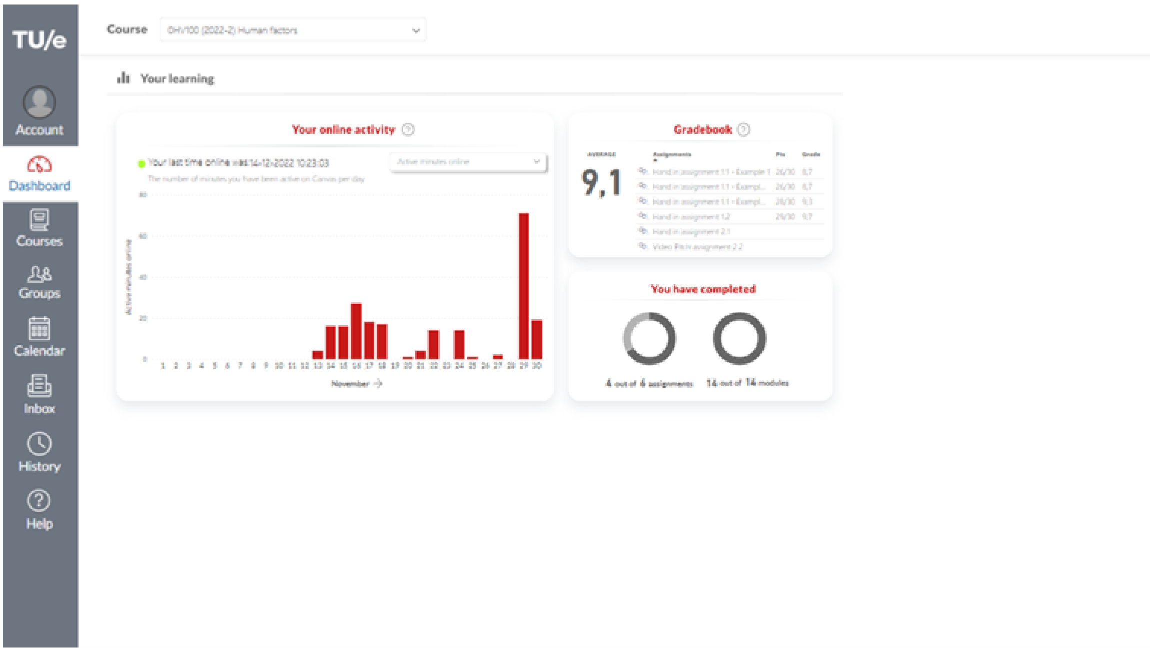


Figure 3: PRFdashboard only included features supporting the performance phase of self-regulated learning

dashboard and PRFdashboard. The online activity visualizations, presented under the 'Your online activity'-tab, include a bar chart of the students' online minutes per day, a categorical pie chart of the files they accessed the most within the course, a timeline of assignment deadlines and submission dates, and two tables with hyperlinked files and pages that the student has accessed the most (Figure 4). Each of these visualizations are accessible through a drop-down menu. All performance phase features use an individual frame of reference such that any graph, visualization or textual feature was based solely on the student's own data and not that of others (e.g., their time online that week, how they used Canvas, etc.) This was a deliberate choice, as PRFdashboard was considered to be a baseline-condition to which SRLdashboard would be compared, and therefore an individual reference frame was selected to rule out any confounding, uncontrolled effect that using a social comparison reference frame could have on cognitive and affective learning outcomes (i.e., as discussed in section 2.3.4). By using an individual frame of reference in performance phase features, it was ensured that the baseline condition (i.e., PRFdashboard) was as neutral as possible.

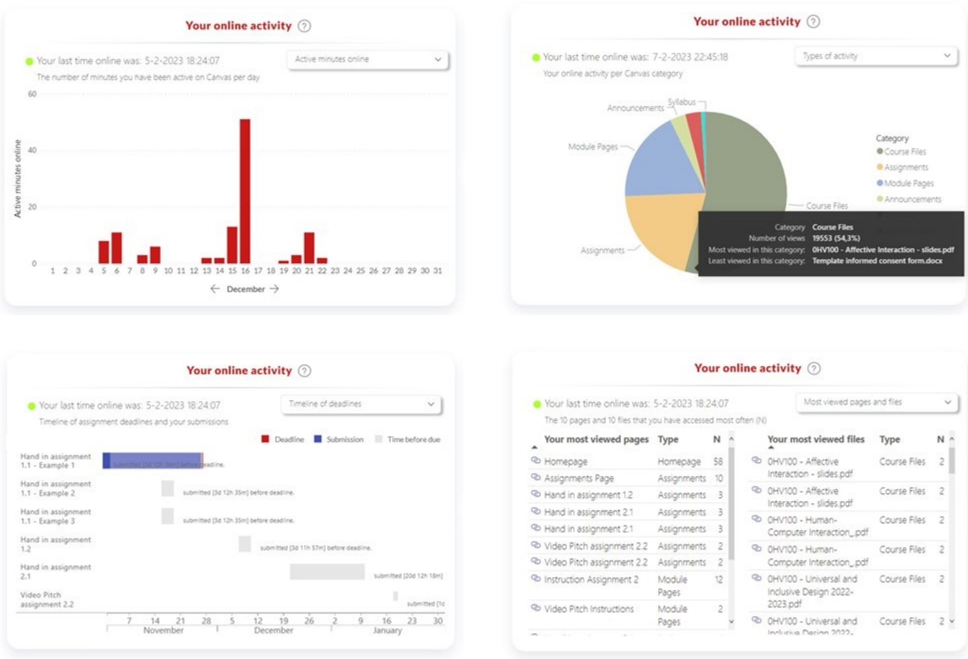


Figure 4: The four tabs of the 'Your online activity' visualization

4.3.3 Self-Reflection Phase Features

Self-reflection phase features – implemented in the 'What's next'-panel of the user-interface (Figure 2) - include three cards. Two of these cards include tables with the files and pages on the course page that other students have accessed the most. In addition, a third card displays how other students have used Canvas differently in terms of online activity and assignment deadlines. The data used in these features are filtered on the grade goal that the student has selected in the 'Plan your studies'-panel. That is, any data on how other students used Canvas differently or how their online learning differed from the student was based only on other students that have the grade that the user is aiming for. Due to the nature of the experiment and the practical limitations imposed by the study being part of a master thesis project (i.e., a strict and concise schedule), using only an individual reference frame for self-reflection phase features would have made it difficult to provide meaningful functionality. Only few changes were made to the Canvas pages of both courses during the five weeks that the experiment ran (i.e., uploads, published grades, etc.) The ability of the dashboard provide substantial feedback using just an individual reference frame was therefore limited (e.g., 'last week, you used Canvas differently in this way'). Instead, a social comparison reference frame was 'cautiously' implemented for self-reflection phase features (e.g., 'other students have used Canvas differently in this way'). Because social comparison can have unexpected and varying effects on affective learning outcomes (Kim et al., 2020), the goal-setting slider was implemented in the forethought phase features to allow students to set a grade goal for themselves. By adjusting the frame of reference in the self-reflection phase features to a subset of reference points that are relevant to that user only (i.e., other students that match what they're trying to achieve), the confounding effects of using a social comparison reference frame are expected to even out (Fleur et al., 2020). In this way, SRLdashboard was able to provide meaningful feedback in the 'What's next'-panel while minimizing confounding and unexpected effects on cognitive and affective learning outcomes due to using social comparison.

4.4 Piloting

Two small pilot studies were conducted to evaluate the preliminary design of SRLdashboard. In the first study, a wireframe was presented to two students, which included the same functionality as the final prototype described in section 4.3, but the functionality was not categorized into the three panels yet. The feedback from the first study suggested that the prototype did not have a clear structure to organize the overwhelming amount of functionality. In the second pilot study, three students were presented with a preliminary prototype that closely resembled the final SRLdashboard, and the features were categorized according to the three panels. However, the overwhelming amount of information displayed on the dashboard was still an issue. None of the students intuitively clicked the (?)-icons to retrieve more information about the dashboard's functionality, despite the icons being implemented. To address this issue, a walkthrough tutorial was developed, which prompted the user to complete it upon their first log-in (Appendix C). The tutorial followed a chronological structure, starting with the students indicating their grade goal and then demonstrating how the results in the lower-left panel were filtered based on that grade goal. Subsequently, students were guided through the rest of the user interface one feature at a time.

5. Methods

5.1 Experimental Design

In order to address the research questions and test the hypotheses, an experimental study was conducted using both the SRLdashboard and PRFdashboard. This study employed a randomized, repeated-measures, between-subjects experimental design, with participants randomly assigned to either the SRLdashboard-group or the PRFdashboard-group. Measurements were taken at three different points in time: prior to participants receiving access to the learning dashboard (referred to as "pre"), immediately after participants received access to the dashboard (referred to as "start"), and three weeks after participants had access to the dashboard (referred to as "end"). Throughout these three phases, differences in cognitive and affective learning outcomes were assessed through three surveys which participants were asked to complete (Figure 5). In order to account for time-varying differences in the measurements, an additional point of measurement was included right when participants received access to the dashboard. This allowed researchers to monitor how these differences progressed over time. In the second and third survey, participants were also asked to evaluate their dashboard in terms of perceived support for self-regulated learning, usability and usefulness, and trustworthiness. Clickstream data from Canvas and the dashboard were continuously tracked and updated once per day, from the moment that participants registered for the study until completion.

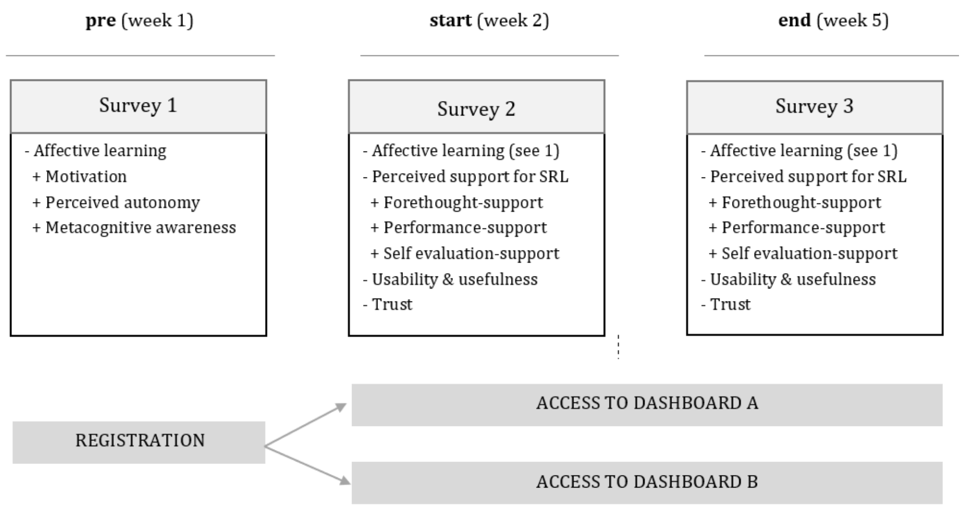


Figure 5: Experimental design and measurements

5.2 Measurements

Motivation [Survey 1, Survey 2, Survey 3] was measured using the Motivated Strategies for Learning Questionnaire (MSLQ) (Duncan & Mckeachie, 2010). The MSLQ was developed to measure learning strategies and academic motivation amongst college students. The entire survey, consisting of 81 questions in total, covers many aspects of motivation and learning strategies, including goal orientation, perceived task value, self-efficacy as well as test anxiety. A subset of 15 questions was used which the authors argue to be a reliable scale of academic motivation alone (α=.93). The questions consisted of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I participate in this course because I want to.', in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). The MSLQ-scale was chosen for this study as it makes a clear distinction between intrinsic and extrinsic motivation in academic settings. The questions were believed to be an appropriate measure of motivation as they reflect the nature of a learning dashboard essentially being a form of external feedback for students.

Perceived autonomy [Survey 1, Survey 2, Survey 3] was measured using the Personal Autonomy Scale (PAS) (Bei, Mavroidis & Giossos, 2020). The PAS-scale was developed to serve as a reliable psychometric scale to measure the perceived autonomy of college students. The 7 questions consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I can easily adapt to difficult situations.' (α=.74) , in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). The PAS-scale was chosen for this study as it was originally tested for distant-learning environments. Considering the nature of learning dashboards being a medium for online external feedback, the PAS-scale was thought to be a contextually appropriate scale to measure perceived autonomy.

Metacognitive awareness [Survey 1, Survey 3] was measured using the Metacognitive Awareness Inventory (MAI) (Schraw & Dennison, 1994). The MAI-scale was developed to measure students' self-assessments of how well they use strategies when working with academic material. The questions consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I try to use strategies that have worked in the past.' (α=.78), in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). The MAI-scale was chosen because it makes a clear distinction between knowledge of cognition as well as regulation of cognition. Both forms of metacognitive awareness are operationalized differently in a learning dashboard supporting self-regulated learning. As discussed in the literature review, a learning dashboard providing knowledge of one's own learning does not necessarily guarantee a change in learning strategy. Since a learning dashboard's aim is to encourage students to reflect on their learning and adjust their strategies accordingly, it is valuable to distinguish between the development of metacognitive awareness in terms of both knowledge of cognition as well as regulation of cognition.

Average grade [continuously] was measured using Canvas-data provided by the IMS-team at TU/e. This data included the grades of submitted assignments for each individual student. The average grade was constructed by computing a weighted average of these grades according to the grading criteria provided in the course's study guide.

Online activity LMS [continuously] was measured using timestamped clickstream activity within Canvas. Every unique timestamp was counted per day and time differences in between timestamps were calculated. If any time difference between two timestamps was less than 5 minutes, the timestamps were seen as part of a single session. The total amount of sessions – including each session's length – was used to calculate the number of minutes online per day. The Canvas Data warehouse dictionary also provides the number of sessions and corresponding session-length during a given period. However, session-length is based on the single sign-on system that the learning management system uses. As long as the student is logged in to any TU/e platform, the system registers it as a single session. This was considered to be an unreliable measure of online activity as these sessions persist as long as the student doesn't log-out manually which may span several weeks.

Support for self-regulated learning [Survey 2, Survey 3] was measured using a self-constructed scale that was based on the Online Self-Regulated Learning Questionnaire (OSLQ) (Barnard-Brak, Lan, To, Paton & Lai, 2009). The OSLQ-scale was chosen as it was developed explicitly to measure self-regulated learning skills in online learning environment. Due to time constraints, the experiment of this study ran for a period of only five weeks. Many studies in educational sciences take the midpoint of courses as a critical measurement point to predict academic success and development of self-regulated learning skills. Using the original OSLQ-scale only – which uses a pre-post style measurements over a prolonged period of time – would have made it difficult to measure clear differences in the development of self-regulated skills during the short time span of the experiment. Instead, it was decided to construct an adaptation of the OSLQ-scale that diverts from measuring the development self-regulated learning skills in general during a course, but rather focuses on the extent to which the learning dashboard is perceived to support the development of self-regulated learning skills. 12 questions from the original scale were chosen that were categorized by the three phases of self-regulated learning, and adopted to a format that focuses on perceived support for certain self-regulated learning skills (α=.87). The questions consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I feel like this learning dashboard can help me with managing my time appropriately in this course.', in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E).

Usability [Survey 2, Survey 3] was measured using the System Usability Scale (SUS) (Brooke, 1995) and the End-User Computing Satisfaction-scale (EUCS) (Aggelidis & Chatzoglou, 2012). The SUS-scale was originally developed to measure the perceived usability of a website or application. This scale was chosen as it allows for the reformulation of questions so that they fit a certain context of use. The SUS-scale also covers many different aspects of usability, including comprehensiveness, learnability, memorability and ease-of-use. The 12 questions consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I found this learning dashboard unnecessarily complex.' in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). In addition, the EUCS-scale was chosen to strength the measurement of perceived user satisfaction, which is an essential part of usability but is not constructively measured in the SUS-scale. It was originally developed as a reliable instrument to measure end-user satisfaction in a web-based environment. The six consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'this learning dashboard provides me the information that I need.' in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). The set of adapted questions from the SUS-scale and the EUCS-scale together form a new scale of usability with a strengthened measurement of user-satisfaction (α=.84).

Trust [Survey 2, Survey 3] was measured using the scale developed by Corritore, Marble, Wiedenbeck, Kracher and Chandran (2005) for measuring online trust, credibility, and risk. The 8 questions consist of 7-point Likert scale questions where respondents are asked to rate how much they agree with statements such as 'I feel like there could be negative consequences from using this learning dashboard.' (α=.73), in which responses ranged from 'strongly disagree' to 'strongly agree' (Appendix E). The scale was chosen for this study as it was specifically developed by Corritore et al. (2005) as new instrument for MIS-researchers to study trust of web-based products.

Online activity Dashboard [continuously] was measured using clickstream data provided by the PowerBI webhosting service. Due to technical limitations and restricted access to sensitive user data within Power BI, it was only possible to retrieve the daily views per dashboard, but not per user. Therefore, the difference in daily views between the two dashboards was tested, but no further analyses could be conducted to examine how daily views relate to other individual measurements.

Open ended questions [Survey 3] relating to the participants' evaluations of their learning dashboard were included on the final page of Survey 3. The three questions included "How would you describe your experiences with this learning dashboard?", "based on the last couple of weeks within course, which features do you think were most valuable?", and "what did you feel was missing in the learning dashboard?" (Appendix E).

5.3 Privacy and Security

Access to the Canvas-data and dashboard-interaction data used in this study was strictly limited to the researcher and the technical team at TU/e that makes the data available. The researcher worked only with pseudonymized data in which all directly identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset or replaced by one or more artificial identifiers. To ensure the data was handled with care and integrity, the study was evaluated through a Data Protection Impact Assessment (DPIA). If students experienced negative consequences because of the feedback provided by the dashboard (e.g., stress or anxiety), they were encouraged to contact the researcher to resolve any unclarity or ambiguity with regards to the results that the dashboard presented and what these results meant. If they faced difficulties proceeding with their studies after participation, they were encouraged to contact their faculty's study-advisor. All research conducted at the Human-Technology Interaction Group adheres to the Code of Ethics of the NIP (Nederlands Instituut voor Psychologen – Dutch Institute for Psychologists). This study has been approved by the Ethical Review Board of Eindhoven University of Technology.

5.4 Participants

Before participant recruitment for the experiment started, an a-priori power analysis was performed using G*Power 3.1.9.7 to estimate the minimum sample size needed to ensure adequate power using a two-way repeated-measures ANOVA within-between interaction factors (Erdfelder, Faul & Buchner, 1996). Using an expected effect-size for the manipulation on cognitive learning outcomes of d=0.409; f=0.205 (Kim et al., 2015) results in a required sample size of N=54. Using an expected effect-size for the manipulation on affective learning outcomes d=0.456; f=0.228 (Fleur et al., 2020) results in a required sample size of N=44.

Participants were recruited through an invitation announcement on the Canvas-pages of the 0HV100 Human Factors and 0HV30 Social Psychology Consumer Behaviour courses at TU/e. They were required to speak English, do full-time studies in the BSc Psychology Technology-program during the first semester of the academic year 2022/2023, and be familiar with using the university's online learning platform Canvas. Participants could only participate in the study through one of two courses. Previous experience with learning dashboards was not required. 62 students registered for participation and were randomly allocated to either the SRLdashboard-group or the PRF-dashboard group after completing the first survey (Table 2). Participants were provided access to their learning dashboard once the second survey opened (Figure 5). 55 out of these 62 students completed all three surveys. All submissions were checked on validity by assessing whether answers to open-ended questions indicated a serious attempt, the total time spent on the survey was enough to be able to read and answer the questions meaningfully, and whether answers given to Likert scale questions were not unrealistically repetitive. No submissions had to be excluded from the analyses.

Table 2: 62 students from two courses registered for participation in the current study. The near equal distribution of participants across both courses in this experiment was coincidental and not intentionally planned.

	0HV100	0HV30	*Total*
SRLdashboard	15 (24%)	15 (24%)	*30 (48%)*
PRFdashboard	17 (28%)	15 (24%)	*32 (52%)*
Total	*32 (52%)*	*30 (48%)*	*62 (100%)*

5.5 Procedure

The study was run fully online and aimed to span 5 weeks in total. Students of the 0HV100 and 0HV30-courses during the 2022 autumn semester at TU/e were invited to participate in the study through the course's online Canvas-pages by completing a LimeSurvey with the registration form – including the information sheet and the informed consent form (Appendix D) - as well as the first out of three surveys. This survey closed after two weeks, after which all registered participants were randomly allocated to either the SRLdashboard-group or the PRFdashboard-group. Accordingly, participants were sent the link through which they could access their learning dashboard using their TU/e-account. The dashboard was accessible for a period of four weeks from the moment participants were granted access. The link to the second survey (i.e., accessible for one week) was sent over email one day after access to the learning dashboard was granted. The link to the third survey (i.e., also accessible for one week) was sent three weeks after participants first received access to the learning dashboard. The two-week gap between the second and third survey was due to the Christmas-break, where participants were not expected to engage with the dashboard. Even though participants were allowed and encouraged to use the dashboard whenever they wished during the four weeks that they had access, they were asked to interact with the dashboard at least once before every survey.

5.6 Data Preparation

Data preparation and analysis was performed using STATA IC 14.2 (*Stata Statistical Software: Release 14. College Station, TX: StataCorp LP*, 2015). The data from the three surveys was aggregated into one dataset where participants were identifiable by their TU/e-email. This dataset was pseudonymized by TU/e's IMS-team where student-emails were converted into pseudo-user-ids. This way, the survey data could be aggregated with the Canvas- and dashboard-clickstream data without the researcher knowing what data belonged to which student-email specifically. Canvas-data and dashboard-data were averaged per experimental phase (i.e., 'pre', 'start' and 'end'). Variables were transformed into long-format according to these phases so that every single pseudo-user-id had three rows of measurements.

Before the analyses were conducted, all measurements were checked for normality, outliers and homogeneity of variance. Due to the technical limitations mentioned earlier, it was only possible to retrieve the daily views per dashboard, but not per participant. In addition, views were often densely restricted to only four or five days and highly skewed towards the first two days of each experimental phase (i.e., because participants were asked to complete a survey around that time). As a result, normality and homogeneity of variances were rejected for daily views (Dashboard) but there were simply too few datapoints to be able to resolve any of them appropriately and meaningfully. To be able to deal properly with the rejected assumptions and the limited datapoints, a two-sample Wilcoxon rank-sum (Mann-Whitney) test was conducted to test for differences in dashboard views. Furthermore, a Shapiro-Wilk test rejected normality for motivation and daily minutes_online on Canvas. To resolve this, the logarithm was computed for each score of motivation and minutes_online to make them follow a normal distribution. The data was checked for outliers by standardizing the measurements and checking where the absolute value of each standardized measurement exceeded 3. One participant was detected with a total amount of online minutes that exceeded 14 hours on a single day. Further inspection of this participant's clickstream data revealed that all activity consisted of toggling one of the course's modules on Canvas. It was assumed that this was a glitch in the Canvas data recordings, and the participant was excluded for analyses on minutes_online. Homogeneity of variances was determined using Levene's test and was not rejected for any dependent variable.

5.7 Data Analysis

To test for differences in the dependent variables between the SRLdashboard-group and the PRFdashboard-group, several mixed-ANOVA analyses were conducted with dashboard-group as between-subject factor, time as within-subject factor, and each of the measurements as the repeated-measures dependent variable. Interactions were considered between the factors and further post-hoc analyses were conducted in case there was a significant effect. As differences in measurements between the two dashboard-groups might be affected by the course through which the participant registered in the study, course was included as an uncontrolled-for categorical covariate in all analyses. Two additional multi-level regression analyses were conducted to see if notable trends in measurements could somehow be explained by parts of the data other than the type of dashboard. In addition to previously tested assumptions, the variables were tested for multi-collinearity, homoscedasticity, and linearity for these analyses. Interaction effects were included in the model if predictor variables correlated too strongly with each other.

6. Results

6.1 Descriptive Statistics

Table 3 includes unstandardized descriptive statistics for the measurements conducted during the first phase of the experiment (i.e., before participants had access to their learning dashboard). Before receiving access to the learning dashboard, participants felt considerably motivated to pass their course (Table 3), although participants taking the 0HV30-course were significantly more motivated than those taking the 0HV100-course, $t(53) = -2.30$, $p = .012$. Participants also reported their sense of autonomy in their course to be considerably high. Metacognitive awareness was average for students from both courses. Nonetheless, as shown in Table 3 these 'pre-access' measurements give the impression that generally high levels of motivation and autonomy are reflected in high average grades at the beginning of the course, which on average even surpass the self-reported grade goal. Neither of these measurements were significantly different between both courses. The distribution plots of these measurements across the different phases of the experiment by dashboard-group can be found in Figure F.1 and Figure F.2 (Appendix F).

Table 3: Descriptive statistics of the measurements during the first phase of the experiment (i.e., before participants received access to their learning dashboard)

		Range	**M**	**SD**	**Var(X)**	**Skew(X)**	**Kurt(X)**
motivation	**0HV100**	2 - 6.3	4.71	1.15	1.31	-.67	3.08
	0HV30	2.8 - 6.1	5.11	.75	.57	-1.05	4.04
perceived autonomy	**0HV100**	2.7 - 5	3.74	.56	.31	.19	2.46
	0HV30	3.1 - 4.9	3.77	.47	.22	.24	2.54
metacognitive awareness	**0HV100**	1 - 4	2.51	.75	.56	.15	2.32
	0HV30	1.2 - 3.8	2.45	.69	.48	.34	2.53
average grade	**0HV100**	7.1 - 9.6	8.62	.62	.39	-.81	3.21
	0HV30	-	-	-	-	-	-
grade goal	**0HV100**	6 - 10	7.43	.84	.71	1.03	4.36
	0HV30	6 - 10	7.46	1.04	1.09	.65	3.41
minutes online (LMS)	**0HV100**	.5 - 52.2	8.85	11.57	145.7	2.15	7.4
	0HV30	.17 - 38.2	7.76	9.38	88.3	1.86	6.19

Note. No course grades were pusblished for 0HV30 during the first week of the experiment.

Table 4: Descriptive statistics of the average evaluations of SRLdashboard and PRFdashboard

		Range	**M**	**SD**	**Var(X)**	**Skew(X)**	**Kurt(X)**
support for SRL	**SRLdashboard**	2.08 - 4.58	3.36	.66	.43	.23	2.37
	PRFdashboard	2 - 5	3.29	.62	.38	-.39	2.42
perceived usefulness	**SRLdashboard**	2.2 - 4.8	3.55	.66	.44	.04	2.36
	PRFdashboard	2 - 5	3.48	.69	.48	-.03	2.37
usability	**SRLdashboard**	.63 - 4.18	2.78	.70	.49	-.51	3.79
	PRFdashboard	.72 - 4.36	2.84	.74	.54	-.38	3.01
trust	**SRLdashboard**	.88 - 3.38	2.41	.49	.25	-.61	3.64
	PRFdashboard	1.25 - 3.38	2.48	.53	.28	-.15	2.12

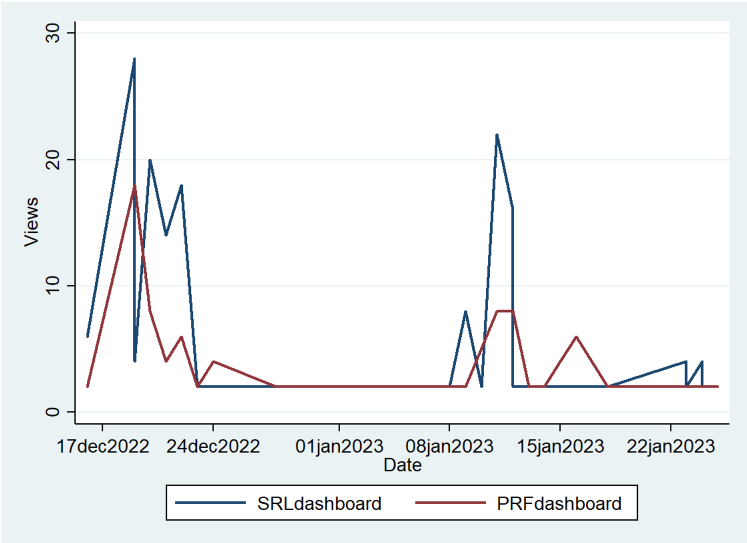


Figure 6: Daily views for SRLdashboard and PRFdashboard. The decline in activity in the middle of the graph can be attributed to the Christmas break, when students were less likely to engage with the dashboard.

After these initial 'pre-access' measurements were conducted, participants had access to their learning dashboard in the 'start'- and 'end'-phase. Figure 6 shows the total number of views per day for SRLdashboard and PRFdashboard within these two phases. There was usually a spike in views for both dashboards at the beginning of each phase. This was likely due to participants receiving access to a new survey in which they were asked to interact with the dashboard. Overall, average daily views were higher for SRLdashboard than for PRFdashboard in both phases. During the first week that participants had access to the dashboard, SRLdashboard had on average 14 views per day (SD=9.9) while PRFdashboard had on average 8 views per day (SD=6.2). After having access for three weeks, SRLdashboard had on average 8 views per day (SD=8.2) while PRFdashboard had on average 4 views per day (SD=2.9). The difference in daily views between SRLdashboard (n=11) and PRFdashboard (n=13) was not significantly different in the 'start'-phase or the 'end'-phase of the experiment (U=280.43, p= .35). Participants were asked to evaluate their learning dashboard twice during the second and third phase (i.e., 'start'- and 'end'-phase). Table 4 includes unstandardized descriptive statistics for the evaluations of SRLdashboard and PRFdashboard. Both dashboards were perceived to be above average supportive of self-regulated learning. In addition, both dashboards were perceived to be useful as well. However, usability and perceived trustworthiness were average for both dashboards.

Table 5: Measurements of all dependent variables by dashboard-group (i.e., *SRLdashboard*, *PRF-dashboard*, and averaged over both dashboards) across time (i.e., *pre*, *start*, *end*, and averaged over all phases).

		pre		**start**		**end**		*overall*	
		M	SD	M	SD	M	SD	M	SD
motivation	**SRLdashboard**	4.99	1.02	4.89	1.12	5.14	1.07	5.00	1.06
	PRFdashboard	4.83	.97	4.90	1.11	4.84	.92	4.85	1.00
	both dashboards	4.91	.99	4.89	1.10	4.98	1.00	4.93	1.03
perceived autonomy	**SRLdashboard**	3.69	.49	3.70	.47	3.70	.57	3.72	.50
	PRFdashboard	3.82	.53	3.90	.57	3.87	.54	3.86	.54
	both dashboards	3.76	.51	3.84	.52	3.79	.56	3.80	.53
metacognitive awareness	**SRLdashboard**	2.39	.71			2.42	.64	2.42	.67
	PRFdashboard	2.59	.72			2.41	.60	2.51	.67
	both dashboards	2.49	.72			2.42	.61	2.46	.67
average grade	**SRLdashboard**	8.67	.66			8.09	.90	8.30	.86
	PRFdashboard	8.62	.61			8.09	.95	8.28	.88
	both dashboards	8.65	.62			8.09	.92	8.29	.87
minutes online (LMS)	**SRLdashboard**	5.04	6.64	5.75	8.83	28.36	50.01	13.74	32.37
	PRFdashboard	11.06	12.79	4.93	9.54	21.87	26.66	13.00	1928
	both dashboards	8.34	10.81	5.28	9.13	24.86	38.87	13.33	25.86
support for SRL	**SRLdashboard**			3.46	.62	3.26	.70	3.37	.67
	PRFdashboard			3.31	.56	3.28	.70	3.30	.62
	both dashboards			3.39	.59	3.27	.70	3.33	.64
usability	**SRLdashboard**			3.08	.62	2.47	.65	2.78	.63
	PRFdashboard			3.12	.70	2.55	.66	2.84	.68
	both dashboards			3.10	.66	2.51	.65	2.81	65
perceived usefulness	**SRLdashboard**			3.47	.78	3.38	.83	3.43	.80
	PRFdashboard			3.28	.76	3.32	.83	3.30	.79
	both dashboards			3.37	.77	3.35	.82	3.36	.79
trust	**SRLdashboard**			2.38	.52	2.42	.48	2.40	.50
	PRFdashboard			2.44	.46	2.52	.60	2.48	.53
	both dashboards			2.42	.49	2.47	.55	2.44	.52
daily views (Dashboard)	**SRLdashboard**			14.3	9.91	7.7	8.2	10.8	9.3
	PRFdashboard			7.6	6.2	4.3	2.9	5.8	4.8
	both dashboards			11.3	8.8	6.2	6.3	8.5	7.8

Note. Measurements were taken at three different points in time: prior to participants receiving access to the learning dashboard (referred to as ”pre”), immediately after participants received access to the dashboard (referred to as ”start”), and three weeks after participants had access to the dashboard (referred to as ”end”). More details on the measurements and the experimental design are described in section 5.2.

6.2 Effect of Dashboard-design

6.2.1 Affective Learning Outcomes

Motivation was high for participants in both dashboard groups. The difference in motivation in the first week that participants had access was minimal although it slightly increased for the SRLdashboard-group while it slightly decreased for the PRFdashboard-group after that (Table 5). At the end of the experiment, motivation was higher for those that had access to SRL-dashboard than for those that had access to PRFdashboard. The mixed-ANOVA results indicate that neither the main effect of dashboard-group or time on motivation was statistically significant (Table 6). However, the interaction effect of dashboard-group and time was nearly significant with SSdashboardphase = 178.63, F(2, 109) = 2.69, p = .072, and 2 partial = .047. There is some evidence to suggest that the effect of the dashboard-group is different for different phases of the experiment and can explain some of the variance in motivation (4.7%), but it is not strong enough to be considered statistically significant, suggesting that further analyses and interpretation of this finding are needed to fully understand its theoretical implications (section 6.3). The covariate effect of the course on motivation was not statistically significant, nor were any of its interaction effects with the other variables (Table 6).

Perceived autonomy was high for participants in both dashboard groups, although it was on average higher for the PRFdashboard-group than for the SRLdashboard-group during all three phases of the experiment (Table 5). In addition to the PRFdashboard-group already reporting higher levels of perceived autonomy before they received access to the dashboard than the SRLdashboard-group, the levels of perceived autonomy for both groups hardly changed. Neither the main effect dashboard-group nor time alone - nor their interaction effect - could therefore explain a significant portion of the variance in perceived autonomy. The effect of the dashboard-group was not significantly different for any phase of the experiment. The covariate effect of the course through which the participant registered in the study and none of the interaction effects were statistically significant (Table F.1, Appendix F).

6.2.2 Cognitive Learning Outcomes

Metacognitive awareness was average for both dashboard groups (Table 5). There was only a minimal difference between the two dashboards after participants received access to their dashboard. Neither main effects of dashboard-group nor time - nor their interaction effects

Table 6: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and motivation as the repeated measured dependent variable.

				R-squared =	0.891
	Partial SS	**df**	**MS**	**F**	**p**
Model	13035.448	65	200.54535	13.72	0.0000
dashboard	81.117592	1	81.117592	0.41	0.5249
id|dashboard	11695.908	59	198.23573		
course	6.5033434	1	6.5033434	0.44	0.5062
phase	35.629553	2	17.814777	1.22	0.2996
phase#dashboard	78.639069	2	39.319534	2.69	0.2996
Residual	1593.2091	109	14.616597		
Total	14628.657	174	84.072742		

- were statistically significant (Table F.2, Appendix F). As such, metacognitive awareness was not significantly different for the dashboard-groups during different phases of the experiment, and therefore could not explain a signification portion of the variance alone. The covariate effect of the course on metacognitive awareness was statistically significant with SScourse = .767, F(1,53) = 5.11, p = .02 and 2 partial = .09. Apparently, at least some of the variance in metacognitive awareness (9%) is explained by the course through which the participant registered for the study. Post-hoc analyses show that at the end of the experiment, participants taking the 0HV30 Social Psychology Consumer Behaviour-course had slightly above average metacognitive awareness (M= 2.64, SD=.55) while participants taking the 0HV100 Human Factors-course had slightly below average metacognitive awareness (M=2.19, SD=.59), t(55) = -2.9, p = .001. This effect was not significantly different for the dashboard-groups, or during the first week of the experiment.

Average grade went down from the moment that participants received access to their learning dashboard (Table 5). The difference in this decrease was minimal between the SRLdashboard-group and the PRFdashboard-group. Neither the main effect of dashboard-group nor time - nor their interaction effect - were statistically significant. However, the covariate effect of the course through which the participant registered in the study was statistically significant with SScourse = 1.08, F(1, 24) = 22.43, p = ¡ .001 and 2 partial = .48. Apparently, the course in which the participant was enrolled significantly explained 48% of the variance in average grade. Further post-hoc analyses indeed show that the average grade for 0HV100 was higher (M=8.65, SD=.54) than for 0HV30 (M=7.27, SD=.71), t(75) = 9.3, p = ¡ .001. This effect was not significantly different for the dashboard-groups (Table F.3, Appendix F).

Minutes online per day in the LMS increased between the first week that participants had access to their learning dashboard and the last week (Table 5). The main effect of time

Table 7: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and minutes online (LMS) as as the repeated measured dependent variable.

				R-squared =	0.6514
	Partial SS	**df**	**MS**	**F**	**p**
Model	199.01856	58	3.4313544	2.35	0.0003
dashboard	1.4869941	1	1.4869941	0.61	0.4380
id|dashboard	116.69316	48	2.4311074		
phase	59.068873	2	29.534437	20.24	0.000
dashboard#phase	3.881458	2	1.940729	1.33	0.2707
course	5.7223882	1	5.7223882	3.92	0.0514
phase#courses	3.6811732	2	1.8405866	1.26	0.2893
Residual	106.49615	73	1.4588513		
Total	305.5147	131	2.3321733		

was statistically significant with SSphase = 73.71, F(2, 86) = 24.78, p ¡ .000 and 2 partial = .365. Apparently, the phase of the experiment explained a significant 37% of the variance in daily minutes online in the LMS. Further post-hoc analyses indicate that the daily minutes online was higher at the end of the experiment, t(99) = -3.71 and p ¡ .001. This increase was higher for participants that had access to SRLdashboard than for those that had access to PRFdashboard, although it was not statistically significant. The covariate effect of the course through which the participant registered in the study was nearly significant with SScourse = 5.72, F(1, 83) = 3.76, p = .056 and 2 partial = .043. Further post-hoc analyses indicate that daily minutes online in the last week of the experiment was higher for 0HV100-students (M=28.30, SD=47.49) than for 0HV30-students (M=19.27, SD=29.34), t(47) = 2.32 and p = .012. Although the p-value reached statistical significance, the portion of the variance in minutes online (LMS) that course could explain was very small (4.3%). None of the interaction effects of course with phase or dashboard-group were statistically significant (Table 7).

6.2.3 Dashboard Evaluations

Perceived support for self-regulated learning was considerably high for both dashboard groups, although in the first week that participants had access it was higher for SRLdashboard than for PRFdashboard (Table 5). The difference between the two dashboards was only minimal towards the end of the experiment. Neither the main effect dashboard-group nor time alone - nor their interaction effect - could explain a significant portion of the variance in perceived autonomy. Therefore, the effect of the dashboard-group was not significantly different during the different phases of the experiment. The covariate effect of the course in which the participant

Table 8 Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and usability as as the repeated measured dependent variable.

				R-squared =	0.8678
	Partial SS	**df**	**MS**	**F**	**p**
Model	50.087723	59	.84894446	5.90	0.0000
dashboard	.25006081	1	.25006081	0.35	0.5582
id|dashboard	39.636135	55	.72065701		
course	.52892556	1	.52892556	3.67	0.0606
phase	8.8702207	1	8.8702207	61.63	0.0000
phase#dashboard	.11259445	1	.11259445	0.78	0.3804
Residual	7.628653	53	.14393685		
Total	57.716376	112	.51532479		

was enrolled was nearly statistically significant with SScourse = 1.74, F(1, 51) = 3.31, p = .074 and 2 partial = .06. Post-hoc analyses show that on average perceived support for self-regulated learning was higher for 0HV100 Human Factors (M=3.36, SD=.62) than for 0HV30 Social Psychology Consumer Behaviour (M=3.08, SD=.91). This effect was not significantly different for the dashboard-groups or phases of the experiment (Table F.4, Appendix F). Although the p-value did not reach statistical significance, the effect size was small to moderate, suggesting that further analysis and interpretation of this finding are needed to fully understand its theoretical implications (section 6.3).

Perceived usefulness was considerably high for both dashboards, although it was slightly higher for SRLdashboard than it was for PRFdashboard (Table 5). These differences in perceived usefulness became smaller towards the end of the experiment. Trust in both dashboards was average, and there seemed to be no difference between the two dashboards (Table 5). In general, trust in the learning dashboard – regardless of dashboard design – went slightly up from the moment that participants first received access until they had access to it for three weeks. Neither the main effect dashboard-group nor time alone - nor their interaction effect - could explain a significant portion of the variance in perceived usefulness or trustworthiness. The effect of the dashboard-group was not significantly different during the different phases of the experiment. The covariate effect of the course in which the participant was enrolled was nearly significantly different for perceived usefulness with SScourse = .081, F(1, 53) = 2.72, p = .091 and 2 partial = .08. Considering that at least some part of the variance in perceived usefulness (8%) can possibly be explained by the course, further post-hoc analyses were conducted and showed that participants in the 0HV100-course perceived SRLdashboard to be significantly more useful (M=3.65, SD=.67) than PRFdashboard (M= 3.33, SD= .65), t(53) = 1.69, p=.04. This difference was not observed in participants enrolled in the 0HV30-course (Table F.5, Appendix F). Further analysis and interpretation of this finding are needed to fully understand its theoretical implications

Usability was above average for both SRLdashboard and PRFdashboard although it decreased significantly from the moment that students first received access to their learning dashboard (Table 5) with SSphase = 8.87, F=(1,53) = 61.63, p ¡ .001 and 2 partial = .53 (Table 8). This decrease was not significantly different between both dashboard groups. The covariate effect of the course in which the participant was enrolled was nearly statistically significant with SScourse = .404, F(1, 53) = 2.81, p = .091 and 2 partial = .09. There is some evidence to suggest that the phase of the experiment and the course can explain some of the variance in usability, suggesting that further analysis are needed (section 6.3). Post-hoc analyses show that participants enrolled in the 0HV100-course evaluated SRLdashboard to be significantly more usable (M=.53, SD=.66) than PRFdashboard (M=.21, SD=.57), t(56) = 1.91, p = .03. This difference was not observed for participants enrolled in the 0HV30-course.

6.3 Predicting a Learner's Motivation

Some interesting patterns were observed in motivation that suggested an unexplained effect. Even though motivation seemed largely the same for both dashboard-groups at the beginning of the experiment, it increased for the SRLdashboard-group while it slightly decreased for the PRFdashboard-group (Table 5). In addition, levels of motivation were significantly different in both courses. A multi-level regression analysis was conducted to see if these observations can somehow be explained by other parts of the data than the dashboard-group or phase of the experiment alone. Two empty variance component models to predict motivation were run where id and course were considered as clusters separately. The variance explained by both clusters was significantly different from zero. Therefore, a multi-level regression analysis was conducted where individual-level and course-level variances were partialled out (i.e., with repeated measures nested in id's and id's clustered in courses). Whenever minutes online in the LMS was included in the model, the coefficients and significance of other predictors changed drastically. However, there was no significant increase in model fit by including it as a predictor (i.e., as shown by no difference in conditional intraclass correlations). Minutes online in the LMS was therefore excluded from the model as it was believed to capture too much noise in the data and therefore overfitting the model. All other standardized measurements were included as predictors. No interaction effects were included in the model as there was no multicollinearity amongst the predictors. Trust was

Table 9: Results of the multi-level regression analysis predicting motivation.

Model (motivation)	β	***SE***	***z-value***	***p***	***95% CI***
metacognitive awareness	-.02	.141	.01	.991	-.273, .276
perceived autonomy	.568 *	.129	4.40	.001	.315, .821
support for SRL	.236 *	.112	2.11	.035	.315, .821
usefulness	.343 *	.139	2.47	.014	.071, .615
usability	.001	.121	.01	.994	-.237, 239
trust	-.148	.126	-1.17	.241	-.396, .0994
dashboard; "PRFdashboard"	-.578 *	.167	-3.47	.001	-.904, -.251
constant	.406 *	.206	1.97	.049	.001, .809

Random-effects parameters	***Estimate***	***SE***	***95% CI***
course: Independent			
var(constant)	.059	.075	.047, .732
id: Independent			
var(trust)	.213	.179	.041, 1.10
var(constant)	.162	27.51	7e-14, 3e+14
var(Residual)	.066	27.50	0, .

Note. Predictors are measurements clustered in id's, clustered in courses. All predictors are standardized for interpretation purposes and random slopes are included for trust. The course-cluster explains 12% of the variance in motivation; the total model explains 41% of the variance in motivation.

found to have random slopes at the individual level.

The model predicted a statistically significant portion of the variance in motivation compared to a null model with no predictors (likelihood ratio test: chi2(7) = 57.92, p ¡ .001). Interestingly, the strongest predictor of motivation is dashboard-group (i.e., SRLdashboard or PRFdashboard). This suggests that dashboard-group was not a significant predictor of motivation on its own, but its effect is rather moderated by the other predictors. Participants in the PRFdashboard-group are predicted to be significantly less motivated than those in the SRLdashboard-group only if all other variables are included in the model. In addition, a participant's perceived autonomy as well as the dashboard's perceived usefulness and perceived support for self-regulated learning were significant and positive predictors of motivation. The usability and trustworthiness of the dashboard were not significant predictors, nor was the participant's metacognitive awareness. In summary, motivation was generally higher whenever participants reported higher levels of perceived autonomy, and evaluated the dashboard to be more useful and to provide more support for self-regulated learning. In addition – only when all these measurements are included in the model – motivation was predicted to be higher for the SRLdashboard-group than for the PRFdashboard-group.

Table 10: Results of the multi-level regression analysis predicting a dashboard's perceived usefulness.

Model (usefulness)	*β*	***SE***	***z-value***	***p***	***95% CI***
motivation	.27 *	.111	2.47	.013	-.057, .496
metacognitive awareness	.011	.134	.08	.937	-.252, .274
perceived autonomy	-.174	.126	-1.39	.166	-.421, .072
support for SRL	.609 *	.079	7.69	.001	.453, .764
usability	.301	.099	3.04	.002	.107, .496
trust	.046	.101	0.46	.642	-.149, .242
course; "0HV30"	-.0737	.181	-0.41	.684	-.428, -.281
dashboard; "PRFdashboard"	-.102	.162	.63	.528	-.216, .422
constant	-.023	.143	-.17	.868	-.305, .257

Random-effects parameters	***Estimate***	***SE***	***95% CI***
id: Independent			
var(constant)	.259	14.19	7e-14, 1e+46
var(Residual)	.037	14.19	0, .

Note. Predictors are measurements clustered in id's. All predictors are standardized for interpretation purposes; no evidence was found for random slopes. The fixed effect accounts for 78% of the variance in perceived usefulness; random effects and residual variance account for 22% of the variance in perceived usefulness.

6.4 Understanding the Dashboard's Perceived Usefulness

Some interesting patterns were observed amongst the dashboard evaluations that suggested an unexplained effect. Perceived usefulness of SRLdashboard was above average - and in fact significantly higher for participants enrolled in the 0HV100-course -while for PRFdashboard it was below average. In addition, the dashboard's perceived support for self-regulated learning was higher for SRLdashboard than for PRFdashboard, where this difference seemed to be partly explained by the course through which the participant registered. The level of trust in both dashboards increased equally from the moment that participants received access until the end of the experiment. A multi-level regression analysis was conducted to see if these observable patterns in the dashboard's evaluations somehow relate and can be explained by other parts of the data than the dashboard-group alone. Two empty variance component models to predict perceived usefulness were run where id and course were considered as clusters separately. The variance explained by id was significantly different from zero; for the course it was not. Therefore, a multi-level regression analysis was conducted where individual-level variances were partialled out (i.e., with repeated measures nested in id's) and the dashboard type and the course were included as categorical predictors. Minutes online in the LMS was again excluded as it was believed to overfit the model. No interaction effects were included in the model as there was no multicollinearity amongst the predictors. Metacognitive awareness was found to have random slopes at the individual level.

The model predicted a statistically significant portion of the variance in motivation compared to a null model with no predictors (likelihood ratio test: chi2(8) = 124.65, p ¡ .001). Perceived support for self-regulated learning was the strongest predictor of perceived usefulness, followed by the dashboard's usability. In contrast to the previous multi-level model, neither the dashboard nor the course are significant predictors of perceived usefulness. A participant's metacognitive awareness, motivation or perceived autonomy was not a significant predictor, nor was their trust in the dashboard. In summary, the dashboard's perceived usefulness was generally higher whenever the dashboard was perceived as more usable and supportive for self-regulated learning.

6.5 Qualitative Findings

This section will briefly discuss some of the most notable findings from the open-ended questions at the end of the third survey (Appendix E). The open-ended questions asked participants to provide more detailed feedback on their experiences with SRLdashboard and PRFdashboard, and their responses provided valuable insights into the strengths and weaknesses of each dashboard.

6.5.1 SRLdashboard: Balancing Usefulness and Overwhelming Information

Participants found SRLdashboard to be comprehensive and well-designed. One participant said, "the interface is clean and visually pleasing. I was given plenty of information, but the layers of individual information were also well represented and arranged". One participant said, "It's organized and after the short explanations easy to understand." Another remarked, "The information presented to the user is a lot but at the same time the whole structure is clear and does not make people feel confused." However, it still seems that SRLdashboard is a mixed bag in terms of user experience. While some participants found the dashboard to be clear and comprehensive, others felt that it was too chaotic and overwhelming. They appreciated the large amount of information presented, but some felt that not all of it was useful. One participant remarked that "there [was] way too much information on the screen at once," and that it was "hard to find what you are looking for." One participant commented "It feels a bit messy, but it was well-made." Despite these criticisms, most participants still found the dashboard easy to use and visually pleasing. Several participants noted positive aspects of the dashboard, such as the "you might want to check out" feature for files and the visualization of the agenda with colors, dates, and deadlines. One participant said, "I like that it's just one page and it's easy to see everything I need to do for the course". Another participant said, "it's a helpful tool to have, especially if you're not great at planning and need some extra help".

6.5.2 PRFdashboard: Compact and Visually Pleasing with Limited Usefulness

In contrast, participants found PRFdashboard to be more compact and easier to navigate. They appreciated the clear visualization of grades and progress, with one participant saying, "I like the gradebook and progress pie charts." Another noted, "It looks quite good and there is not too much information expressed at once." One participant said, "I like the layout, it's very clear and easy to use". However, one participant felt that the dashboard was "not the best". Overall, the feedback for PRFdashboard was generally positive. However, some participants suggested improvements such as a tool to help them achieve their target score and a way to keep track of learning progress and testing knowledge.

6.5.3 Suggested Improvements for A New Canvas Dashboard

For both SRLdashboard and PRFdashboard, participants commented on the familiarity of the dashboard's layout, which resembled Canvas. They suggested having a clear page that directs them to the modules, since they used them most on Canvas. Another participant suggested having a part in the calendar that shows what materials are coming with every week/lecture. Participants had mixed opinions on the usefulness of certain dashboard features. For example, while some found data about the time spent on Canvas to be irrelevant, others appreciated being able to see their online activity and upcoming deadlines. One participant even suggested that the time of deadline graph could be even more helpful if it were integrated with all of the their courses in one graph. Additionally, while some participants liked the comparison with other participants based on grade goal, others didn't see how this feature could help them improve their own learning. One participant even remarked that the calendar helped them know the exact time of deadlines and the schedule related to the course. Some participants also commented on the need for more customization and personalization options.

Overall, the qualitative findings provide valuable feedback for improving both dashboards and making them more useful for students. Participants did not recommend removing functionalities for either SRLdashboard or PRFdashboard, but rather optimizing what already exists and adding new ones, like a more interactive design and more useful commands. Some participants suggested that they would like to have more homework and interactive quizzes to test their knowledge. When it came to improvements, some participants suggested having recent updated files of the course, while another wanted to see the average grade for the assignments. One participant suggested having a part of the dashboard that shows how many points they need to get on their next assignment to achieve their target score.

7. Discussion

This study focused on how the design and evaluation of student-facing learning dashboards can be grounded in self-regulated learning theory, and to what extent the (in)consideration of the theory's phases yields different cognitive and affective learning outcomes. In an experimental study, differences in dashboard evaluations and cognitive and affective learning outcomes were examined between a group of students that had access to a learning dashboard that supported all three phases of self-regulated learning (SRLdashboard) and a group of students that had access to a learning dashboard that only supported the performance phase (PRFdashboard). The three research questions were: ***RQ1:*** *'How is a learning dashboard that supports all three phases of self-regulated learning evaluated differently compared to a learning dashboard that only supports the performance phase?'*, ***RQ2:*** *'What cognitive learning outcomes are impacted differently when having access to a learning dashboard providing support for all three phases of self-regulated learning compared to a learning dashboard only supporting the performance phase?'*, and ***RQ3:*** *'What affective learning outcomes are impacted differently when having access to a learning dashboard providing support for all three phases of self-regulated learning compared to a learning dashboard only supporting the performance phase?'*.

It was expected that differences in cognitive and affective learning outcomes would be observed between participants that had access to SRLdashboard and participants that had access to PRFdashboard. However, the findings suggest that the dashboard that participants had access to was not enough to explain observed differences in learning outcomes alone. For one, the difference between the two learning dashboards could have been simply too small to be able to detect them with the measurements used. Nonetheless, the data that was obtained from the experimental study speak of an interesting story as some trends in the measurements were observed that were explainable by parts of the data other than the dashboard group. The multi-level regression analyses that were conducted provided some interesting discoveries.

7.1 General Findings

7.1.1 Increase Dashboard Usefulness through More SRL-Support

The findings suggested a relationship between how much the learning dashboards supported self-regulated learning and how useful they were perceived to be. The perceived usefulness of SRLdashboard was above average while for PRFdashboard it was below average, although the differences between the two dashboards could not explain the differences in perceived usefulness alone. However, participants enrolled in the 0HV100 course found SRLdashboard to be significantly more useful than PRFdashboard, and similarly, also found SRLdashboard to provide more support for self-regulated learning. These findings do not seem to be coincidental. Whereas the Canvas-page for 0HV30 was only used to post slides and assignments, the Canvas-page for 0HV100 contained a lot more content (e.g., instructional videos, additional literature). As a result, participants that were enrolled in 0HV100 had access to a dashboard that was a lot more dynamic. Likely, the information that SRLdashboard presented was also very different between students within that course because it was based on a lot more course content, and therefore varied much more. As such, the data that was visualized in the performance phase and self-reflection phase functionality (i.e., recommendations) was a lot more specific to a specific students (e.g., what instructional videos on YouTube other students accessed more). For students enrolled in 0HV30, it is likely that the difference in the dashboards between students was much smaller. Since only slides and assignments were posted to the 0HV30 course page, recommendations included what lecture slides other students accessed more. As a result, a dashboard's perceived ability to support in the self-reflection phase of self-regulated learning – regardless of the dashboard's design – was at least to some degree affected by the course content onto which it builds. These findings served as an interesting start for further analyses, which revealed that perceived support for self-regulated learning was the strongest predictor of perceived usefulness, followed by the dashboard's usability. Neither the dashboard nor the course were significant predictors of perceived usefulness in this analysis, suggesting that the differences that were observed in perceived usefulness between the two courses could be attributed to differences in perceived support for self-regulated learning. The first hypothesis to research question one was the following:

H1.1: *Perceived usefulness will be higher for a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the performance phase.*

Since the dashboard-group was not a significant predictor of perceived usefulness - also not after including all the other predictors - the current study could not provide enough support for H1.1. However, the results highlight the importance of designing learning dashboards that are both usable and supportive for self-regulated learning, as these factors were found to be strong predictors of perceived usefulness and motivation. As discussed in the literature, the dashboards developed by Santos et al. (2013) and Loboda et al. (2014) were found to be more effective in supporting self-regulated learning than the dashboards developed by Kim et al. (2015) and Corrin and Barba (2014) as they provided more information on course material and tasks, as well as the learner's study tactics and goals. These findings suggest that dashboards that are more closely aligned with course material and provide more assistance in goal-setting and task planning may lead to greater user satisfaction. Differences in the dashboard's perceived usefulness between the two courses in this study could be largely attributed to differences in perceived support for self-regulated learning, which also was different for both courses. Participants in the interview study already expressed concerns about the extent to which the learning dashboard would be able to assist them in goal-setting and task-planning, as they were questioning its ability to consider differences between the organization and goals of different courses. In order to provide actionable feedback on how to increase performance, a learning dashboard must properly embed the course's learning objectives in the presentation of performance data, so that situations where students deviate from the course structure can be appropriately considered without being misinterpreted as a decline in learning. These findings are in line with Winne and Hadwin (1998) which emphasizes the importance of incorporating external task conditions in learning dashboard design.

The experimental findings on differences in perceived usefulness align with responses to the open-ended questions, particularly relating to certain features. While some participants found data about the time spent on Canvas to be irrelevant, others appreciated being able to see their calendar and upcoming deadlines (i.e., only available to SRLdashboard). One participant even suggested that the time of deadline graph could be even more helpful if it were integrated with all of the their courses in one graph. When it comes to the performance phase features, the visualizations that were most appreciated were usually the ones that related to the course material specifically; the Canvas-indicators (i.e., clickstream data from the LMS) were rarely mentioned as a valuable indicator of performance. As (Jivet et al., 2018) argued, it is unclear to what extent this 'observable' behaviour accurately represents 'learning' behaviour. While previous research has shown these types of indicators to be most commonly used in learning dashboards (Matcha, Uzir et al., 2020), studies have shown that presenting quantitative metrics like these rarely offer any actionable guidance to a learner's study strategies (Kovanovic et al., 2016). This was reflected in more appreciation for features that reflected the course material, and not just clickstream data.

It was expected that by incorporating all three phases of self-regulated learning, a learning dashboard can enhance its trustworthiness in several ways. To begin with, it can offer a more comprehensive perspective of the learning process, thus enhancing the transparency and credibility of the dashboard. This is because by providing information on how learners plan their learning, monitor their progress, and reflect on their learning, the dashboard can present a more complete portrayal of the learner's exertions and advancements. However, the data did not provide support for the second hypothesis to research question one:

***H1.2**: Perceived trust will be higher for a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the performance phase.*

7.1.2 Exploring the Relationship between Dashboard Design and Affective Learning

Motivation increased for the SRLdashboard group but slightly decreased for the PRF-dashboard group as soon as participants received access to their learning dashboard. Although both dashboard groups had high motivation before receiving access to their learning dashboard, the marginal impact of the dashboard on motivation was not clear. It is possible that the differences in the dashboard design may not have been strong enough to significantly influence motivation levels. However, the results indicated that the effect of the dashboard on motivation could be moderated by other predictors. Only when all specific dashboard evaluation measurements were taken in to account, was the SRLdashboard-group significantly more motivated than the PRFdashboard-group. Overall, motivation was generally higher when learners also reported higher levels of perceived autonomy, and evaluated the dashboard to be more useful and to provide more support for self-regulated learning. One of the hypotheses to research question 3 was:

***H3.1**: Motivation will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

The data does not provide sufficient evidence to support this hypothesis. However, a caveat should be noted here. While the differences in implementation of self-regulated learning was not strong enough of an effect to yield differences in motivation, more perceived support for self-regulated learning in combination with a higher perceived usefulness did have significant effect on motivation. This suggests that in order to increase motivation, learning dashboard should aim to incorporate more support for all phases of self-regulated learning. The extent to which differences in motivation were truly because of the different degrees of support for self-regulated learning, is yet still debatable. The concerns that were addressed in the literature review still hold up. Kim et al. (2015) discovered that the effect of a learning dashboard on student motivation varied depending on their academic achievement level. Low-achievers generally experienced a greater increase in motivation from comparing themselves to others, while high-achievers were already highly motivated and thus showed a lower increase in motivation. Students who fell into these extremes tended to prefer a learning dashboard that contextualized their performance in a more self-referenced way, such as in comparison to their own previous performance. Fleur et al. (2020) demonstrated that accounting for differences in goal orientation reduces the fluctuations in motivation in learning dashboards that use social-comparison reference frames, and can therefore ensure a more stabilized increase in motivation. While the current dashboard attempted to account for these fluctuations in motivation through a target grade slider, it is uncertain whether this was the sole reason for the differences in motivation that were observed, as differences in goal orientation were not measured in the study. Even though motivation was higher for SRLdashboard than for PRFdashboard if also accounting for differences in the dashboard's evaluation, it is difficult to come to clear conclusions on the extent to which this increase in motivation was truly because of the dashboard's implementation of self-regulated learning.

How much the participant felt in control during their course was a significant predictor of motivation. Although the PRFdashboard-group had higher levels of perceived autonomy than the SRLdashboard-group throughout the experiment, the difference in autonomy could not be attributed to the dashboard. The PRFdashboard-group already felt more autonomous before they had access to the learning dashboard, and therefore the difference could not be explained by the dashboard, nor having access at all (i.e., perceived autonomy hardly changed during the experiment). Therefore, the second hypothesis to research question 3 was not supported:

H3.2: *Perceived autonomy will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

7.1.3 Impact of Learning Dashboards on Cognitive Learning

It was expected that SRLdashboard would result in higher metacognitive awareness than PRFdashboard through incorporating features that provide a comprehensive and integrated view of the learning process, which can improve the user's understanding and awareness of their learning progress. For instance, including features that enable planning and goal-setting was expected to assist learners in comprehending what they need to do to achieve their learning objectives. Similarly, features that facilitate monitoring and reflection were expected to help learners recognize areas for improvement and adjust their learning approaches accordingly. This reasoning was not reflected in the data, so that the first hypothesis to research question two could not be supported. The dashboard supporting only the performance phase may have included some features that indirectly supported metacognitive awareness, which could have contributed to the lack of difference between the two dashboards. This is just speculative, however, and further examination of varying effects on metacognitive awareness is needed.

H2.1*: Metacognitive awareness will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for a learning dashboard that only supports the performance phase.*

It was expected that SRLdashboard would yield higher average course grades than PRF-dashboard. Previous studies suggest that goal setting and monitoring play a crucial role in self-regulated learning and are linked to better learning outcomes (Jivet et al., 2018). Feedback and support for the self-reflection phase can improve study strategy adjustments, which can optimize learning and performance (Bodily & Verbert, 2017). Based on these findings, it was believed that a learning dashboard that supports all three phases of self-regulated learning would lead to higher course grades compared to a dashboard that only supports the performance phase. However, the data did not support this hypothesis. The results showed that participants' average grades decreased once they received access to their learning dashboard, but this decrease was similar between the two dashboard groups. The last hypothesis could therefore not be suppported:

H2.2*: Course grades will be higher for students that have access to a learning dashboard that supports all three phases of self-regulated learning than for students that have access to a learning dashboard that only supports the performance phase.*

7.2 Limitations

One of the main limitations of the current study was the time constraint under which the final experiment was conducted, which restricted the way in which the development of self-regulated learning skills could be measured. Ideally, self-regulated learning skills are measured continuously over a prolonged period of time. Many studies in educational sciences take the midpoint of courses as a critical measurement point to predict academic success and development of self-regulated learning skills. Using any pre-validated scale – which commonly uses pre-post style measurements spaced out across a larger time span - would have made it difficult to measure clear differences in the development of self-regulated skills during the short time span of the experiment (i.e., five weeks). Instead, an adaptation of the OSLQ and MSLQ-scale was developed specifically for the current study. Naturally, even though the alpha value of the newly developed scale was considerably high, it is open to interpretation whether the scale as a whole measured exactly what it was supposed to measure.

Another limitation due to the time constraint was a restriction in measuring differences in cognitive learning outcomes. The experiment – including access to the learning dashboard – started halfway during the two courses. For the 0HV30-course, this meant that no grades or assignments were published when the experiment started. This negatively impacted two parts of the experiment. For one, the dashboard that 0HV30-students had access to was populated with a lot less data (e.g., the gradebook, assignments to be completed). This might have affected the results, such as the difference in perceived usefulness between the two courses. Secondly, this also meant that the average grade measurements in the 'pre'- and 'start'-phase of the experiment were based only on students in the 0HV100-course, which dramatically reduced the amount of datapoints for those analyses. In addition, grade changes were published during the 'start'-phase of the experiment, meaning that the final analysis was conducted only on the 'pre'-phase and the 'end'-phase of the experiment.

In addition, the resources to analyse interaction with the dashboard were restricted. Interaction with the dashboard could only be measured using clickstream data provided by the PowerBI webhosting service. Due to technical limitations and restricted access to sensitive data within PowerBI, it was only possible to retrieve the daily views per dashboard, but not per user. Therefore, the difference in daily views between the two dashboards was tested, but no further analyses could be conducted to examine how daily views relate to other individual measurements. Being able to 'connect' user-specific clickstream data (e.g., how they navigated through the dashboard, where they clicked, how long they were online) with the rest of the measurements could have provided some valuable insights. It would have been interesting to see if any of the effects that were found in the multi-level regression analyses could be explained by the clickstream-data of the dashboards. This would have opened up for more interesting analyses; do participants that perceive the dashboard to be more supportive of self-regulated learning also use it more frequently?

Is more clickstream activity in the SRLdashboard good indicator of more perceived usefulness? Was a sudden decrease in clickstream activity also an indicator of a decrease in motivation? In addition, the researcher worked with sensitive data and the experimental study went through an extensive DPIA. Practically, the learning dashboards that the were developed were easily scalable to other courses at TU/e, which would have increased the number of participants and therefore the study's statistical power.

7.3 Implications and Future Research

The discussion highlighted the importance of designing learning dashboards that are both usable and supportive for self-regulated learning. The findings suggest that dashboards that provide more information on course material and tasks, as well as the learner's study tactics and goals, may lead to greater user satisfaction. However, the study could not provide support for the hypothesis that perceived usefulness is higher for a learning dashboard that supports all three phases of self-regulated learning than for a dashboard that only supports the performance phase. Further research should therefore focus on properly embedding the course's learning objectives in the presentation of performance data. This is needed to explore the relationship between the course content and the effectiveness of learning dashboards. This could be done by comparing the effectiveness of dashboards across different courses with varying levels of course content available.

Similar to Santos et al. (2013) and Loboda, this study followed a theory-driven, user-centered design approach in which the results from the interview -were used as guidance for constructing design features that support each phase of self-regulated learning theory in a way that is appropriate to the given context of use. Positive findings in perceived usefulness and support for self-regulated learning seem to suggest that taking this approach is more effective than a data-driven design approach. The qualitative evaluations reflect these findings, as the most appreciated features were those that directly related to the course material. On the other hand, clickstream indicators were not frequently mentioned as a useful performance indicator. These results are consistent with previous research, which suggests that it is unclear whether observable behavior accurately represents learning behavior. While previous research has shown that these types of indicators are commonly used in learning dashboards, other studies have indicated that presenting quantitative metrics like these does not provide practical guidance for a learner's study strategies. Therefore, future dashboard designs should evaluate how visualizations could align with the course material more.

The results underscore the importance of taking a more holistic approach to designing learning dashboards, one that accounts for a range of factors that influence learning outcomes. For instance, dashboard designers may need to focus on providing more personalized support to learners, through features that allow learners to set personalized goals, track progress, and receive tailored feedback. They may also need to focus on enhancing learner autonomy, by giving them more control over their learning experiences and providing them with opportunities to take ownership of their learning goals and progress.

The discussion highlighted the importance of designing learning dashboards that are both usable and supportive for self-regulated learning. The findings suggest that dashboards that provide more information on course material and tasks, as well as the learner's study tactics and goals, may lead to greater user satisfaction. However, the study could not provide support for the hypothesis that perceived usefulness is higher for a learning dashboard that supports all three phases of self-regulated learning than for a dashboard that only supports the performance phase. Further research should therefore focus on properly embedding the course's learning objectives in the presentation of performance data. This is needed to explore the relationship between the course content and the effectiveness of learning dashboards. This could be done by comparing the effectiveness of dashboards across different courses with varying levels of course content available.

Previous learning dashboard studies have rarely used AB-testing in their experiments (Matcha, Uzir et al., 2020). Rather the dashboard is evaluated using pre-access and post-access measurements only, leaving open the question of different design elements contribute to differences in learning outcomes. Additionally, given the heterogeneity of students across different dashboard studies, it is difficult to compare the effectiveness of different implementations of self-regulated learning. The study highlights the importance of comparing different designs with each other, rather than simply assessing the impact before and after providing access to a particular design. Studies including those by Kim et al. (2015), Fleur et al. (2020), Corrin Barba (2014), and Tan et al. (2016), have reported positive effects on course grades after implementing self-regulated learning principles in their learning dashboards. However, it is important to question whether these effects are truly due to differences in implementation of self-regulated learning theory, as many other dashboard studies also report an increase in course grades without measuring differences in self-regulated learning effectively. This raises concerns about whether the positive effects on grades are actually due to improved self-regulation or to other factors, such as higher motivation, as suggested by Kim et al. (2015). AB-testing can help isolate the effects of specific design features, such as supporting all three phases of self-regulated learning, from other potential confounding variables.

Many participants indicated in the open-ended questions that even though SRLdashboard was perceived to be more useful and more supportive of learning, the information it presented was also a lot more overwhelming and confusing. In future research, it is important to consider cognitive load as an essential component when evaluating the effectiveness of learning dashboards. Although learning dashboards provide a rich source of information about students' progress, they can also overwhelm learners with too much data, leading to cognitive overload and reduced learning outcomes. Therefore, measuring cognitive load could provide valuable insights into the design of effective learning dashboards that balance information overload with learners' ability to process information efficiently. By incorporating cognitive load measures into future evaluations, we can gain a deeper understanding of how learners interact with learning dashboards and develop more effective designs that enhance learning outcomes. Ultimately, this could lead to the development of personalized learning dashboards that adapt to learners' cognitive capabilities, improving their overall learning experiences. In addition, the current study was restricted in the way that self-regulated learning could be measured. Suggestions for future research are conducting a longitudinal study to investigate the long-term impact of learning dashboards on student learning outcomes, where the development of self-regulated learning skills are measured over a prolonged period of time tat students have access to a learning dashboard. This could involve measuring student performance before and after the implementation of a learning dashboard and tracking changes in their learning behaviors and strategies.

7.4 Conclusion

This study focused on how the design and evaluation of student-facing learning dashboards could be grounded in self-regulated learning theory, and to what extent the (in)consideration of the theory's phases would yield different cognitive and affective learning outcomes. While the difference in the implementation of self-regulated learning could not explain observable patterns in the data, the results have contributed to a wider understanding of how dashboards should assist learners online. The data told of an interesting story, in which a dashboard's perceived ability to support self-reflection in self-regulated learning was affected by the course content onto which it was built, which greatly affected perceived usefulness. Together with a learner's perceived autonomy, these variables were found to be significant predictors of a learner's motivation. For future research, it is recommended to take a more holistic approach to designing learning dashboards by accounting for a range of indicators that influence learning outcomes. Design directions include providing more personalized support to learners, enhancing learner autonomy and aligning visualizations with the course material. The continuous advancements in learning dashboard research open up a fruitful ground for new understandings on how optimize students' learning processes more effectively. This study hoped to contribute to these understandings by providing design recommendations that are better aligned with students' needs and preferences for self-regulated learning assistance.

References

Aggelidis, V. P. & Chatzoglou, P. D. (2012). Hospital information systems: Measuring end user computing satisfaction (eucs). *Journal of Biomedical Informatics*, *45*(3), 566-579. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1532046412000391` doi: https://doi.org/10.1016/j.jbi.2012.02.009 40

Allen, I. E. & Seaman, J. (2014). Grade change tracking online education in the united states. Retrieved from `www.favazza.com` 1

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*, 600-614. Retrieved from `https://doi.org/10.1007/s40593-016-0105-0` doi: 10.1007/s40593-016-0105-0 1

Barnard-Brak, L., Lan, W., To, Y., Paton, V. & Lai, S.-L. (2009, 01). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*, *12*, 1-6. doi: 10.1016/j.iheduc.2008.10.005 39

Bei, E., Mavroidis, I. & Giossos, Y. (2020, 01). Development of a scale for measuring the learner autonomy of distance education students. *European Journal of Open, Distance and E-Learning*, *22*, 133-144. doi: 10.2478/eurodl-2019-0015 38

Bodily, R. & Verbert, K. (2017, 10). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, *10*, 405-418. doi: 10.1109/TLT.2017.2740172 1, 6, 7, 10, 11, 20, 63

Boud, D. & Molloy, E. (2012, 01). Rethinking models of feedback for learning: The challenge of design. *Assessment Evaluation in Higher Education - ASSESS EVAL HIGH EDUC*, *38*, 1-15. doi: 10.1080/02602938.2012.691462 7

Brooke, J. (1995, 11). Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, *189*. 40

Canvas Data Portal. (n.d.). *Canvas Data warehouse documentation Version 4.2.9, year = n.d., url = https://portal.inshosteddata.com/docs,*. 31

Corrin, L. (2022). *Evaluating students' interpretation of feedback in interactive dashboards.* Routledge. doi: 10.4324/9781351136501-13 28

Corrin, L. & Barba, P. D. (2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards closing the loop view project supporting self-regulated learning through personalised analytics-based feedback in stem education view project. Retrieved from `https://www.researchgate.net/publication/271769111` 9, 10, 11, 13, 14, 16, 17, 19, 60

Corritore, C. L., Marble, R. P., Wiedenbeck, S., Kracher, B. & Chandran, A. (2005). Measuring online trust of websites: Credibility, perceived ease of use, and risk. In *Americas conference on information systems.* 40

Duncan, T. & Mckeachie, W. (2010, 06). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, *40*, 117-128. doi: 10.1207/s15326985ep4002_6 38

Erdfelder, E., Faul, F. & Buchner, A. (1996). Gpower: A general power analysis program. *Behavior Research Methods, Instruments, Computers*, *28*, 1-11. Retrieved from https://doi.org/10.3758/BF03203630 doi: 10.3758/BF03203630 41, 42

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117 - 140. 15

Fleur, D. S., van den Bos, W. & Bredeweg, B. (2020). Learning analytics dashboard for motivation and performance. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12149 LNCS*, 411-419. doi: 10.1007/978-3-030-49663-0_51 17, 18, 19, 20, 42, 62

Gašević, D., Dawson, S. & Siemens, G. (2015, 1). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*, 64-71. doi: 10.1007/s11528-014-0822-x i, i, i, 1

Hardebolle, C., Jermann, P., Pinto, F. & Tormey, R. (2019). Impact of a learning analytics dashboard on the practice of students and teachers. In (p. 11). Retrieved from http://infoscience.epfl.ch/record/273370 (Proceedings at: https://www.sefi.be/wp-content/uploads/2019/10/SEFI2019$_{P}$*roceedings.pdf*) 18

Jivet, I., Scheffel, M., Drachsler, H. & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10474 LNCS*, 82-96. doi: 10.1007/978-3-319-66610-5_7 2, 6, 11

Jivet, I., Scheffel, M., Specht, M. & Drachsler, H. (2018, 3). License to evaluate: Preparing learning analytics dashboards for educational practice. *ACM International Conference Proceeding Series*, 31-40. doi: 10.1145/3170358.3170421 1, 2, 6, 7, 8, 11, 13, 15, 19, 20, 27, 60, 63

Kim, J., Jo, I.-H. & Park, Y. (2015, 11). Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review*, *17*. doi: 10.1007/s12564-015-9403-8 9, 10, 11, 13, 14, 15, 16, 17, 19, 28, 42, 60, 62

Konert, J., Bohr, C., Bellhauser, H. & Rensing, C. (2016, 11). Peerla - assistant for individual learning goals and self-regulation competency improvement in online learning scenarios. In (p. 52-56). Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICALT.2016.100 16, 17, 18, 20

Kovanovic, V., Gašević, D., Dawson, S., Joksimovic, S. & Baker, R. (2016, 2). Does time-on-task estimation matter? implications on validity of learning analytics findings. *Journal of Learning Analytics*, *2*, 81-110. doi: 10.18608/jla.2015.23.6 11, 60

Kuo, Y., Walker, A., Schroder, K. & Belland, B. (2014). Interaction, internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *Internet and Higher Education*, *20*, 35–50. doi: 10.1016/j.iheduc.2013.10.001 18

Loboda, T., Guerra, J., Hosseini, R. & Brusilovsky, P. (2014, 09). Mastery grids: An open source social educational progress visualization.. doi: 10.1007/978-3-319-11200-8_18 12, 13, 14, 20, 59

Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., Pardo, A., Lim, L., … Tsai, Y.-S. (2020, 9). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics*, *7*, 45-71. Retrieved from `https://learning-analytics.info/index.php/JLA/article/view/7008` doi: 10.18608/jla.2020.72.3 2

Matcha, W., Uzir, N. A., Gasevic, D. & Pardo, A. (2020, 4). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*, 226-245. doi: 10.1109/TLT.2019.2916802 i, i, i, 1, 2, 6, 7, 8, 10, 13, 15, 18, 19, 28, 29, 60, 66

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory Cognition*, *39*, 462-476. Retrieved from `https://doi.org/10.3758/s13421-010-0035-2` doi: 10.3758/s13421-010-0035-2 8

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*. doi: 10.3389/fpsyg.2017.00422 4

Ranalli, J. (2012, 12). Alternative models of self-regulation and implications for l2 strategy research. *Studies in Self-Access Learning Journal*. doi: 10.37237/030403 5, 8

Ryan, R. M. & Deci, E. L. (2000, 1). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*, 54-67. doi: 10.1006/CEPS.1999.1020 15, 19

Sansom, R. L., Bodily, R., Bates, C. O. & Leary, H. (2020, 6). Increasing student use of a learner dashboard. *Journal of Science Education and Technology*, *29*, 386-398. doi: 10.1007/s10956-020-09824-w 11

Santos, J., Verbert, K., Govaerts, S. & duval, e. (2013, 04). Addressing learner issues with stepup!: An evaluation. In (p. 14-22). doi: 10.1145/2460296.2460301 11, 12, 13, 14, 20, 59, 65

Schmitz, B. (2001). Self-monitoring to support the transfer of a self-regulation instruction for students: A process study; [self-monitoring zur unterstützung des transfers einer schulung in selbstregulation für studierende eine prozessanalytische untersuchung]. *Zeitschrift fur Padagogische Psychologie*, *15*, 181 – 197. Retrieved from `https://www.scopus.com/inward/record.uri?eid=2-s2.0-0035603597&doi=10.1024%2f%2f1010-0652.15.34.181&partnerID=40&md5=9ed7caabb8de69c764a2cd2b04f78047` (Cited by: 42) doi: 10.1024//1010-0652.15.34.181 4

Schraw, G. & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*, 460-475. 38

Schumacher, C. & Ifenthaler, D. (2018, 1). Features students really expect from learning analytics. *Computers in Human Behavior*, *78*, 397-407. doi: 10.1016/j.chb.2017.06.030 7, 8

Schunk, D. H. (2008, 12). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review*, *20*, 463-467. doi: 10.1007/s10648-008-9086-3 4, 20

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., . . . Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*, 30-41. doi: 10.1109/TLT.2016.2599522 1

Stata statistical software: Release 14. college station, TX: StataCorp LP. (2015). 43

Tan, J., Yang, S., Koh, E. & Jonathan, C. (2016, 11). Fostering 21st century literacies through a collaborative critical reading and learning analytics environment: User-perceived benefits and problematics. , 430-434. doi: 10.1145/2883851.2883965 16, 17, 19, 28

Venkatesh, V. & Bala, H. (2008, 05). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences - DECISION SCI*, *39*, 273-315. doi: 10.1111/j.1540-5915.2008.00192.x 14

Verbert, K., Govaerts, S., Duval, E., Santos, J. L. & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*, 1499-1514. doi: 10.1007/s00779-013-0751-2 1, 6, 11

Winne, P. H. & Hadwin, A. F. (1998). *Studying as self-regulated learning* (1st ed., Vol. 93). 5, 7, 27, 28, 29, 60

Wise, A. F., Zhao, Y. & Hausknecht, S. N. (2014). Learning analytics for online discussions: Embedded and extracted approaches. *Journal of Learning Analytics*, *1*, 48-71. 15

Zimmerman, B. & Schunk, D. (2001). Self-regulated learning and academic achievement : theoretical perspectives.
2, 4, 6, 7, 20

A. Informed Consent Interviews

Informed consent form for the study 'Online learning dashboards'.

This study is performed by Bram Peters, a MSc student of the Human-Technology Interaction group at Eindhoven University of Technology under the supervision of Uwe Matzat and Rianne Conijn.

Before participating, you should understand the procedure followed in this study, and give your informed consent for voluntary participation. Please read this page carefully.

About this study

This study has the goal to understand what students find important to have in online learning dashboards. Online learning dashboards are visual representations of a students' progress through a course or program and is used to give feedback to students about their study progress and strategies where needed, if necessary. You will be asked a number of open-ended questions by the interviewer that focus on how best to design such a learning dashboard, including what features a learning dashboard should or should not have. For this study you do not have to share your Canvas page, study results or a personal dashboard whatsoever.

This study will take 15-20 minutes to complete and does not involve any risks.

Voluntary Participation

Your participation is completely voluntary and anonymous. In the consent form, you can indicate whether you agree to also use your data for scientific purposes. You can stop participation at any time, but you will only receive payment if you complete the interview. You can also withdraw your permission to use your data up to 1 month after completing this study. You will be compensation €5,- if you complete this study, which will be transferred to you through Tikkie or another form of 'betaalverzoek'. Any contact details shared through one of these platforms will be deleted after successful payment.

Confidentiality and use, storage, and sharing of data

This study has been approved by the Ethical Review Board of Eindhoven University of Technology. In this study personal data (e.g. age, study program) and interview data, including audio recordings, will be stored temporarily. Audio recordings will be made during the interview with a speech recorder application to transcribe the data and use it for analysis later. After processing and analyzing, the audio file will be deleted and will not be used for any other purpose. The anonymized dataset generated with these audio recordings, that to the best of our knowledge and ability will not contain information that can identify you, will only be discussed in the final research paper and not be made publicly available.

Further information

If you want more information about this study, the study design, or the results, you can contact Bram Peters (contact email: b.a.peters@student.tue.nl). You can report irregularities related to scientific integrity to confidential advisors of the TU/e, whose contact information can be found on www.tue.nl.

Certificate of consent

By starting this study, ☐ I indicate that I have read and understood the study procedure, and I agree to voluntarily participate. ☐ I also give permission to make my anonymized recorded data available to others in a public online data repository. ☐

Name **Signature**

B. Interview Guide

1. Do you have any past experiences with learning dashboards?

2. What are the features and kinds of visual information that first come to mind, and that you would find valuable, when you think of a learning dashboard like the one that I just described?

3. How do you think a learning dashboard could help you with setting goals for yourself? What kind of visual information or features do you think would help with this?

4. How do you think a learning dashboard could help you with your planning during a course? What kind of visual information or features do you think would help with this?

5. How do you think a learning dashboard could help you with evaluating how well you are doing during a course? What kind of visual information or features do you think would help with this?

6. How would it affect your performance and motivation if you could see the results of other students? So, you could see how well they are doing compared to you?

7. Do you think you would trust the advice that a learning dashboard gives to you?

8. What kind of information would you definitely not want in a dashboard? Or information that you would never want to share with a teacher or another student?

C. Prototype Demonstration

VIDEO DEMONSTRATION (*PRFdashboard*):
https://1drv.ms/v/s!Av-QTsRkdvNsj4gZvOkEJMxK8uvCFA?e=M9QZhv

[TUTORIAL WALKTHROUGH] *

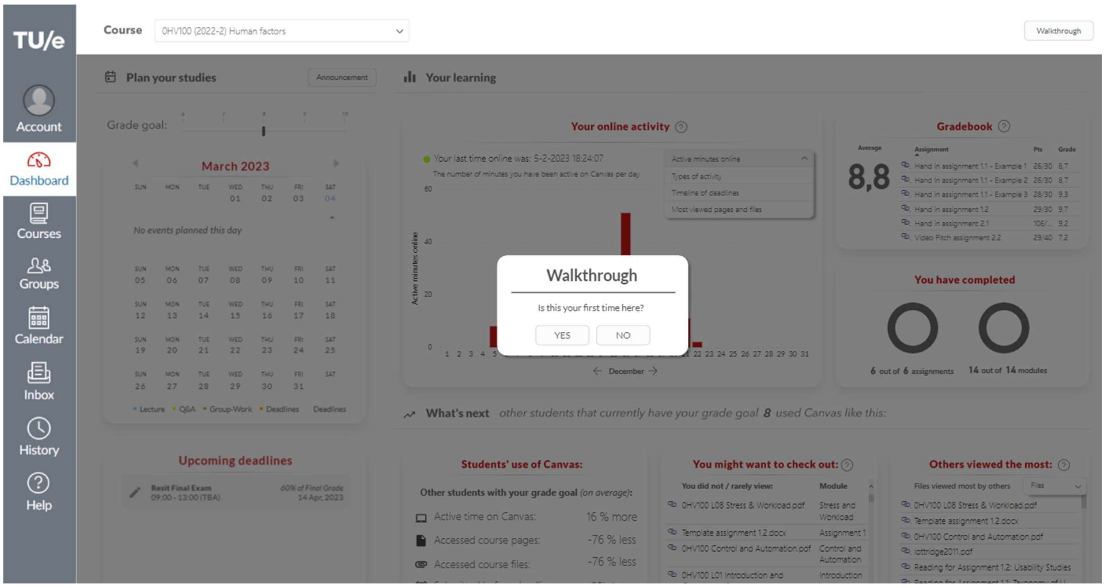


Figure C.1: New users are prompted to complete a walkthrough tutorial at first log-in

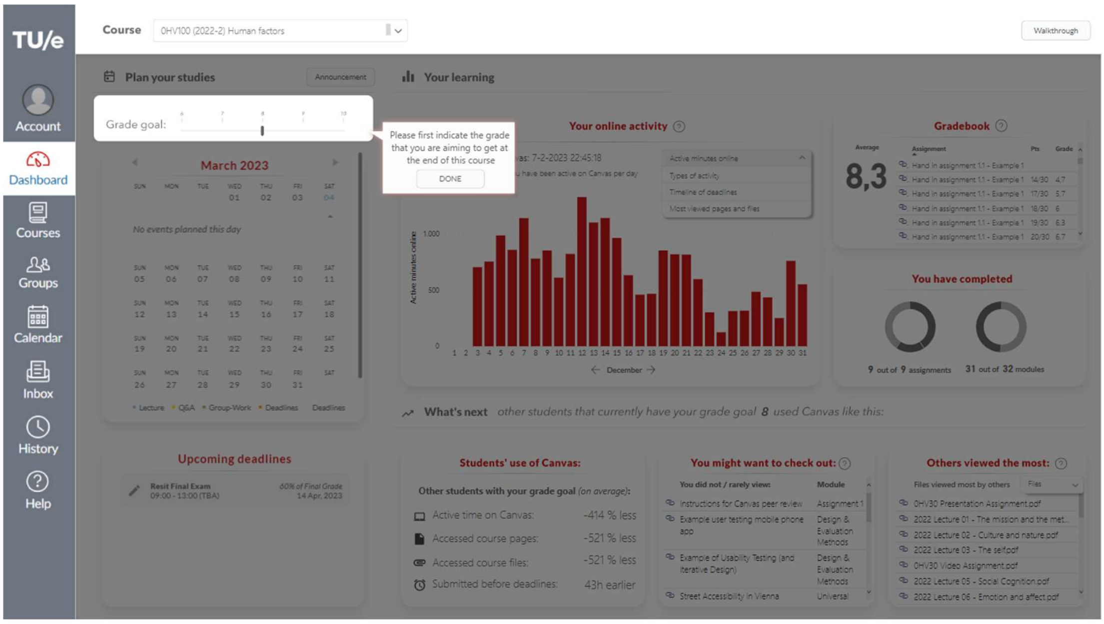


Figure C.2: Users are asked to indicate their grade goal through a slider in the 'Plan your studies'-panel (i.e., *Forethought phase features*).

* *The screenshots display a learning dashboard that is based on the pseudonymized data of a real student. All individually identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset and visualizations, and replaced by one or more artificial identifiers.*

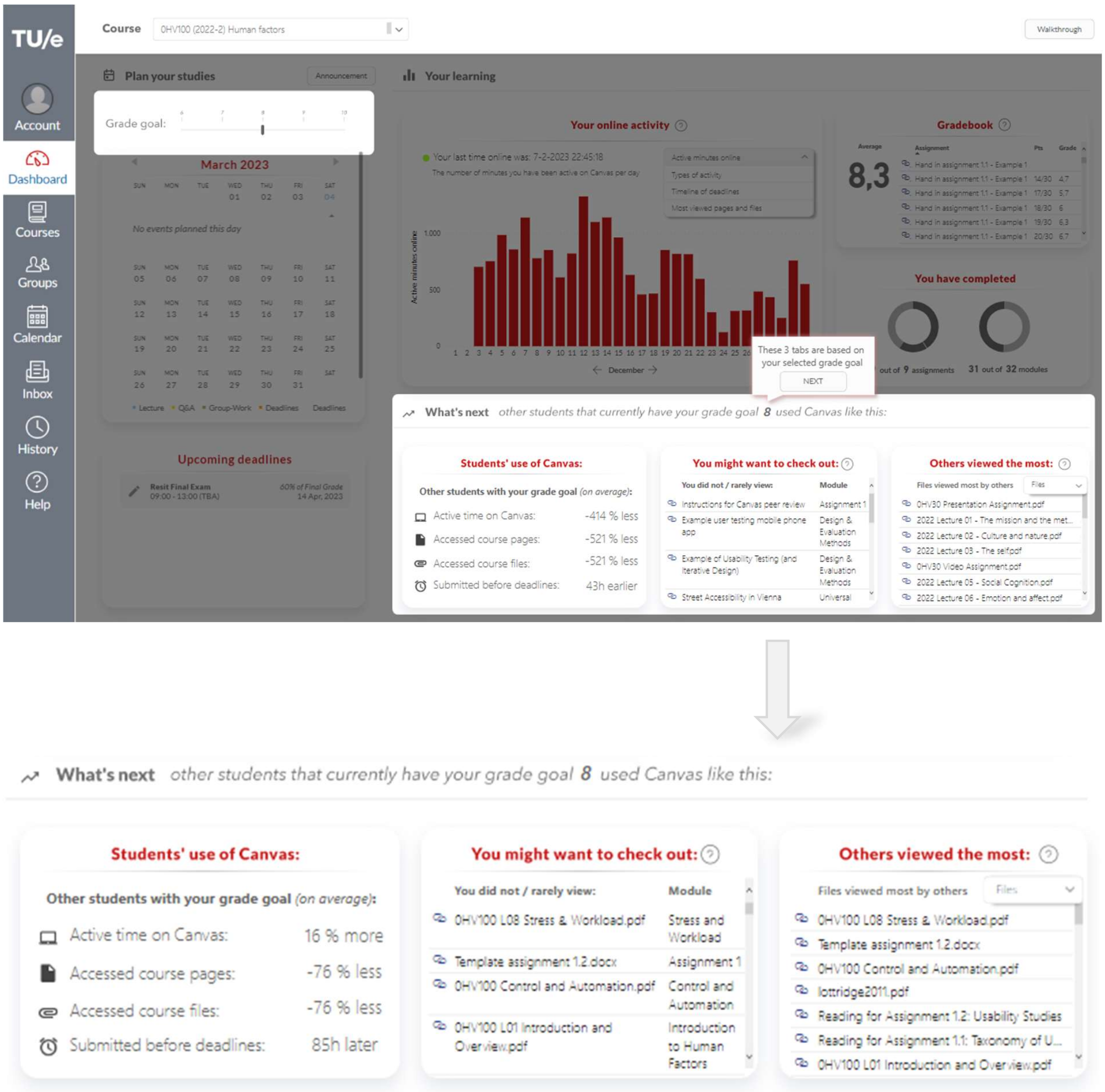
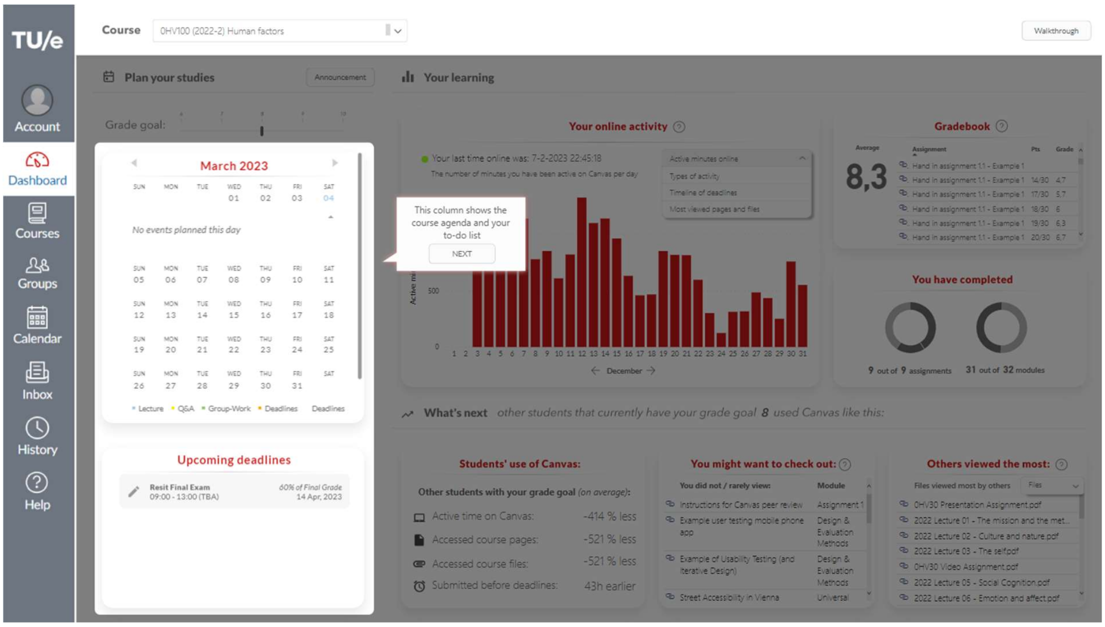


Figure C.3: The visualizations in the *Self-Reflection phase features* (i.e., the 'What's next'-panel) are filtered on the users grade goal. Data used in these visualizations are now based on a subset of students that have the grade that the user is aiming for.

** The screenshots display a learning dashboard that is based on the pseudonymized data of a real student. All individually identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset and visualizations, and replaced by one or more artificial identifiers.*



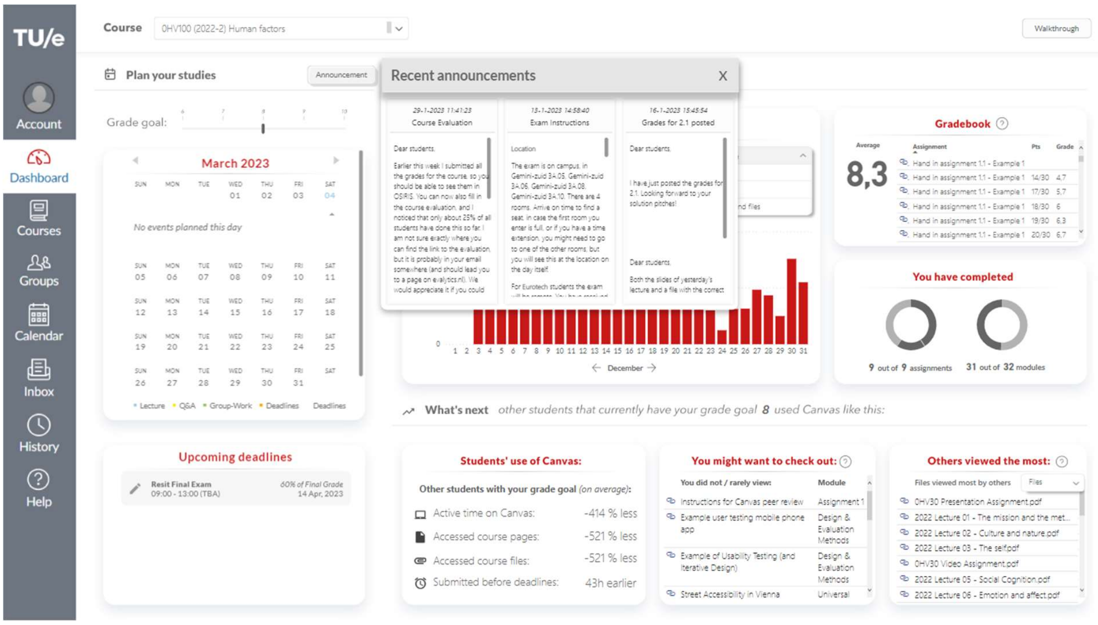
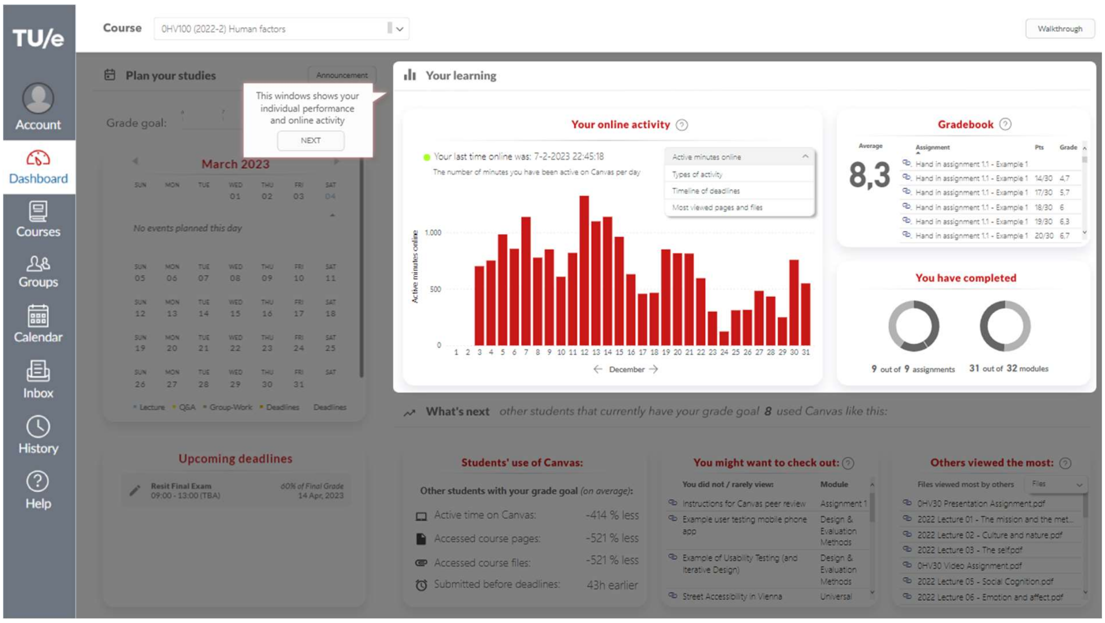


Figure C.4: The 'Plan your studies'-panel (i.e., *Forethought phase features*) also contains a course calendar, an upcoming deadlines list, and a recent course-announcements panel.

** The screenshots display a learning dashboard that is based on the pseudonymized data of a real student. All individually identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset and visualizations, and replaced by one or more artificial identifiers.*



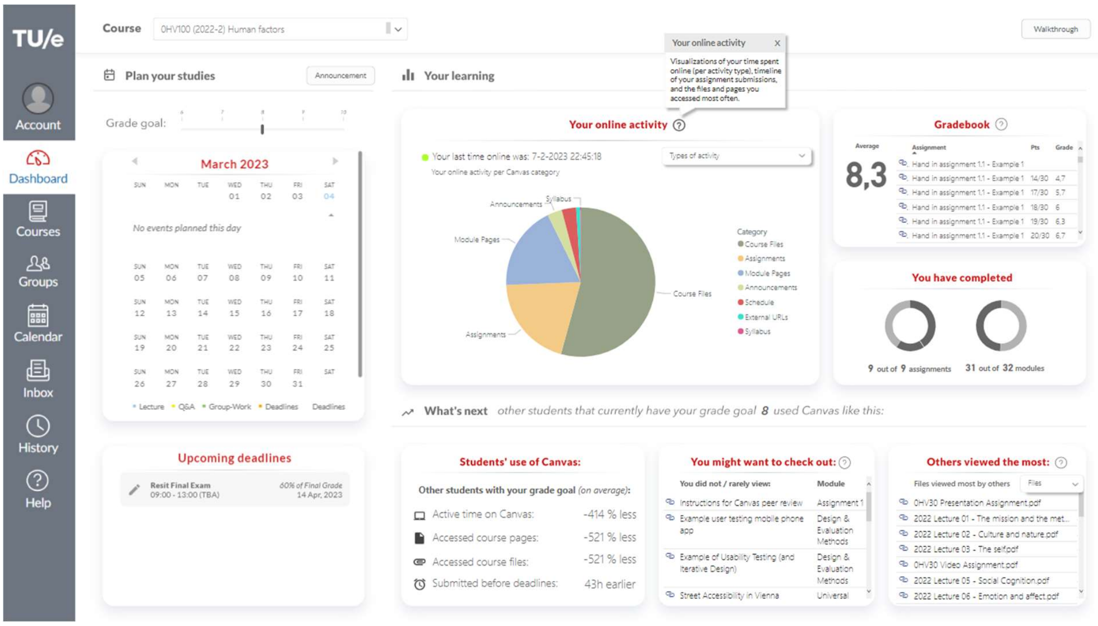


Figure C.5: The tutorial ends with the 'Your learning'-panel (i.e., *Performance phase features*). Visualizations include a bar chart displaying the user's daily online minutes, a pie chart illustrating the types of files they accessed and a timeline graph in which submissions are visualized with respect to assignment deadlines. In addition, the 'Your learning'-panel includes a gradebook and two Canvas progression indicators (i.e., completed modules and completed assignments).

* *The screenshots display a learning dashboard that is based on the pseudonymized data of a real student. All individually identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset and visualizations, and replaced by one or more artificial identifiers.*

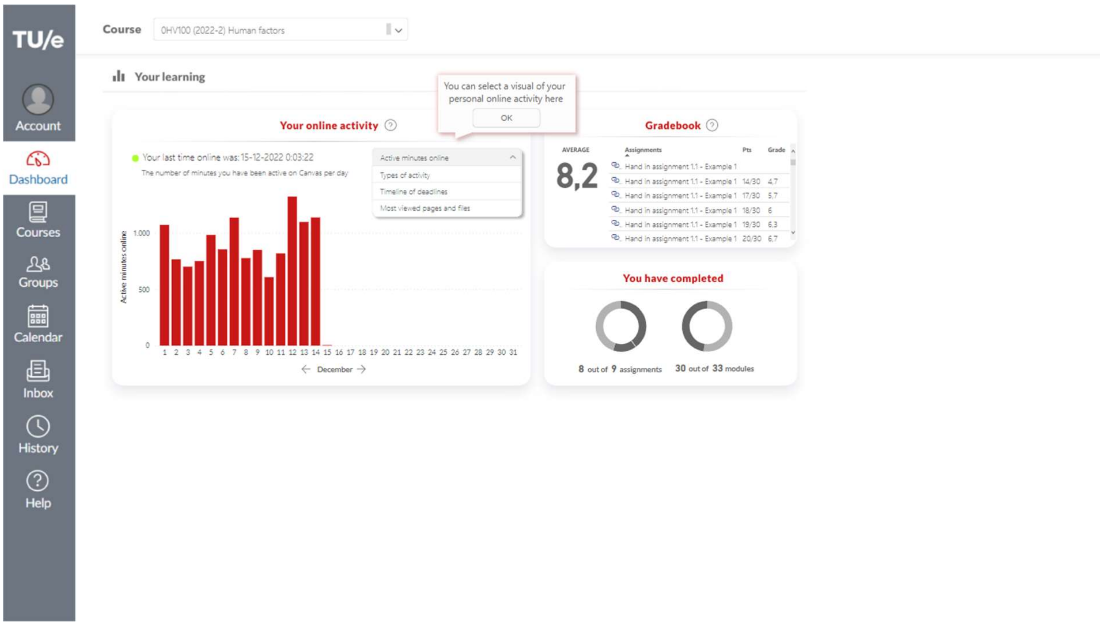


Figure C.6: PRFdashboard included a similar tutorial; although it only included the 'Your learning'-panel (i.e., the *Performance phase* features).

* *The screenshots display a learning dashboard that is based on the pseudonymized data of a real student. All individually identifiable data (e.g., name, e-mail, IP-address) was removed from the dataset and visualizations, and replaced by one or more artificial identifiers.*

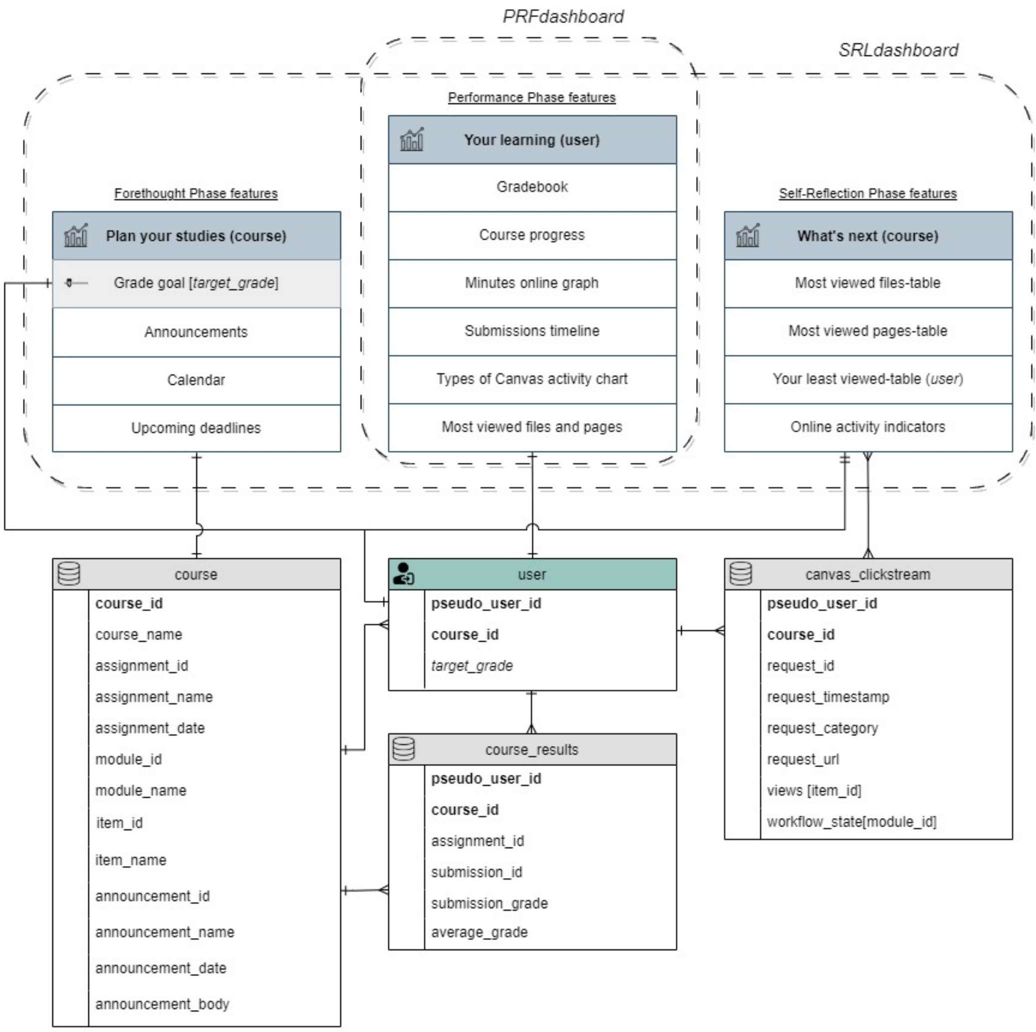


Figure C.7: Schematic diagram displaying the data model underlying the dashboard visualization. The features that support the different phases of self-regulated learning respectively (i.e., and thus define the differences for both dashboard) are displayed in the blue-headered boxed on top. Underneath, the data model as described in section 4.3 is displayed that was used to populate the visualizations.

D. Information Sheet Experiment

Information sheet and informed consent form for study: 'Design and Evaluation of Student-Facing Learning Dashboards Using Theory of Self-Regulated Learning'.

This study is performed by Bram Peters, a MSc Human-Technology Interaction-student at Eindhoven University of Technology under the supervision of Uwe Matzat and Rianne Conijn. You have been invited to take part because you are a student in one of the following courses:

- 0HV30 (Social Psychology & Consumer Behaviour)
- 0HV100 (Human Factors)

Before giving your informed consent, you should understand the procedure followed in this study, be aware of potential risks involved, and be informed about the way that personal data is collected, stored and processed.

Voluntary participation

Your participation is completely voluntary, but participation requires you to consent to using your personal Canvas-data and the potential risks involved in this study. You can stop participation at any time, but you will only receive a full compensation if you complete all parts of the study. Cancelling your participation will have no negative consequences for you or your study/course results. You can also withdraw your permission to use your data up to 1 month after completing of this study. You will be compensated €15,- after full completion, which will be transferred to you through a payment-request app (e.g. Tikkie). You will not receive any compensation for partial completion of the experiment. Any contact details shared through one of these applications will be deleted after successful payment.

Aim of the study

The overall goal of this study is to understand how a new online learning dashboard for Canvas could help and motivate students with their learning. Online learning dashboards are visual representations of a students' performance through a course and are used to give feedback to students about their study progress and strategies where needed, if necessary. The research project to which this study belongs aims to understand how best to design a learning dashboard that can motivate students and provide them with personalized feedback to improve their learning.

Procedure

For students of the 0HV30 or 0HV100-course in Q2 of academic year 2022/2023 that participate in this study, a learning dashboard will be created based on their Canvas-data of the course. Canvas automatically keeps track of your course results as well as the way

that you interact with the system (i.e. through clickstream-data). You will get access to a learning dashboard that is based on this data and shows you how well you are performing in the course in comparison to others. After giving your informed consent, you will be provided access to this learning dashboard for a period of 5 weeks.
During these 5 weeks, you will be asked to fill in three surveys based on your learning:

Survey 1 (available from December 5th until December 15th)
Survey 2 (available from December 19th until December 23rd)
Survey 3 (available from January 9th until January 13th)

Survey 1 will immediately follow this consent form. If you have succesfully filled in this consent form and Survey 1 before the deadline (see above), you will get access to your learning dashboard on December 19th and be notified by email. After this, you will immediately get access to **Survey 2**. After the Christmas break, you will be asked to complete **Survey 3**. Survey 2 and survey 3 will be sent to you over email and will take approximately 15-20 minutes to complete. Before you start filling in these surveys, you are kindly requested to access and review your learning dashboard for a moment. Of course, accessing your learning dashboard outside of the surveys is allowed and encouraged. You do not need to do anything else except participate in this course as usual. After you filled in **all** surveys successfully, you will be debriefed and rewarded your compensation.

Note that the learning and study materials for these courses are equal for all students, whether you participate in the study or not. Hence, the quality of provided education is equal for participating and non-participating students.

Duration

The three surveys are expected to take 65 minutes in total.

Which personal data will be collected?

The researcher will collect your Canvas-data, dashboard-interaction data, and your survey responses. The Canvas-data (i.e. for 0HV30 or 0HV100) that will be collected includes:

- user-id Canvas;
- student-number;
- course-results (all results of the course);
- course information (e.g. course codes, deadlines, examination dates, lectures);
- course setup (information on modules, lectures, discussion forums, wikis, assignments);
- students' answers to tests and assignments;
- log-in data (timestamps of when you logged into the course and for how long);
- clickstream data (timestamps of every click within the course);

The dashboard-interaction data that will be collected includes:

- log-in data (timestamps of when you logged into the dashboard and for how long);
- clickstream data (every click within a specific element of the dashboard with time stamps);

From the surveys, the following information is collected:

- self-regulated learning skills in relation to the course (e.g. goal-setting, task-planning, self-monitoring, self-reinforcement, self-assessment skills)
- your affective evaluations of the course (e.g. motivation, engagement, sense of control, perceived autonomy)
- your affective evaluations of the dashboard (e.g. how easy it is to use)

Who will have access to my data?

Access to the Canvas- and dashboard-data used in this study is strictly limited to the researcher and the technical team who makes the data available (i.e. Learning Analytics-department at TU/e Information Management Services). The technical team has access to personal non-pseudonymized data, but the researcher works only with pseudonymized data. This means that all directly identifiable data is either removed from the dataset (such as name, gender, email address, IP-address, personal comments/feedback, and personal input) or pseudonymized (such as student number). Individually identifiable data within a dataset are replaced by one or more artificial identifiers to ensure that the possibility to identify a student based on the data is lowered to a minimum. Teachers of the 0HV100 and 0HV30-courses do not have access to this data nor the dashboard.

Privacy and security measures

Even though all directly identifiable data is removed from the dataset or pseudonymized, combinations of data values may possibly unintentionally result in the ability to identify a student. To ensure your data is handled with care and integrity, this study has been evaluated through a Data Protection Impact Assessment (DPIA), and the researcher has signed the TU/e Code of Conduct of Scientific Integrity. Your personal data is processed on the lawful ground of 'consent of the data subject' laid down in Art. 6 (1) (a) GDPR as well as for research purposes (Article 9(2)(j) of the GDPR of the GDPR). In the consent form, you can give optional permission to store your pseudonymized research data (i.e. Canvas- and dashboard-interaction data and survey responses) for use in future research on learning motivation at the Human-Technology Interaction-group or the Learning Analytics-department of Information Management Services at TU/e, with due regard for recognized ethical standards for scientific research, and for education purposes of your research data. You can withhold your permission and still participate in this study. The research results that may be published will not in any way include personal data or confidential information through which anyone can recognize you.

The Canvas-data that is used for this study is encrypted and securely stored in TU/e administration and learning management systems. The combined dataset(s) used for this study are securely stored on storage solutions offered by TU/e IMS Services. In this study, your data will be presented through a learning dashboard developed in Microsoft Power BI. Microsoft complies with the standards that apply for storing data on European servers. By participating in this study, you also consent to the terms and conditions of Microsoft. For more information about Microsoft's privacy statement, click this link.

If you have specific questions concerning the handling of personal data you may direct these to the privacy team (privacy@tue.nl) or data protection officer of TU/e by sending a mail to: dataprotectionofficer@tue.nl. Furthermore, you have the right to file a complaint with the Data Protection Authority (in Dutch: *Autoriteit Persoonsgegevens*).

Risks involved

This study will provide you feedback on your study progress and performance in comparison to others through a learning dashboard. If you experience negative consequences because of this feedback (e.g. stress, anxiety), you are encouraged to contact the researcher (Bram Peters: b.a.peters@student.tue.nl) to resolve any unclarity or ambiguity with regards to what results the dashboard showed you and what they mean. If you feel that you face difficulties proceeding with your studies after participating in this study, please contact your study-advisor (find your study-advisor here).

Confidentiality

All research conducted at the Human-Technology Interaction Group adheres to the Code of Ethics of the NIP (Nederlands Instituut voor Psychologen – Dutch Institute for Psychologists). This study has been approved by the Ethical Review Board of Eindhoven University of Technology.

Contact information

If anything remains unclear about the purpose of the study, its design, data privacy or potential risks, or if you want to cancel your participating during the study, please contact Bram Peters (contact email: b.a.peters@student.tue.nl). You can report irregularities related to scientific integrity to confidential advisors of the TU/e, whose contact information can be found on www.tue.nl.

Informed consent-form for study: 'Design and Evaluation of Student-Facing Learning Dashboards Using Theory of Self-Regulated Learning'.

Through this consent form I recognize the following:

1. I am sufficiently informed about the study. I have read the information about this study and the procedure and have subsequently had the opportunity to ask questions. These questions have been answered satisfactorily.

2. I take part in this study voluntarily. I do not take part under any kind of explicit or implicit duress. It is clear to me that I can cancel my participation at any moment without having to provide any reason. I do not have to answer a question against my wish.

3. I give permission to process the personal data that are collected from me during this study in the way described in the information sheet.

   YES ☐ NO ☐

Beside the above, you can below give optional permission for further use of your research data. You can withhold your permission and still participate in this study.

4. I give permission to store the pseudonymized research data collected from me (i.e. Canvas- and dashboard-interaction data and survey responses) for use in future research on learning motivation at the Human-Technology Interaction-group or the Learning Analytics-department of Information Management Services at TU/e, with due regard for recognized ethical standards for scientific research, and for education purposes.

   YES ☐ NO ☐

Name of participant: **Date:**

Signature:

E. Survey Measurements

[MOTIVATION] (Survey 1, Survey 2, Survey 3)

Motivation Level

Within {placeholder_course}, please indicate what describes your feelings about following the course best:

1.	Motivated	1 2 3 4 5	Unmotivated
2.	Interested	1 2 3 4 5	Uninterested
3.	Involved	1 2 3 4 5	Uninvolved
4.	Stimulated	1 2 3 4 5	Unstimulated
5.	Inspired	1 2 3 4 5	Uninspired
6.	Enthused	1 2 3 4 5	Unenthused
7.	Excited	1 2 3 4 5	Unexcited
8.	Fascinated	1 2 3 4 5	Unfascinated

Intrinsic/Extrinsic Motivation

How much do you feel like the following statements correspond with why you are participating in {placeholder_course}? Please answer using the following scale: *(1) Corresponds not at all, (2) Corresponds very little, (3) Corresponds a little, (4) Corresponds moderately, (5) Corresponds enough, (6) Corresponds a lot, (7) Corresponds exactly*.

I participate in this course…

1. Because I think that this course is interesting.
2. Because I think that this course is pleasant.
3. Because I feel good when doing this course.
4. Because I'm doing it for my own good.
5. Because I believe that this course is important for me.
6. Because I am supposed to do it.
7. Because I don't feel like I have a choice.

[PERCEIVED AUTONOMY] (Survey 1, Survey 2, Survey 3)

Within {placeholder_course}, how much do you agree with the following statements? Please answer using the following scale: *(1) Strongly disagree, (2) Moderately disagree, (3) Neither disagree nor agree, (4) Moderately agree, (5) Strongly agree*.

1. I am aware of my abilities and limits in this course.
2. I feel like I can easily adapt to difficult situations in this course.
3. I find it easy to work on my own without the help of others in this course.
4. I feel like I can set realistic goals that meet what I want in this course.
5. I am aware of the steps I should take in order to pursue my goals in this course.
6. I know where to look for resources that I need to find to meet my goals.

[METACOGNITIVE AWARENESS] (Survey 1, Survey 3)

Within {placeholder_course}, how well do the following behaviors, thoughts, and feelings describe you? Please answer using the following scale: *(1) Not at all characteristic of me, (2) Not really characteristic of me, (3) Moderately characteristic of me, (4) Characteristic of me, (5) Very characteristic of me*.

1. I understand my strengths and weaknesses in this course.
2. I know what kind of information is most important to learn in this course.
3. I am aware of what strategies I use when I study for this course.
4. I find it easy to set specific goals before I begin a task in this course.
5. I think it is difficult to decide what to do in this course.
6. I find myself analyzing the usefulness of strategies while I study.

[PERCEIVED SUPPORT FOR SELF-REGULATED LEARNING] (Survey 2, Survey 3)

How much would you agree with the following statements while using this learning dashboard for {placeholder_course}? Please answer using the following scale: *(1) Strongly disagree, (2) Moderately disagree, (3) Neither disagree nor agree, (4) Moderately agree, (5) Strongly agree*.

Goal-setting

1. I feel like this learning dashboard represents the structure of the learning material in this course well.
2. I feel like this learning dashboard can help me with setting appropriate goals for myself in this course.
3. I feel like this learning dashboard supports me in making a realistic list of things I want to do in this course.

Task-planning

4. I feel like this learning dashboard can help me with managing my time appropriately in this course.
5. I feel like this learning dashboard could help me to decide what tasks have more priority than others in this course.
6. I feel like this learning dashboard supports me in making a realistic planning of when I need to complete certain tasks in this course.

Performance

7. I feel like this learning dashboard supports me in monitoring how well I'm doing in this course.
8. I feel like this learning dashboard reflects my performance in this course well.
9. I feel like this learning dashboard reflects my knowledge in this course well.

Self-Evaluation

10. I feel like this learning dashboard understands how I study in this course well.
11. I feel like this learning dashboard can assist me in achieving the goals I set for myself in this course.
12. I feel like this learning dashboard is able to show me where I can improve on what I do in this course.

[USABILITY] (Survey 2, Survey 3)

How much would you agree with the following statements while using this learning dashboard for {placeholder_course}? Please answer using the following scale: *(1) Strongly disagree, (2) Moderately disagree, (3) Neither disagree nor agree, (4) Moderately agree, (5) Strongly agree*.

1. I think that I would like to use this learning dashboard frequently.
2. I found this learning dashboard unnecessarily complex.
3. I thought this learning dashboard was easy to use.
4. I think that I would need the support of a technical person to be able to use this learning dashboard.
5. I found the various functions in this learning dashboard were well integrated.
6. I thought there was too much inconsistency in this learning dashboard.
7. I would imagine that most people would learn to use this learning dashboard very quickly.
8. I needed to learn a lot of things before I could get going with this learning dashboard.
9. I am satisfied with having this learning dashboard for this course.
10. This learning dashboard provides me the information that I need.
11. I feel like this learning dashboard is a useful tool for my studies in this course.
12. I think the learning dashboard presents my data in a useful manner.

[TRUST] (Survey 2, Survey 3)

How much would you agree with the following statements while using this learning dashboard for {placeholder_course}? Please answer using the following scale: *(1) Strongly disagree, (2) Moderately disagree, (3) Neither disagree nor agree, (4) Moderately agree, (5) Strongly agree*.

1. I feel like there could be negative consequences from using this learning dashboard.
2. I feel it is unsafe to interact with this learning dashboard.
3. This learning dashboard provides truthful information.
4. The information provided by this learning dashboard is believable.
5. The content of the learning dashboard is what I expected.
6. There were no surprises in how the learning dashboard responded to my actions.
7. I believe this learning dashboard is trustworthy.
8. I don't expect this learning dashboard to take advantage of me.

[OPEN-ENDED QUESTIONS] (Survey 2, Survey 3)

1. How would you describe your experiences with this learning dashboard?

   [________LONGTEXT______]
2. Based on the last couple of weeks within {placeholder_course}, which features do you think were most valuable?

   [________LONGTEXT______]
3. Based on the last couple of weeks {placeholder_course}, what did you feel was missing in the learning dashboard?

   [________LONGTEXT______]
4. Any additional remark:

   [________LONGTEXT______]

[EXAMPLE PAGE] (Survey 1: Welcome Screen)

TU/e

Welcome!

This survey includes the registration form and first part of the study: **'Design and Evaluation of Student-Facing Learning Dashboards Using Theory of Self-Regulated Learning'.**

As a student from the **0HV100** and/or **0HV30**-course, you are invited to participate in a new study that will give you the opportunity to get **insight** into the way that you study online and use Canvas. You will get access to your own **personal Learning Dashboard** that will give you information and visualizations of your course results and interaction with Canvas, all based on **real** data. This is an exciting opportunity for you to see your **performance** in 0HV100 or 0HV30, but also to be part of the **research** on Learning Dashboards as a potential new Canvas-feature! And on top of that, you will receive a **€15,- reward** for participation!

This study consists of **3 surveys** and includes access to your own personal Learning Dashboard. Registration for this study will be possible until **Friday December 16th 10:00.**

By clicking **Next**, you will be taken to the Information Sheet and the Informed Consent-form.

Next

[EXAMPLE PAGE] (Survey 3: Usability Questionnaire)

58%

*How much would you agree with the following statements while using this learning dashboard for ?

	strongly disagree	moderately disagree	neither agree nor disagree	moderately agree	strongly agree
I think that I would like to use this learning dashboard frequently.	○	○	○	○	○
I found this learning dashboard unnecessarily complex.	○	○	○	○	○
I thought this learning dashboard was easy to use.	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this learning dashboard.	○	○	○	○	○
I found the various functions in this learning dashboard were well integrated.	○	○	○	○	○
I thought there was too much inconsistency in this learning dashboard.	○	○	○	○	○
I would imagine that most people would learn to use this learning dashboard very quickly.	○	○	○	○	○
I needed to learn a lot of things before I could get going with this learning dashboard.	○	○	○	○	○

F. Results Tables and Figures

Table F.1: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and perceived autonomy as the repeated measured dependent variable.

	Partial SS	df	MS	R-squared = F	0.8372 p
Model	40.492026	65	.62295425	8.62	0.0000
dashboard	.7762545	1	.7762545	1.18	0.2818
id|dashboard	38.816881	59	.65791324		
course	.12538708	1	.12538708	1.74	0.1904
phase	.12496224	2	.06248112	0.86	0.4239
phase#dashboard	.03469658	2	.01734829	0.24	0.7869
Residual	7.8735023	109	.07223397		
Total	48.365528	174	.27796281		

Table F.2: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and metacognitive awareness as the repeated measured dependent variable.

	Partial SS	df	MS	R-squared = F	0.8461 p
Model	43.815115	65	.69547801	4.62	0.0000
dashboard	.39012305	1	.39012305	0.55	0.4629
id|dashboard	42.166694	59	.71468973		
course	.76793931	1	.76793931	5.11	0.0280
phase	.1636394	1	.1636394	1.09	0.3016
phase#dashboard	.19192225	1	.19192225	1.28	0.2637
Residual	7.9712088	53	.15040017		
Total	51.786324	116	.44643383		

Table F.3: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and perceived support for self-regulated learning as as the repeated measured dependent variable.

	Partial SS	df	MS	R-squared = F	0.7621 p
Model	35.501195	59	.60171516	2.88	0.0000
dashboard	.31824507	1	.31824507	0.51	0.4801
id|dashboard	34.621545	55	.62948263		
course	.69444451	1	.69444451	3.32	0.0740
phase	.25965235	1	.25965235	1.24	0.2702
phase#dashboard	.06647051	1	.06647051	0.32	0.5753
Residual	11.082141	53	.20909699		
Total	46.583335	112	.41592264		

Table F.4: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and perceived trustworthiness as as the repeated measured dependent variable.

				R-squared =	0.8267
	Partial SS	**df**	**MS**	**F**	**p**
Model	24.569135	59	.41642601	4.29	0.0000
dashboard	.27518032	1	.27518032	0.63	0.4321
id|dashboard	24.15912	55	.43925673		
course	.00390625	1	.00390625	0.04	0.8418
phase	.17253606	1	.17253606	1.78	0.1884
phase#dashboard	.05719515	1	.05719515	0.59	0.4463
Residual	5.1493389	53	.09715734		
Total	29.718473	112	.26534351		

Table F.5: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and perceived usefulness as as the repeated measured dependent variable.

				R-squared =	0.8096
	Partial SS	**df**	**MS**	**F**	**p**
hline *Model*	41.77064	59	.70797696	3.82	0.0000
dashboard	.09530628	1	.09530628	0.13	0.7234
id|dashboard	41.424057	55	.75316467		
course	.08999997	1	.08999997	0.49	0.4890
phase	.03014228	1	.03014228	0.16	0.6884
phase#dashboard	.15232418	1	.15232418	0.82	0.3688
Residual	9.8244033	53	.1853661		
Total	51.595044	112	.46067003		

Table F.6: Results table of a repeated measures mixed-ANOVA with dashboard-group as between-subject factor, time as within-subject factor, and average grade as as the repeated measured dependent variable.

				R-squared =	0.9789
	Partial SS	**df**	**MS**	**F**	**p**
Model	57.292831	52	1.1017852	23.24	0.0000
dashboard	.037939	1	.037939	0.07	0.7971
id|dashboard	27.256364	48	.56784092		
course	1.0986018	1	1.0986018	23.17	0.0001
phase	.103176	1	.103176	2.18	0.1522
phase#dashboard	.01811841	1	.01811841	0.38	0.5419
Residual	1.2328971	26	.04741912		
Total	58.525728	78	.75032985		

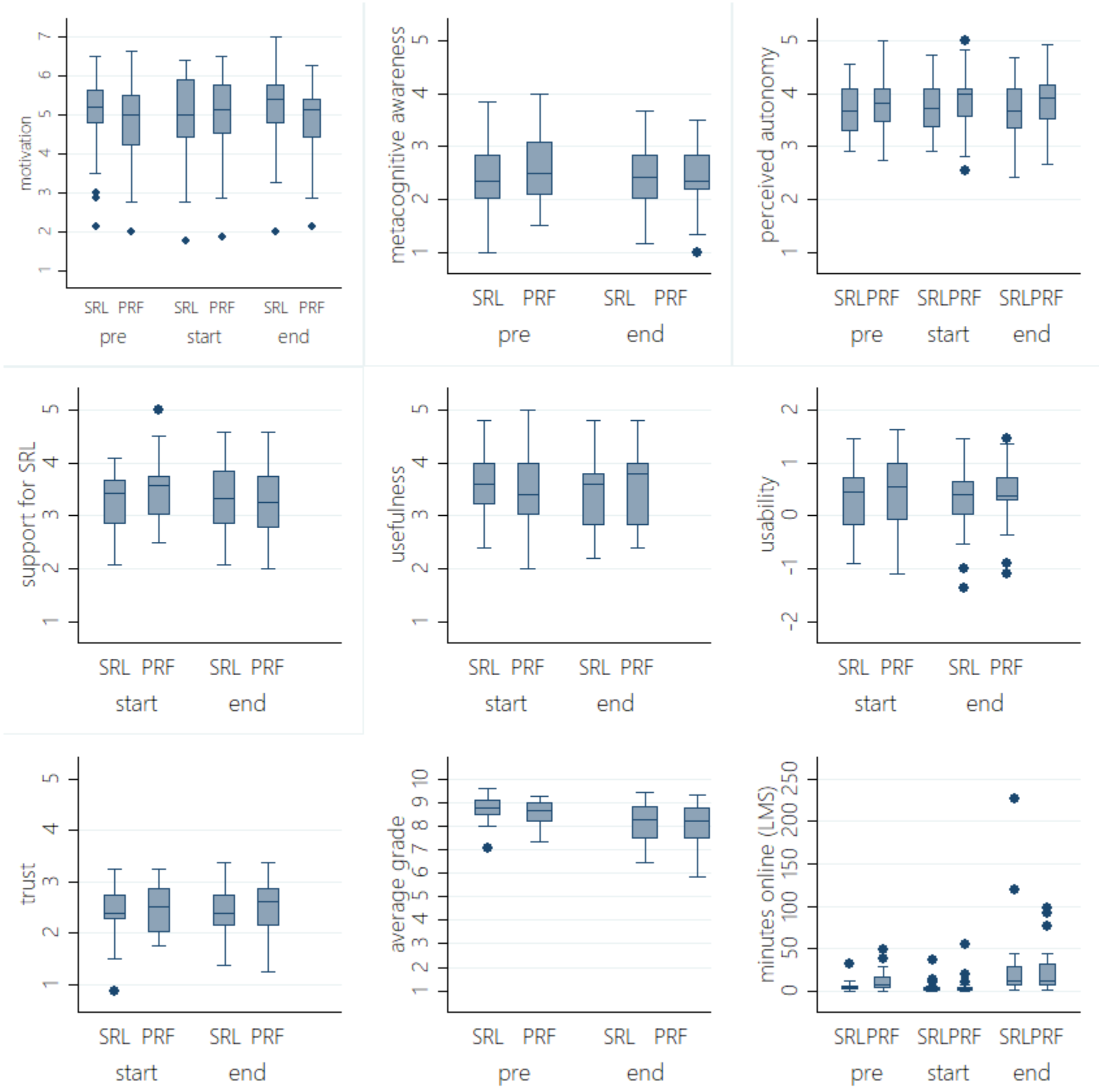


Figure F. 1: Boxplots displaying the distribution of the measurements across different phases of the experiment by dashboard-group.

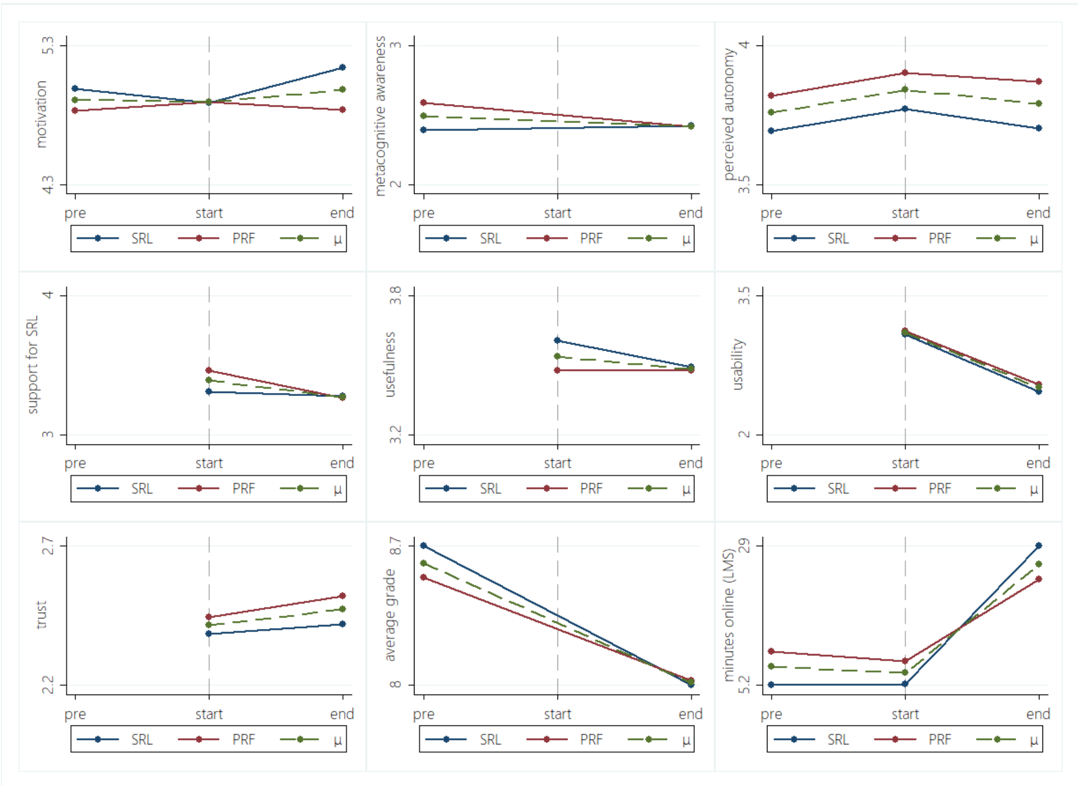


Figure F. 2: Plots of the measurements across different phases of the experiment by dashboard-group.