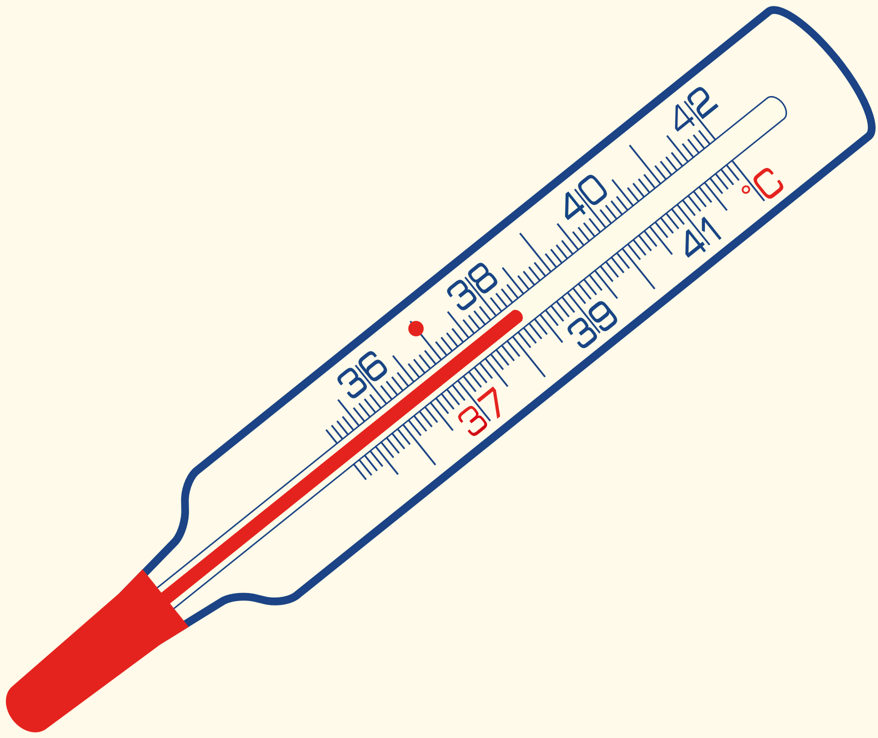


STUDIA PAEDAGOGICA

VOLUME 28 / NUMBER 3 / YEAR 2023

LEARNING ANALYTICS TO STUDY AND SUPPORT SELF-REGULATED LEARNING

MASARYK UNIVERSITY PRESS

STUDIA PAEDAGOGICA
VOLUME 28 / NUMBER 3 / YEAR 2023

LEARNING ANALYTICS TO STUDY AND SUPPORT SELF-REGULATED LEARNING

SPECIAL ISSUE

LIBOR JUHAŇÁK
SREĆKO JOKSIMOVIĆ
DIRK IFENTHALER
EDITORS

2023

MASARYK
UNIVERSITY

Studia paedagogica cooperates with the Emerging Researchers' Group
of the European Educational Research Association (EERA).

Citation: *Studia paedagogica* 28, 3, 2023

ISSN 1803-7437 (print)
ISSN 2336-4521 (online)

CONTENTS

EDITORIAL

LEARNING ANALYTICS TO STUDY AND SUPPORT SELF-REGULATED LEARNING

The theme of the current issue of Studia paedagogica is learning analytics and its potential for research and support of self-regulated learning.

Over the past ten to fifteen years, the field of learning analytics has undergone tremendous development, capturing the interest of an expanding community of researchers from diverse fields including educational sciences, psychology, computer science, and others. As the widely accepted definition suggests, learning analytics is concerned with the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purpose of understanding and optimizing learning and the environments in which it occurs (Ifenthaler & Yau, 2020; Joksimović et al., 2019; Juhaňák & Zounek, 2019; Siemens, 2013). Similarly, self-regulated learning has received increasing attention in educational research. Especially in the last two decades, extensive theoretical and conceptual development has taken place, and several distinct definitions and models of self-regulated learning have been proposed and developed (Boekaerts et al., 2000; Panadero, 2017; Zimmerman & Schunk, 2011).

Existing research on self-regulated learning has repeatedly demonstrated the importance and impact of self-regulation on student performance and learning outcomes, as well as its implications for student well-being. Students engaging in self-regulated learning are able to manage their own learning and adapt their learning behaviors effectively, and they exhibit more positive motivational characteristics, leading to their better performance on learning tasks and academic success in general (Boekaerts et al., 2000; McInerney et al., 2012). At the same time, research on self-regulated learning has become more prominent in recent years in relation to different online learning environments, as several studies have found that students' ability to self-regulate their learning differs in online versus traditional settings (Broadbent & Poon, 2015; Sedrakyan et al., 2018; Wong et al., 2019). While the use of learning analytics to study self-regulated learning is still in its early stages, researchers in the field have already begun to systematically explore what indicators of self-regulation can be tracked in online learning environments and which computational and analytic tools can be used to analyze these data, with the goal of accurately measuring self-regulated learning in online learning environments and better understanding the various aspects of students' learning behaviors (Viberg et al., 2020).

Still, despite the increasing number of studies applying learning analytics to study and support self-regulated learning (Park et al., 2023), many questions remain unanswered. Therefore, this monothematic issue provides a space for researchers to present their research and discuss current issues and questions related to self-regulated learning research.

We are pleased that the articles in this issue approach the intersection of learning analytics and self-regulated learning from a variety of directions, highlighting the breadth of the topic and the range of possible research approaches that can be used to study self-regulated learning.

Natalie Borter, in her study *Differential Effects of Additional Formative Assessments on Student Learning Behaviors and Outcomes*, adopts a quasi-experimental approach to examine whether additional formative assessments completed by students lead to improved learning outcomes and changes in students' self-regulated learning behaviors. The study, conducted in a real-world blended learning environment, employed a learning analytics approach by combining both behavioral and self-reported data and using several analytical techniques such as exploratory factor analysis and cluster analysis. The results support the notion that the additional formative assessments lead to improved learning outcomes, but at the same time, suggest that the change in students' self-regulated learning behaviors based on their participation in additional formative assessments can be both positive and negative.

As an integral part of self-regulated learning, Libor Juhaňák, Karla Brücknerová, Barbora Nekardová, and Jiří Zounek focused on goal setting and goal orientation in their study *Goal Setting and Goal Orientation as Predictors of Learning Satisfaction and Online Learning Behavior in Higher Education Blended Courses*. Using a relatively large sample of hundreds of students and dozens of different blended courses, and employing multilevel modeling, the authors examine the relationship between goal setting and goal orientation and student behavior in the online learning environment, as well as the effects of these two measures on student learning satisfaction.

Another example of using learning analytics to study online learning behavior is presented by Ricardo Santos and Roberto Henriques in the article *Decoding Student Success in Higher Education: A Comparative Study on Learning Strategies of Undergraduate and Graduate Students*. Similar to Juhaňák et al.'s study, Santos and Henriques analyze student behavioral data extracted from a learning management system (LMS); however, they focus primarily on uncovering various self-regulated learning behaviors and learning strategies adopted by students in the LMS. Using k-means clustering, the authors were able to distinguish five different learning strategy profiles among the undergraduate and graduate students. Further, the authors examine how the identified learning strategy profiles relate to student learning outcomes.

Mattias Wickberg Hugerth, Jalal Nouri, and Anna Åkerfeldt use a different methodological approach in their study *"I Should, but I Don't Feel Like It": Overcoming Obstacles in Upper Secondary Students' Self-Regulation Using Learning Analytics.* The authors explore how learning analytics can be helpful for students to support their self-regulation, paying primary attention to the challenges experienced by students in the process of self-regulation and focusing on the data and information students need to better regulate their own learning. Based on the findings, the authors suggest that learning analytics systems need to be designed with students' self-regulation needs in mind, incorporating support for scaffolding self-regulation, while taking into account that support for the planning and performance phases appears to be most critical.

Nicol Dostálová and Lukáš Plch contributed to this special issue with a review study entitled *A Scoping Review of Web-Cam Eye Tracking in Learning and Education.* The study builds on existing research in the area of multimodal learning analytics, where different types of technologies are employed to capture and analyze student behavioral data, including eye-tracking technology, which is used to study students' gaze and eye movements. The authors focus specifically on webcam eye-tracking, which can be considered a relatively new technology, and show that webcam eye-tracking has great potential in self-regulated learning research and educational research in general.

The issue concludes with the emerging researcher section, which contains Barbora Al Ajeilat Kousalová's study *Vocabulary Learning Strategies, Self-Regulated Learning, and Learners' Outcomes in Primary School Pair Work.* Al Ajeilat Kousalová addresses a research gap identified in previous approaches to studying vocabulary learning strategies used by learners during pair work in the context of foreign language learning and, inspired by the research area of multimodal learning analytics, employs a qualitative design study approach. By analyzing audio and video recordings, the author was able to uncover different patterns of vocabulary learning strategies and distinguish between successful and unsuccessful strategy applications.

We believe that this special issue of Studia paedagogica has the potential to contribute to the current discussion on the use of learning analytics in self-regulated learning research and to enrich the research practices of using learning analytics in self-regulated learning research, thereby enabling higher education students to be more effectively and qualitatively supported in developing their ability to regulate their own learning.

Libor Juhaňák, Srećko Joksimović, and Dirk Ifenthaler
Editors

References

Boekaerts, M., Pintrich, P. R., & Zeidner, M. (2000). *Handbook of Self-Regulation*. Academic Press.

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, *68*(4), 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Joksimović, S., Kovanović, V., & Dawson, S. (2019). The Journey of Learning Analytics. *HERDSA Review of Higher Education*, *6*, 37–63.

Juhaňák, L., & Zounek, J. (2019). Learning analytics: Challenges and opportunities of using data analysis in education. In P. Ordóñez de Pablos, M. Lytras, X. Zhang, & K. Chui (Ed.), *Opening up education for inclusivity across digital economies and societies* (pp. 52–72). IGI Global.

Panadero, E. (2017). A Review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(422), 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Park, E., Ifenthaler, D., & Clariana, R. (2023). Adaptive or adapted to: Sequence and reflexive thematic analysis to understand learners‘ self-regulated learning in an adaptive learning analytics dashboard. *British Journal of Educational Technology*, *54*(1), 98–125. https://doi.org/10.1111/bjet.13287

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. (2018). Linking learning behavior learning analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, *107*, 1–15. https://doi.org/10.1016/j.chb.2018.05.004

Schumacher, C., & Ifenthaler, D. (2021). Investigating prompts for supporting students' self-regulation – A remaining challenge for learning analytics approaches? *The Internet and Higher Education*, *49*, 100791. https://doi.org/10.1016/j.iheduc.2020.100791

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Viberg, O., Khalil, M., & Baars, M. (2020). *Self-regulated learning and learning analytics in online learning environments: A review of empirical research*. The 10th International Learning Analytics and Knowledge Conference, March 23–27, 2020, Frankfurt, Germany.

Winne, P. (2017). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, Wise, A., & D. Gasevic (Eds.), *Handbook of Learning Analytics* (pp. 241–249). Society for Learning Analytics Research.

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G.-J., &Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction*, *35*(4–5), 356–373. https://doi.org/10.1080/10447318.2018.1543084

Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of Self-Regulation of Learning and Performance*. Routledge.

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-1

DIFFERENTIAL EFFECTS OF ADDITIONAL FORMATIVE ASSESSMENTS ON STUDENT LEARNING BEHAVIORS AND OUTCOMES

Natalie Borter[a,b]

[a] Institute for Psychology, University of Bern, Switzerland
[b] School of Business, HSLU Luzern, Switzerland

ABSTRACT

It is well-established that formative assessments with accompanying feedback can enhance learning. However, the degree to which additional formative assessments on the same material further improve learning outcomes remains an open research question. Moreover, it is unclear whether providing additional formative assessments impacts self-regulated learning behavior, and if the benefits of such assessments depend on students' self-regulated learning behavior. The current study, conducted in a real-world blended learning setting and using a Learning Analytics approach, compares 154 students who completed additional formative assessments with 154 students who did not. The results indicate that the additional formative assessments led to an improvement in learning outcomes, but also had both positive and negative effects on students' self-regulated learning behavior. Students who completed additional formative assessments performed better on the assessments but reported lower levels of subjective comprehension and devoted more time to completing exercises. Simultaneously, they devoted less effort to additional learning activities (additional investment), such as class preparation and post-processing. Furthermore, the impact of additional formative assessments on learning success depended on students' self-regulated learning behavior. It was primarily the students who invested above-average time during formative assessments (time investment) who benefited from the additional exercises. Cluster analysis revealed that high-effort students (those with above-average time investment and above-average additional investment) gained the most from the extra exercises. In contrast, low-effort students and those who achieved high performance with relatively low effort (efficient students) did not benefit from additional formative assessments. In conclusion, providing students with additional formative assessments can enhance learning, but it should be done with caution as it can alter self-regulated learning behavior in both positive and negative ways, and not all students may benefit from it equally.

KEYWORDS

(indirect) testing effect; formative assessments; feedback; self-regulated learning behavior; individual differences

CORRESPONDING AUTHOR
Natalie Borter, Institute for Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland
e-mail: natalie.borter@unibe.ch

Introduction

The testing effect, a well-established learning technique (Jensen et al., 2020), denotes the enhanced learning success observed when students actively engage with learned material, rather than relying on passive repetition or memorization (Schwieren et al., 2017). The effect is typically quantified by comparing the post-learning performance of learners who participated in active retrieval during the learning phase to those who did not engage in such practices. An instance of active retrieval is solving formative assessments, wherein learners apply their knowledge to solve problems (Boston, 2002).

A significant testing effect was demonstrated in both experimental and applied settings (Lamotte et al., 2021; Schwieren et al., 2017). It is evident in tasks ranging from relatively simple ones, like vocabulary memorization, to more complex tasks that involve applying theoretical knowledge to novel situations (Schwieren et al., 2017). The effect occurs when exercises during the learning phase are identical to those measuring learning success (Carpenter, 2011; Eriksson, et al., 2011; Karpicke & Roediger, 2007) and when non-identical exercises covering the same material are employed (Batsell et al., 2017; Foss & Pirozzolo, 2017; Francis et al., 2020; Jensen et al., 2020; McDaniel et al., 2013).

One open research question regarding the testing effect concerns whether multiple testing instances result in greater learning success than administering a single test. Multiple testing can include either the repeated use of the same test or the utilization of different tests covering the same content (Yang et al., 2021). A meta-analytic review conducted by Adesope et al. (2017) did not identify a significant difference between the testing effects of multiple tests versus a single test on the same material. However, a more recent review by Yang et al. (2021) discovered that, in applied contexts such as classrooms, the testing effect was more pronounced when the same test (or a similar test with the same content) was administered repeatedly. Additional empirical research is required to determine whether multiple tests on the same content yield a more pronounced testing effect (Yang et al., 2021).

Assessing the effectiveness of additional exercises on identical content is crucial for practical implementation, particularly given the laborious task of generating such exercises. Moreover, the provision of additional learning materials may not always lead to increased learning, and can even have a negative impact on learning in some cases by inducing cognitive overload or stress (Kossen & Ooi, 2021).

As additional formative assessments on the same content have the potential to influence self-regulated learning both positively and negatively, a learning analytics approach was employed to investigate the impact of additional formative assessments on self-regulated learning behavior. This approach involves the collection and analysis of data on students' learning behavior and progress, to enhance learning and teaching (Chatti et al., 2013; Ifenthaler, 2015; Leitner et al., 2017). The utilization of such data has attracted attention in the field of self-regulated learning, as it enables the monitoring of the learner's holistic action without interference in the process (Winne & Baker, 2013).

Prior research suggests that formative assessments have a positive effect on self-regulated study time allocation and monitoring (Clariana & Park, 2021; Fernandez & Jamet, 2017; Perry & Winne, 2006; Soderstrom & Bjork, 2014; Yang et al., 2017). Engagement in solving exercises enhanced students' monitoring of their knowledge, leading to a reduction in the overestimation of their knowledge and an increase in time allocated for studying (Soderstrom & Bjork, 2014). The positive impact of exercise solving on learning success was partially mediated by improved monitoring and learning behavior (Fernandez & Jamet, 2017). This is because additional formative assessments can provide feedback on learning status, aiding learners in metacognitive control, and adapting their self-regulated learning behavior (Clariana & Park, 2021; Perry & Winne, 2006).

Furthermore, the efficacy of additional formative assessments likely depends on individual differences in students' characteristics (Bertilsson et al., 2021). Prior knowledge and experience are also important factors, as students with more prior knowledge tend to benefit more from the testing effect than those with less prior knowledge (Cogliano, et al., 2019; Francis et al., 2020). According to the elaborative retrieval hypothesis, mental effort during recall predicts the magnitude of the testing effect, and practicing with exercises that are challenging but within the learner's abilities can enhance the effect (Carpenter, et al., 2009; Greving et al., 2020; Minear et al., 2018). In addition, students who already possess effective learning strategies may not benefit as much from formative assessments as those who lack such strategies (Robey, 2019). This is because they have already achieved high learning success even without formative assessments, and the effect of testing may not significantly enhance their learning outcomes.

Not only the number but also the timing of formative assessments can influence their effectiveness (Karpicke & Bauernschmidt, 2011). In line with current research on spaced learning (Greene, 2008; Jost et al., 2021), evenly distributed learning throughout the semester is more effective in promoting learning success than cramming right before a test (Adesope et al., 2017; Karpicke & Bauernschmidt, 2011).

In summary, the positive impact of testing on learning can be influenced by multiple factors, including individual student characteristics (e.g., hope of success, prior knowledge, investment, timing). It is plausible to assume that students with limited prior knowledge, low motivation, low investment, frequent incorrect responses, last-minute study habits, and shallow feedback processing may not benefit as much from extra exercises with feedback as other students. Therefore, various types of students may exist, and some may not experience the same benefits from additional formative assessments.

In this study, a clustering based on the data collected through the learning analytics approach was employed to examine whether the impact of solving additional formative assessments on end-of-semester knowledge test performance is influenced by student self-regulated learning behavior. Cluster analysis is a method that divides students into groups based on similarities within a cluster and dissimilarities between clusters (Dalmaijer et al., 2021; Shin & Shim, 2021).

To sum up, although there is significant evidence supporting the idea that formative assessments improve learning outcomes, it is still uncertain whether administering additional assessments on the same content leads to a more substantial effect and how this practice influences self-regulated learning behavior. Additionally, studies indicate that the benefits of additional formative assessments may differ based on students' self-regulated learning characteristics. As a result, this study aims to examine the following three hypotheses:

1. Administering additional formative assessments of the same content leads to higher performance on an end-of-course knowledge test.
2. Administering additional formative assessments of the same content influences self-regulated learning behavior.
3. The relationship between administering additional formative assessments and learning success depends on students' self-regulated learning behavior.

1 Method

1.1 Participants

To be included in the study, students had to take both the prior knowledge test at the beginning of the semester and the end-of-semester knowledge test. They also had to complete at least four of the five formative assessments (not including the additional formative assessments). In addition, students with very low performance in the formative assessments or in the end-of--semester knowledge test (clear outliners) were excluded (three students). Of the 276 students enrolled in the course in 2020, a total of 194 met the inclusion criteria, while in 2021 a total of 166 of the 234 students enrolled met the inclusion criteria. This represents approximately 70% of all enrolled students in both courses. Out of the total 360 students evaluated (194 + 166), 324 had no missing values and 36 students were missing one of the five formative assessments. To address this, missing values for the 36 students (21 from 2020's 194 and 15 from 2021's 166) were imputed using the mice function (van Burren & Groothuis-Oudshoorn, 2011) employing a predictive mean matching approach.

The aim of the current study was to compare students solving most additional exercises with students not solving any additional exercises. Thus, students without access to any extra exercises were labeled "non-solvers" while those who finished at least four out of the five additional assessments were referred to as "solvers". From the 2021 cohort, 12 students who completed fewer than four additional formative assessments were excluded. Accordingly, 194 students were identified as "non-solvers" and 154 as "solvers". The study was approved by the local ethics committee.

1.2 Procedure

The mandatory "Psychological Diagnostics" course for master's students in psychology focused on complex methodological content such as equivalence analysis and item response theory. Students were permitted to choose the learning approach they found most suitable for the specific learning situation, as their learning behavior was not explicitly manipulated. Consequently, this study examined the impact of extra exercises compared to any other learning approach, including no learning at all.

Due to ethical concerns, students were not randomly assigned to groups. Instead, this study adopted a quasi-experimental approach, evaluating the same course across two successive years, 2020 and 2021. The only variation introduced between the two years was the inclusion of optional additional formative assessments in 2021. The additional formative assessment consisted of new exercises covering the same content as the initial assessment. All other elements of the course, including initial exercises, podcasts, literature, and instructions, were kept consistent across both years. Participation in the study was voluntary. All students had access to the standard learning materials. However, those who volunteered to participate in the study received additional feedback post-exam: z-standardized values of all variables specified in the method section, enabling them to compare their performance with that of the entire group. Non-participants did not receive this supplementary information. All data were pseudo-anonymized using unique pseudonyms. Based on the pseudonyms, there were no students who attended the course in both years, 2020 and 2021.

A blended learning approach was employed in both years, allowing students to engage with course material at their own pace and participate in timely online discussions. The curriculum included 12 weekly podcast lectures, each 90 minutes long, a prior knowledge test, and five biweekly formative assessments covering two lectures each. Data collection occurred during the formative assessments.

Upon completing each exercise, students received immediate feedback and suggestions for supplementary resources, including relevant literature, lecture slides, podcast excerpts, and additional links or references. Students could ask further questions on the feedback page, which were addressed in a forum or, if necessary, through scheduled online discussions.

Students were advised to adhere to a one-week submission window for formative assessments, facilitating an even distribution of their learning throughout the semester. This structure was consistent over both years; however, the frequency of formative assessments varied. In 2020, formative assessments were assigned every two weeks, whereas in 2021, with the addition of supplementary formative assessments, they occurred weekly. Despite the change in frequency, the exercises, including the additional assessments, could be repeated and remained accessible to students until the final exam.

Two weeks before the final exam, a comprehensive end-of-semester knowledge test with new exercises covering the entire course content was administered. All exercises and self-reports were designed and hosted using Qualtrics (Qualtrics, Provo, UT). Response latency, accuracy, and time spent on feedback pages were assessed. Links to the Qualtrics questionnaires were embedded into the ILIAS learning management system, where all essential learning resources, such as podcasts and literature, were made available to students.

To identify clusters of response behavior, data from the initial formative assessments were utilized. Additional formative assessments were not included because they were only accessible to the solvers.

In the next section, the behavioral and self-reported learning analytics data gathered are presented. Self-reports were used to capture perceptions such as subjective knowledge, subjective investment, and subjective importance.

Furthermore, for variables over which I did not have full control, as I intentionally permitted students to learn in ways they found most suitable for the specific learning situation, such as downloading or printing materials; self-reports were employed. The drawbacks of self-reports were minimized by employing pseudonymization to reduce the impact of social desirability and by using precise questions to decrease the likelihood of recall errors. For our main emphasis, the formative assessments, offline solutions were not allowed, so behavioral data were utilized.

1.3 Behavioral data

Prior knowledge

Prior knowledge was assessed with 19 multiple-choice exercises. The test contained mainly theoretical exercises and calculations concerning real--world applications of the knowledge acquired during the bachelor's program (e.g., reliability, validity). One sum score was built for the 13 exercises covering theoretical exercises and one for the six exercises covering calculations.

Performance in the formative assessments

For each of the five formative assessments, covering different content such as item response theory, confirmatory factor analysis, equivalence analysis, and criterion-referenced testing, a sum score was built. The number of exercises per formative assessment ranged from 10 to 24. The exercises consisted primarily of multiple-choice exercises, in which the theoretical knowledge acquired from the podcast was applied to concrete situations. The sum scores of the five assessments were highly related, with a Cronbach's alpha of 0.76. The sum of all five formative assessments was used for further analyses.

End-of-semester knowledge test

The dependent variable of the current study was the performance in the end-of-semester knowledge test. It consisted of 22 exercises. In contrast to the exam, the knowledge test covered only the content of the formative assessments and was identical in 2020 and 2021. The correlation between the end-of-semester knowledge test and the final grades was comparable in both cohorts – 2020 ($r = 0.52$, $p < 0.05$) and 2021 ($r = 0.57$, $p < 0.03$).

Time investment

Response latency was recorded for each task and the feedback page. To reduce the effect of strong outliers, for each time measure, all values greater than the 95th percentile were trimmed to the 95th percentile. As the response latencies were still highly right-skewed, each time measure was logarithmized. Thereafter, all response latencies were z-standardized and the first strong principal component of the response latencies on the exercise page (explaining 43% of the variance), and the first strong principal component of the response latencies on the feedback page (explaining 46% of the variance) were extracted and used for further analyses.

Number of completions

The "completions initial exercises" variable was computed for the initial formative assessments, considering the count of completions for both identical and distinct formative assessments. For the variable "completions overall" the total number of completions (also including the additional formative assessments) was calculated. The number of completions overall was categorized into six groups: 1 = up to ten completions; 2 = 11–15 completions; 3 = 16–20 completions; 4 = 21–25 completions; 5 = 26–30 completions; 6 = more than 30 completions.

Questions for the forum

Across all exercises of the formative assessments, the frequency with which questions were posed by students on the feedback page was recorded. This variable was highly right-skewed, and therefore the values were logarithmized.

On time / regularity

As a measure of regularity, it was counted how often students finished the formative assessments during the recommended one-week submission window.

1.4 Self-reported data

Subjective knowledge

At the beginning of each formative assessment, students rated their subjective understanding of the content covered in the respective exercise session on a five-point scale (1 = I don't know this concept, 2 = I don't understand this concept well, 3 = I understand this concept less well, 4 = I understand this concept well, 5 = I understand this concept very well). First, the average of these ratings was taken for each formative assessment, and then the first strong principal component (explaining 65% of the variance) was extracted from the averaged ratings across all five formative assessments (excluding the additional formative assessment) and used for further analyses.

Subjective investment

After each formative assessment, students rated on a four-point scale their effort level in attempting to complete the exercises to the best of their ability (1 = I didn't try hard, 2 = I tried a little, 3 = I tried a lot, 4 = I tried hard). The average of these ratings was calculated and used for further analyses.

Lectures

At the beginning of each formative assessment, students indicated whether they had listened to the podcasts of the two lectures covered in the formative assessment (1 = I listened to neither of the two podcasts, 2 = I listened to parts of both podcasts, 3 = I listened to at least one of the podcasts completely, 4 = Yes, I listened to both podcasts completely). The mean value of this variable, computed across the five formative assessments, was utilized for subsequent analyses.

Reading forum

At the beginning of the end-of-semester knowledge test, students indicated on a three-point scale whether they had read the forum posts before (0 = I never read the forum, 1 = I read the forum only when I had questions, 2 = I read all forum posts at least once).

Compulsory literature

At the outset of each formative assessment, students specified their engagement with the mandatory literature, which, in combination with lectures, served as the foundational preparation for the assessment: (1) indicated they read at least some part of the mandatory literature, while (0) denoted they did not engage with it.

The mean value of this variable, computed across the five formative assessments, was utilized for subsequent analyses.

Relevance of content

On a four-point scale (false, somewhat false, somewhat true, true) students responded to the following questions about the content of the course:

- I find "Psychological Diagnostics" interesting.
- I think my knowledge of "Psychological Diagnostics" will be useful to me in the future.
- I think it is important to learn "Psychological Diagnostics" in psychology education.

The average of the three items was used for further analyses.

Learning hours during semester holidays

The students reported the number of hours they dedicated to studying for the exam following the final lecture of the semester. As data were highly right-skewed, they were logarithmized.

2 Results

All analyses were conducted in R version 3.6.1 (R Core Team, 2021).

2.1 Descriptive statistics

Given the quasi-experimental design of the study, it was crucial to establish that there were no initial differences between the students from 2020 and 2021 in terms of "prior knowledge" and "subjective relevance of the content" at the beginning of the course. To compare the means for these measures, an equivalence analysis was conducted (Bentler & Satorra, 2010). For "prior knowledge," a two-factor solution (theory and calculations) was compared to a one-factor solution. The significant Chi-square difference ($\Delta\chi^2(1) = 67.70$, $p < 0.001$) indicated that the two-factor model ($\chi^2(151) = 178.43$, $p = 0.063$, CFI = 0.934, RMSEA = 0.022, SRMR = 0.047) provided a better fit to the data than the one-factor model ($\chi^2(152) = 237.07$, $p < 0.001$, CFI = 0.796, RMSEA = 0.039, SRMR = 0.055). Consequently, prior knowledge is more accurately represented by a two-factor solution. The two factors, theory and calculations, were correlated ($r = 0.53$, $p < 0.01$).

A measurement invariance analysis using lavaan (Rosseel, 2012) confirmed scalar equivalence (configural vs. metric fit: $\Delta\chi^2(17) = 17.34$, $p = 0.43$; scalar vs. metric fit: $\Delta\chi^2(17) = 16.72$, $p = 0.47$; scalar model fit $\chi^2(336) = 367.15$, $p = 0.12$, CFI = 0.926, RMSEA = 0.023, SRMR = 0.066), allowing for comparison of the means between the two groups (2020 vs. 2021 course). Accordingly, prior knowledge in calculations was measured using the sum score of all items loading on the calculations factor, while the sum score of all items loading on the theory factor was employed as a measure of theoretical prior knowledge.

Non-solvers differed from solvers in both scales of prior knowledge, calculations ($t(358) = -2.14$, $p < 0.05$; non-solvers: $M = 4.62$, $SD = 1.45$; solvers: $M = 4.98$, $SD = 1.20$,), and theory ($t(286.03) = -2.15$, $p < 0.05$; non-solvers: $M = 7.77$, $SD = 1.39$; solvers: $M = 8.14$, $SD = 1.76$) as well as in the subjective relevance of the content ($t(325.37) = -2.30$, $p < -0.05$; non-solvers: $M = 2.96$, $SD = 0.64$; solvers: $M = 3.12$, $SD = 0.66$). To ensure comparability of prior knowledge and subjective relevance of the content between solvers and non-solvers, a matching approach was employed. The matching was conducted using the function matchit from the MatchIt package, with a nearest neighbor method, distance logit, and an "ATT" estimate (Pishgar et al., 2021). The 194 "non-solvers" were matched to the 154 "solvers". The matched samples, each consisting of 154 students, did not differ in the prior knowledge scale calculations (non-solvers: $M = 4.86$, $SD = 1.35$ versus solvers: $M = 4.98$, $SD = 1.20$, $p = 0.40$) and theory (non-solvers: $M = 7.97$, $SD = 1.39$ versus solvers: $M = 8.14$, $SD = 1.76$, $p = 0.34$) nor in subjective relevance

($t(325.37) = -2.30, p = .67$; non-solvers: $M = 3.09, SD = 0.59$; solvers: $M = 3.12$, $SD = 0.66$). Subsequent analyses were carried out exclusively on the matched samples.

In Table 1, mean (standard deviation), skewness and kurtosis of the variables considered in the study are provided for the entire sample ($N = 308$), for the solvers ($N = 154$) and for the non-solvers ($N = 154$). The skewness of all variables was between −3 and 3 and the kurtosis between 10 and −10. According to Kline (2011), this indicates approximately normally distributed variables. Parametric methods were applied in this study as they are generally robust to scale assumption violations, especially when likert scales have seven or more categories (Norman, 2010; Dolan, 1994; Robitzsch, 2020). The majority of our ordinal variables had seven or more categories due to aggregation. The sole exception, "reading forum" with three categories, showed negligible differences between Pearson and Spearman correlations (maximum difference: 0.0165; average difference: < 0.0016). Hence, parametric methods were used.

2.2 Solving additional formative assessments, self-regulated learning behavior and learning success

With a *t*-test I investigated whether the solvers performed better in the end-of-semester knowledge test than the non-solvers. As shown in Table 1, solvers reached a higher performance in the knowledge test than non-solvers ($t(305.06) = -2.92, p < 0.01, d = 0.33$), confirming the first hypothesis.

Consistent with the hypothesis, the findings indicate that engagement with additional formative assessments significantly influences self-regulated learning behavior (see Table 1). Specifically, it was observed that those who solved these assessments demonstrated enhanced performance, invested more time in the completion of exercises, and posed fewer questions about those exercises.

Albeit not statistically significant, in tendency, solvers demonstrated a lower level of subjective understanding and less dedication to reading the mandatory literature than the non-solvers.

When analyzing the "total completions", which is the total number of completed exercises from both the initial and the additional formative assessments (where multiple attempts were possible), solvers completed significantly more exercises. This was expected since they had access to both initial and additional assessments.

However, when considering the "initial completions" (which both groups could attempt multiple times), solvers completed fewer exercises than non-solvers. This suggests that while having access to additional assessments led to more completions overall, it resulted in fewer completions of the initial assessments that were available to everyone.

Table 1
Descriptive statistics for entire sample, non-solvers and solvers as well as correlations with end-of-semester knowledge test (r).

	Mean (*SD*)	Skew	Kurtosis	Non-solvers	Solvers	*p*	*r*
Knowledge test	17.83 (2.82)	−0.95	1.00	17.36	18.29	<0.01	–
Prior knowledge	12.98 (2.25)	−0.28	−0.14	12.83	13.12	0.25	0.33***
Formative assessments	48.21 (5.96)	−0.86	0.91	47.28	49.13	<0.01	0.51***
Completions initial exercises	12.84 (6.19)	1.65	3.69	13.64	12.04	<0.05	0.08
Completions overall	2.93 (1.54)	0.59	−0.71	2.23	3.63	<0.001	0.18**
Subjective understanding	0.03 (1.02)	−1.01	3.47	0.14	−0.08	0.06	0.28***
Time investment on exercises	0.07 (2.45)	−1.74	−6.57	−0.35	0.59	<0.01	0.10
Time investment on feedback	−0.02 (2.59)	−0.54	0.40	−0.04	−0.01	0.93	0.04
Subjective investment	3.16 (0.49)	−0.16	−0.31	3.20	3.13	0.17	0.15*
Completing on time	0.56 (0.37)	−0.17	−1.53	0.58	0.54	0.45	0.15*
Lectures	3.94 (0.15)	−2.30	3.80	3.94	3.94	0.64	0.24***
Read forum	0.84 (0.69)	0.21	−0.92	0.88	0.81	0.32	0.10
Compulsory literature	0.51 (0.40)	−0.06	−1.58	0.55	0.47	0.09	−0.06
Questions	0.18 (0.27)	2.12	4.86	0.23	0.13	<0.01	0.01
Relevance of content	3.11 (0.62)	−0.57	−0.02	3.09	3.12	0.67	0.22***
Learning hours after course	3.08 (0.99)	−1.21	2.24	3.15	3.01	0.22	0.06

Note. r – correlation between the corresponding variable and performance in the end-of-semester knowledge test; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. The remaining correlations were not significant ($p > 0.10$). For subjective understanding and the two-time investment measures, scores on the first principal component are reported.

2.3 Students' characteristics, solving additional formative assessments and learning success

To identify meaningful clusters of self-regulated learning behavior, understanding the interrelations of the learning variables detailed in the Method section was crucial. An exploratory factor analysis was conducted to reduce the variables to a few interpretable factors. By decreasing the number of variables in the model, the cluster analysis can more effectively detect clusters within the dataset (Dalmaijer et al., 2021). The z-standardized variables were inputted into the fa.parallel function from the psych package (Revelle, 2022), resulting in a three-factor solution that best described the correlations between the thirteen manifest variables. The factor solution, following an oblimin rotation, is presented in Table 2.

Table 2

Standardized loadings of the measures on the three factors extracted by exploratory factor analysis with oblimin rotation

Variable	Performance	Time investment	Additional investment	h^2
Formative assessments	**0.88**	0.06	−0.06	0.80
Subjective understanding	**0.52**	−0.10	0.20	0.31
Time investment exercises	0.26	**0.65**	−0.04	0.57
Time investment feedback	−0.08	**0.85**	0.03	0.71
Subjective investment	0.25	**0.38**	**0.33**	0.44
Completing on time	0.28	−0.23	**0.41**	0.25
Lectures	**0.37**	0.19	0.12	0.25
Read forum	−0.04	0.02	**0.42**	0.18
Compulsory literature	−0.13	0.09	**0.53**	0.30
Questions	−0.07	0.04	**0.43**	0.19
Prior knowledge	**0.49**	−0.06	−0.05	0.23
Relevance of content	**0.35**	−0.10	0.10	0.13
Learning hours after course	−0.10	0.12	**0.39**	0.18
R^2	0.14	0.12	0.09	
Proportion R^2	0.41	0.33	0.26	

Note. R^2 – variance explained by the corresponding factor, h^2 – explained variance of the corresponding measurement, loadings of at least 0.30 are in bold.

To comprehend the three factors, they will be described based on the measures exhibiting the highest loadings (Table 2). The first factor is associated with performance, as evidenced by substantial loadings of performance in formative assessments, subjective understanding, and prior knowledge. The second factor is connected to time investment, which includes time spent on exercise pages and feedback pages. This factor is related to the time investment in content learning, a critical self-regulation skill identified by Kim et al. (2018) to effort regulation (Baker et al., 2020) or organization (Mega et al., 2014).

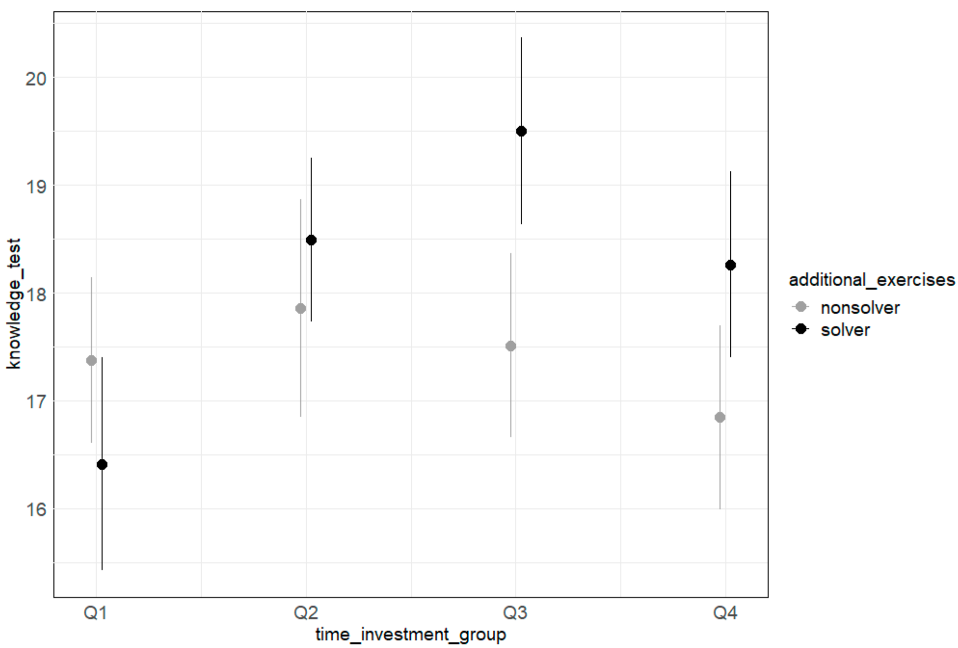
The third factor pertains to additional investment, as demonstrated by engagement in reading the literature, posing questions, reading the forum, dedicating learning hours during the semester break and timely completion of exercises. Accordingly, additional investment is a combination of help seeking (Kim et al., 2018), time management (Kim et al., 2018; Li et al., 2018) and investment in content learning (Kim et al., 2018).

The three factors were slightly correlated (performance and time investment $r = 0.27, p < 0.001$; performance and additional investment $r = 0.14, p < 0.05$; time investment and additional investment $r = 0.23, p < 0.001$) and together explained 35% of the variance. For further analyses, factor scores extracted using the regression method were used (DiStefano et al., 2009). Solvers scored higher on the performance factor ($t(305.35) = -2.03, p < 0.05, d = 0.013$) and lower on the additional investment factor ($t(305.37) = 3.55, p < 0.001, d = 0.04$) while there was no difference in scores on the time investment factor ($t(292-76) = -0.99, p = 0.32, d = 0.003$).

To examine whether the effect of additional formative assessments depends on students' self-regulated learning behavior, two approaches were employed. First, each of the three extracted factors was divided into four equal groups (quartiles) and the dependency of the effect of solving additional formative assessments on that split variable was investigated for each factor. Second, a cluster analysis was conducted across all three factors, and the dependency of solving additional formative assessments on cluster membership was examined.

In the first approach, a two-way ANOVA was conducted for each factor group (quartiles), with factor group membership and solving additional formative assessments as the between-subject factors and performance in the end-of-semester knowledge test as the dependent variable. A significant interaction would indicate that the effect of additional formative assessments on performance in the end-of-semester knowledge test depends on student characteristics. The interaction term was not significant for the performance factor ($F(3, 300) = 0.41, p = 0.75, \eta^2 = 0.003$) or the additional investment factor ($F(3, 300) = 0.23, p = 0.87, \eta^2 = 0.002$); however, it was significant for the time investment factor ($F(3, 300) = 4.14, p < 0.01, \eta^2 = 0.04$).

Figure 1 displays the interaction between the time investment group and solving additional formative assessments. In the lowest time investment quartile, solving additional formative assessments was associated with slightly lower performance in the end-of-semester knowledge test, whereas in all other quartiles, it was associated with higher performance. The performance difference in the end-of-semester knowledge test between non-solvers and solvers was −0.96 (p = 0.13) for the first quartile, 0.63 (p = 0.32) for the second quartile, 1.99 (p < 0.01) for the third quartile, and 1.42 (p < 0.05) for the fourth quartile. However, when applying a Bonferroni-corrected alpha level of 0.0125, the difference in the fourth quartile was no longer statistically significant. Overall, solving additional formative assessments appeared to be more beneficial for students who invested more time in solving the exercises.



Note. Q1 = first quartile, Q2 = second quartile, Q3 = third quartile, Q4 = fourth quartile; means and standard deviations are displayed.

Figure 1
Interaction between the completion of additional formative assessments and students' quartile ranking in time investment, in relation to performance on the end-of-semester knowledge test

It is important to note that solvers and non-solvers were not equally distributed across the four time investment groups ($\chi^2(3) = 10.44$, $p < 0.05$). Fewer solvers ($n = 29$) than non-solvers ($n = 48$) were in the first quartile, and more solvers ($n = 49$) than non-solvers ($n = 28$) were in the second quartile. In the other two groups, solvers and non-solvers were similarly distributed (either $n = 38$ or $n = 39$).

In the second approach, which is based on all three factors (performance, time investment, additional investment), a k-means cluster analysis was conducted to identify distinct student types. Initially, the number of clusters was determined using the NbClust function (Charrad et al., 2014), followed by the execution of the k-means cluster analysis using the stats package (R Core Team, 2021). The NbClust function helps determine the number of clusters in a dataset by evaluating 22 distinct fit indicators. Among these fit indicators, eight suggested a two-cluster solution and six recommended a three-cluster solution. Higher numbers of clusters were proposed by fewer than three fit indicators each. Consequently, both the two and three-cluster solutions were further examined. To circumvent local minima, 1,000 random starting positions were utilized.

For both the two and three-cluster solutions, an investigation was conducted to determine if the positive effect of additional formative assessments depended on cluster membership, or in other words, whether a significant interaction existed between cluster membership and the positive effect of solving additional formative assessments on performance on the end-of-semester knowledge test. To this end, a two-way ANOVA was performed for both the two and three-cluster solutions, with cluster membership and solving additional formative assessments as between-subject factors, and performance in the end-of-semester knowledge test as the dependent variable. The interaction was not significant for the two-cluster solution ($F(1, 304) = 1.09$, $p = 0.29$, $\eta^2 = 0.003$) but it was for the three-cluster solution ($F(2, 302) = 3.13$, $p < 0.05$, $\eta^2 = 0.02$, see Figure 2). Therefore, the three-cluster solution was further investigated. In addition to the significant interaction, there was a main effect of cluster membership ($F(2, 302) = 20.72$, $p < 0.001$, $\eta 2 = 0.11$) and a significant main effect of completing additional formative assessments ($F(1, 302) = 9.77$, $p < 0.01$, $\eta 2 = 0.03$).

In the three-cluster solution (see Table 3), one cluster ($n = 66$) exhibited low performance, low time investment, and relatively low additional investment.

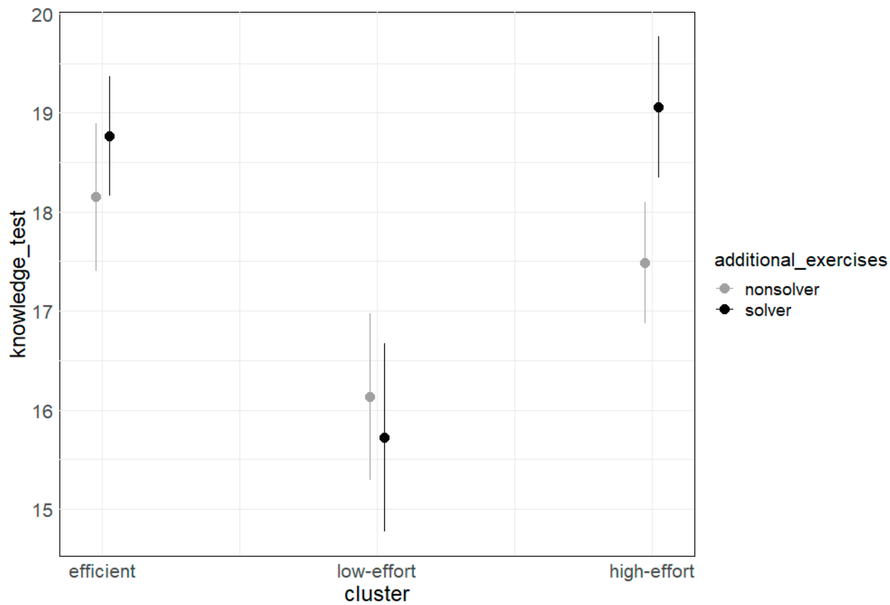


Figure 2
Interaction between the completion of additional formative assessments and the students' cluster membership in relation to their performance on the end-of-semester knowledge test

For subsequent analyses, this cluster will be denoted as the "low effort cluster." Another cluster (n = 120) was characterized by high performance, moderate time investment, and low additional investment. Accordingly, this cluster achieved high performance with comparatively low investment and is therefore referred to as the "efficient cluster." The last cluster (n = 122) exhibited above-average performance and considerable effort in both time investment and additional investment. This cluster will be referred to as the "high effort cluster" in subsequent analyses and discussions.

Table 3
Characterization of the three clusters identified as well as size of the entire sample (N) the sample of solvers, and the sample of non-solvers

	Cluster name	Performance	Time investment	Additional investment	*N* (non-solvers, solvers)
Low-performance, low-investment	low effort	−1.30	−1.02	−0.40	66 (37, 29)
High-performance, medium-investment	efficient	0.43	0.07	−0.70	120 (47, 73)
High-performance, high-investment	high effort	0.28	0.48	0.90	122 (70, 52)

Note. N = sample size of the entire sample and in parentheses sample size of non-solvers and solvers.

The performance difference in the end-of-semester knowledge test between non-solvers and solvers was −1.57 ($p < 0.01$) for the high effort cluster, −0.62 ($p = 0.21$) for the efficient cluster, and 0.41 ($p = 0.53$) for the low effort cluster. This pattern of results persists when alpha is adjusted for multiple testing. Accordingly, solving additional formative assessments appeared to be most beneficial for high effort students.

Again, solvers and non-solvers were not equally distributed across the three clusters ($\chi^2(2) = 9.26, p < 0.01$). A smaller proportion of solvers ($n = 52$) relative to non-solvers ($n = 70$) was observed in the high effort cluster, while a greater proportion of solvers ($n = 73$) compared to non-solvers ($n = 47$) was present in the efficient cluster. In contrast, the low effort cluster exhibited a more evenly distributed composition of solvers ($n = 29$) and non-solvers ($n = 37$).

Taken together, the effect of additional formative assessments depended on students' characteristics in both approaches. Both higher time investment alone and belonging to the high effort cluster resulted in a larger positive effect of additional formative assessment on the end-of-semester knowledge test. As shown in Table 4, the low effort cluster consisted mostly of students of the low time investment group (Q1), the efficient cluster consisted mostly of students with medium time investment (Q2, Q3) and the high effort cluster of high time investment students (Q3, Q4).

Table 4
Number of students in the three clusters depending on solving additional formative assessments and time investment group

Cluster	Time investment group				*Total*
	Q1	Q2	Q3	*Q4*	
Low effort	44 (26, 18)	13 (6, 7)	6 (4, 2)	3 (1, 2)	66 (37, 29)
Efficient	21 (14, 7)	40 (11, 29)	36 (15, 21)	23 (7, 16)	120 (47, 73)
High effort	12 (8, 4)	24 (11, 13)	35 (20, 15)	51 (31, 20)	122 (70, 52)

Note. Cells marked light gray contained at least thirty students. In parentheses (the number of non-solvers, the number of solvers).

3 Discussion

This study aimed to investigate the impact of additional formative assessments on students' self-regulated learning behavior and learning success, while also considering the varying impacts on different student groups. The completion of additional formative assessments covering identical content led to improved performance on the end-of-semester knowledge test. Moreover, these assessments had a differential impact on self-regulated learning behaviors across various variables. Notably, solvers exhibited enhanced performance in the formative assessments, yet reported lower levels of subjective comprehension (albeit not significantly so). They dedicated more time to completing exercises within the assessments, asked fewer questions about the exercises, and tended to engage less with the compulsory literature (albeit not significantly so). Furthermore, the influence of additional formative assessments on learning success depended on students' self-regulated learning behaviors. Both increased time investment individually and membership in the high-effort cluster contributed to a more substantial positive effect of additional formative assessments on the end-of- semester knowledge test outcomes.

3.1 Influence of additional formative assessments on self-regulated learning behavior and learning success

The positive effect of additional formative assessments on learning success is consistent with the findings of Yang et al. (2021), who conducted a meta-analytic overview. The current study extends the existing literature by demonstrating this beneficial effect in an applied setting, with complex exercises and even when the learning phase and assessment phase exercises were not identical but covered the same content.

This study found that solvers exhibited differences from non-solvers in certain aspects of self-regulated learning behavior. This was based on the initial assessments that both groups completed. Given that each formative assessment introduced new material, solvers' enhanced performance can only be attributed to an indirect testing effect, since apart from the additional formative assessments, all other conditions were identical for both groups. The indirect testing effect occurs when testing not only enhances performance on the tested material but also on new, related material (Fernandez & Jamet, 2017; Szpunar et al., 2008; Wissman et al., 2011). Consequently, the additional formative assessments impacted the solvers' self-regulated learning behavior with this new content.

In this context, the differential impact of additional formative assessments on self-regulated learning behavior offers interesting insights. Even though solvers performed better than non-solvers in formative assessments, they reported lower estimates of their understanding (albeit not significantly so) compared to non-solvers. This pattern of results indicates that solvers exhibit less overestimation of their own performance, a phenomenon known as the "illusion of knowing" (Avhustiuk et al., 2018), where learners tend to overestimate their understanding relative to their actual performance.

This pattern of results, combined with the solvers' higher time investment in solving formative assessments, indicates that the provision of additional formative assessments promotes better monitoring of one's knowledge, which is consistent with the observation made by Fernandez and Jamet (2017), and more time allocation for studying, which is in line with Soderstrom and Bjork, (2014). This can be attributed to the fact that the provision of additional formative assessments enables students to receive supplementary feedback on their learning status (Clariana & Park, 2021; Perry & Winne, 2006). This feedback assisted them in monitoring which behaviors in the initial assessments were most beneficial for their learning success in the additional formative assessments, prompting them to adjust their strategies and behaviors accordingly.

However, it is crucial to acknowledge the potential less beneficial effects of the additional learning material. The increased cognitive demands associated with the additional formative assessments, in terms of both the material's complexity and the volume of information, could lead to cognitive overload (Kossen & Ooi, 2021), and the extra exercises probably reduced the time available for students to fully engage with the material, causing them to adopt less elaborate learning strategies (e.g., less additional investment).

3.2 Student characteristics and the benefit of additional formative assessments

The impact of additional formative assessments on learning success depended on students' self-regulated learning behavior. It was primarily the students who invested above-average time during formative assessments that benefited from the additional exercises. Cluster analysis revealed that high-effort students (those with above-average time investment and above-average preparation/post-processing) gained the most from the extra exercises.

This outcome aligns with previous research by Greving et al. (2020), which demonstrated that the beneficial effect of solving exercises was most pronounced when retrieving information from memory was difficult but successful. In the high effort cluster, the retrieval of information from memory was generally successful, as overall performance in the investigated formative assessments was high. Furthermore, the retrieval of information from memory was difficult, as indicated by the above-average time investment (Dodonov & Dodonova, 2012; Dunst et al., 2014; Goldhammer, 2015) and the above-average additional effort (e.g., asking numerous questions in the forum).

The observed results align with the retrieval elaboration hypothesis (Carpenter et al., 2009). The high effort cluster demonstrated high time investment, additional investment, and above-average performance in formative assessments. Increased investment is typically linked with enhanced elaboration (Goldhammer et al., 2021). Likely due to their substantial investment, further elaboration or learning occurred during the initial formative assessments. The retrieval of this newly learned or elaborated content through additional formative assessments led to a more pronounced testing effect.

The high effort of this cluster may be correlated with high expectations of success, which is associated with a stronger positive impact of testing (Heitmann et al., 2022). Additionally, their regular learning behavior might also contribute to a more pronounced testing effect (Adesope et al., 2017; Karpicke & Bauernschmidt, 2011).

On the other hand, students in the low effort and efficient clusters did not show significant positive effects from additional formative assessments on their learning success. Low performers probably do not utilize the extra assessments effectively, while efficient learners do not require them, having already comprehended the material (Bjork et al., 2013).

For the low-effort cluster, this lack of effect might be attributed to the difficulty of the assessments, low motivation, or low elaboration of learning content (Carpenter et al., 2009; Heitmann et al., 2022; Minear et al., 2018). Exercises were probably too difficult for those students and retrieval of information was often unsuccessful, as indicated by the low performance in the formative assessments (Minear et al., 2018). According to the Yerkes-Dodson law (Yerkes & Dodson, 1908), when exercises become too difficult, motivation, response latencies and performance decrease (Borter et al., 2016; Dunst et al., 2014; Goldhammer, 2015). Their lack of prior knowledge may have posed challenges in integrating and elaborating on new but related content (Cogliano et al., 2019; Francis et al., 2020). In addition, especially for this group, the increased cognitive demands associated with the additional formative assessments, in terms of both the material's complexity and the volume of information, might have led to cognitive overload (Kossen & Ooi, 2021) or to hasty and unelaborated learning behavior due to the higher investment requirements imposed by the additional formative assessments.

In the efficient cluster, the absence of a significant positive effect could be due to either high ability and abstraction or the assessments being too easy for these students (Goldhammer, 2015) and accordingly no elaboration was needed. Even though retrieval from memory was quite successful in this cluster as indicated by the high performance in the formative assessments, it was not difficult (average time investment, very low additional investment e.g., asking questions). The exercises were probably not difficult enough for those students and after the first formative assessments no additional exercises were needed, as the students already grasped the content. Beside the possibility that formative assessments were too easy for students in this cluster, the high performance associated with rather low investment might be a sign of high ability or abstraction (Goldhammer, 2015). In this case, additional exercises are probably not necessary, as students understand the content on an abstract level and do not need different exercises from different contexts covering the same content. When low exercise difficulty is the reason for the missing effect of testing in this cluster, more difficult exercises would lead to a testing effect, whereas when high abstraction is the reason, more difficult exercises would probably not lead to a stronger testing effect. To differentiate between the two possibilities, further research is needed.

In addition, it was shown that students with poorer learning strategies show a larger testing effect than students with good strategies (Minear et al., 2018, Robey, 2019). The efficient cluster might have particularly good learning strategies as indicated by the high performance reached with rather low investment.

3.3 Solvers and non-solvers not equally distributed across time investment groups or clusters

The impact of solving additional formative assessments on self-regulated learning behavior led to an uneven distribution of students across time investment groups or clusters. Fewer solvers than non-solvers were found in the very low time investment group (Q1), while more solvers than non-solvers were present in the second time investment group (Q2). Furthermore, solvers more frequently belonged to the efficient cluster and less frequently to the high effort cluster.

On one hand, the additional formative assessments might have resulted in high effort students sacrificing additional investment (e.g., asking questions, reading literature) to invest more time in solving formative assessments (indirect testing effect, better monitoring, prioritizing different learning materials). Due to the positive effect of additional formative assessments, this resulted in higher performance. Higher performance in combination with lower additional investment is the behavioral pattern associated with the efficient cluster and led to a shift from the high effort to the efficient cluster (e.g., in Table 4, more solvers in the efficient cluster and higher time investment groups).

On the other hand, solving additional formative assessments prompted low investment students to invest more time in solving exercises and to achieve higher performance in the formative assessments (indirect testing effect). This combination of medium time investment, higher performance, and low additional investment is associated with the efficient cluster (e.g., in Table 4, there are more solvers in high time investment groups of the efficient cluster but fewer in the low effort low time investment group).

In conclusion, due to an indirect testing effect, solvers demonstrated improved monitoring associated with more efficient learning, and as a result, many solvers were part of the efficient cluster, which is linked to high performance on the end-of-semester knowledge test. Additionally, the availability of numerous formative assessments for solvers may have forced them to make decisions on where to allocate their time (Yang et al, 2017). As they spent more time on the exercises and solved a greater number of them, they reduced other activities (additional investment, fewer repetitions of the first formative assessments, but more repetitions when including additional formative assessments).

3.4 Practical relevance of the findings

As a lot of time is invested in solving additional formative assessments and not all students profit from them, it seems unethical to suggest additional assessments to all students. In the future, approaches from adaptive learning analytics (Mavroudi et al., 2018) should be implemented into the course. As indicated by the results of this study, for students with above average time investment, additional formative assessments should be suggested as adding formative assessments probably improves their learning success. For students with below-average time investment, it is important to know whether below-average time investment is associated with low or high performance in the formative assessments. If it is associated with high performance, there is no need to suggest the additional formative assessments as they probably would not lead to greater learning success. However, more difficult exercises might lead to even greater learning success in this cluster, but future research is needed to test those predictions. When low time investment is linked to low performance in formative assessments, interventions to increase content understanding, content elaboration, improve learning strategies, enhance monitoring, or adjust time allocation should be suggested. Only after successfully making these improvements should additional assessments be recommended.

When deciding whether to create additional formative assessments for a course, it is essential to consider that although many students benefited from the extra assessments and nearly all students solved them when available, the effect sizes were relatively small, and providing additional formative assessments influenced students' behavior in both beneficial and less beneficial ways. The present study highlights the importance of considering individual differences in students' self-regulated learning behavior when implementing additional formative assessments.

3.5 Measurement considerations

To investigate learning as comprehensively as possible, a variety of variables were measured, some of which were highly related. Therefore, variables of the same type (e.g., response latencies for exercises) were reduced to a single score. Observations of the same type can be interpreted as a sampling of observations, and combining them leads to a more reliable measure (Goldhammer et al., 2021). For example, when combining 100 response latencies, the influence of measurement error (e.g., taking a coffee break while solving an exercise, leading to longer response latency) is reduced. Moreover, high correlations between similar measures, as indicated by a strong first principal component, suggest that the different variables measured the same construct. The summarized measures of the same type were combined in a factor analysis. First, this resulted in well-interpretable factors (performance, time investment, additional investment), and second, fewer but more reliable measures lead to a better performance in cluster analysis (Dalmaijer et al., 2021). Based on these three factors, three clusters were built. The clusters found were similar to previous studies, in which clusters based on effort and/or processing depth (Jovanović et al., 2017; Kovanovic et al., 2015; Li et al, 2020; Ning & Downing, 2015; Parpala et al., 2021; Sun & Xie, 2020; van Alten et al., 2021; Vanslambrouck et al, 2019; Zheng et al., 2020) based on regularity of learning (Kim et al., 2018; Parpala, 2021), on prior knowledge (Khayi & Rus, 2019), on the pace of learning (Munje et al., 2020), and on performance and learning behavior were found (Waspada et al., 2019). Accordingly, the three clusters of this study fit well into previous research.

3.6 Future work

Future research could investigate how cluster membership and learning behavior evolves throughout the semester and whether adaptive hints or instructions can help students find the learning behavior or strategy that maximizes their learning success. The consistency of these clusters across various courses needs to be investigated. Furthermore, the psychological traits associated with cluster membership should be understood. It has been suggested by a recent study (Heitmann et al., 2022) that quizzing might not be beneficial for learners exhibiting a low hope of success, an attribute that might be prevalent in some of the clusters identified.

Additionally, the behavior data of the extra formative assessments should be examined, and exercise difficulty should be considered. Future research could benefit from a deeper exploration of the potential impact of assessment length on learner engagement, to discern if longer formative assessments might introduce variability in self-regulated learning. Furthermore, integrating various theories of self-regulation into our understanding of self-regulated learning behavior warrants further investigation. In addition, determining whether the positive effect of additional formative assessments can be attributed to an indirect testing effect, a direct testing effect, or a combination of both would be of significant interest in future research.

3.7 Limitations

The study's limitations primarily stem from its quasi-experimental approach in a real-world setting. Consequently, it is challenging to determine the generalizability of the findings to other courses. Furthermore, not all students in the course participated or met the inclusion criteria, which may have affected the results. Additionally, principal component analysis, exploratory factor analysis, and cluster analysis are exploratory instruments bearing the risk of false discoveries (Moosbrugger & Kelava, 2012). As a result, it is necessary to confirm or disprove these exploratory and course-specific findings in future research.

Conclusion

In conclusion, additional formative assessments led to an overall better performance in the end-of-semester knowledge test. However, this effect depended on students' characteristics. Above-average time investment was associated with a more beneficial effect of solving additional formative assessments. As indicated by the results of the cluster analysis, solvers characterized by above-average time investment and additional investment (high effort cluster) benefited from additional formative assessments, while below-average time investment was associated either with low investment/ understanding (low effort cluster) or high understanding with relatively low investment (efficient cluster). In both these clusters, no positive effect of additional formative assessments was identified. Furthermore, engaging in additional formative assessments led to changes in self-regulated learning behavior, both positive and negative, resulting in a higher proportion of solvers in the efficient cluster, which is associated with high performance on the end-of-semester knowledge test. Taken together, solving additional formative assessments is beneficial for some but not all students and is associated with both beneficial and less beneficial changes in self-regulated learning behavior.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Avhustiuk, M. M., Pasichnyk, I. D., & Kalamazh, R. V. (2018). The illusion of knowing in metacognitive monitoring: Effects of the type of information and of personal, cognitive, metacognitive, and individual psychological characteristics. *Europe's Journal of Psychology, 14*(2), 317–341. https://doi.org/10.5964/ejop.v14i2.1418

Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, Ch., Rodriguez, F., Warschauer. M., & Smyth, P. (2020). The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: Opening the black box of learning processes. *International Journal of Educational Technology in Higher Education*, *17*(3), 1–24. https://doi.org/10.1186/s41239-020-00187-1

Batsell Jr., W. R., Perry, J. L., Hanley, E., & Hostetter, A. B. (2017). Ecological validity of the testing effect: The use of daily quizzes in introductory psychology. *Teaching of Psychology*, *44*(1), 18–23. https://doi.org/10.1177/0098628316677492

Bentler, P. M., & Satorra, A. (2010). Testing model nesting and equivalence. *Psychological Methods*, *15*(2), 111–123. https://doi.org/10.1037/a0019625

Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval practice: Beneficial for all students or moderated by individual differences? *Psychology Learning & Teaching, 20*(1), 21–39. https://doi.org/10.1177/1475725720973494

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Borter, N. (2016). *Aufgabenkomplexität und Intelligenz* [Dissertation, Universität Bern]. https://boris.unibe.ch/101386/

Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research, and Evaluation*, *8*(9). https://doi.org/10.7275/kmcq-dj31

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(6), 1547–1552. https://doi.org/10.1037/a0024140

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, *23*(6), 760–771. https://doi.org/10.1002/acp.1507

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*(6), 1–36. https://doi.org/10.18637/jss.v061.i06

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2013). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5–6), 318–331. https://doi.org/10.1504/IJTEL.2012.051815

Clariana, R. B., & Park, E. (2021). Item-level monitoring, response style stability, and the hard-easy effect. *Educational Technology Research and Development*, *69*, 693–710. https://doi.org/10.1007/s11423-021-09981-8

Cogliano, M., Kardash, C. M., & Bernacki, M. L. (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments. *Contemporary Educational Psychology*, *56*, 117–129. https://doi.org/10.1016/j.cedpsych.2018.12.001

Dalmaijer, E. S., Nord, C. L., & Astle, D. E. (2021). *Statistical power for cluster analysis*. arXiv. https://doi.org/10.48550/arXiv.2003.00381

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). *Understanding and using factor scores: Considerations for the applied researcher*, *14*(20). https://doi.org/10.7275/DA8T-4G52

Dodonov, Y. S., & Dodonova, Y. A. (2012). Response time analysis in cognitive tasks with increasing difficulty. *Intelligence*, *40(5)*, 379–394. https://doi.org/10.1016/j.intell.2012.07.002

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, *47*(2), 309–326. https://doi.org/10.1111/j.2044-8317.1994.tb01039.x

Dunst, B., Benedek, M., Jauk, E., Bergner, S., Koschutnig, K., Sommer, M., Ischebeck, A., Spinath, B., Arendasy, M., Bühner, M., Freudenthaler, H., & Neubauer, A. C. (2014). Neural efficiency as a function of task demands. *Intelligence*, *42*, 22–30. https://doi.org/10.1016/j.intell.2013.09.005

Eriksson, J., Kalpouzos, G., & Nyberg, L. (2011). Rewiring the brain with repeated retrieval: A parametric fMRI study of the testing effect. *Neuroscience Letters*, *505*(1), 36–40. https://doi.org/10.1016/j.neulet.2011.08.061

Fernandez, J., & Jamet, E. (2017). Extending the testing effect to self-regulated learning. *Metacognition and Learning*, *12*, 131–156. https://doi.org/10.1007/s11409-016-9163-9

Foss, D. J., & Pirozzolo, J. W. (2017). Four semesters investigating frequency of testing, the testing effect, and transfer of training. *Journal of Educational Psychology*, *109*(8), 1067–1083. https://doi.org/10.1037/edu0000197

Francis, A. P., Wieth, M. B., Zabel, K. L., & Carr, T. H. (2020). A classroom study on the role of prior knowledge and retrieval tool in the testing effect. *Psychology Learning & Teaching*, *19*(3), 258–274. https://doi.org/10.1177/1475725720924872

Goldhammer, F. (2015). Measuring ability, speed, or both? Challenges, psychometric solutions, and what can be gained from experimental control. *Measurement: Interdisciplinary Research and Perspectives*, *13*(3–4), 133–164. https://doi.org/10.1080/15366367.2015.1100020

Goldhammer, F., Hahnel, C., Kroehne, U., & Zehner, F. (2021). From byproduct to design factor. On validating the interpretation of process indicators based on log data. *Large-Scale Assessments in Education*, *9*, 1–25. http://nbn-resolving.de/urn:nbn:de:0111-pedocs-250050

Greene, R. L. (2008). Repetition and spacing effects. In H. L. Roediger (Ed.), *Learning and memory: A comprehensive reference. Vol. 2: Cognitive psychology of memory* (pp. 65–78). Elsevier.

Greving, S., Lenhard, W., & Richter, T. (2020). Adaptive retrieval practice with multiple-choice questions in the university classroom. *Journal of Computer Assisted Learning*, *36*(6), 799–809. https://doi.org/10.1111/jcal.12445

Heitmann, S., Grund, A., Fries, S., Berthold, K., & Roelle, J. (2022). The quizzing effect depends on hope of success and can be optimized by cognitive load-based adaptation. *Learning and Instruction*, *77*. https://doi.org/10.1016/j.learninstruc.2021.101526

Ifenthaler, D. (2015). Learning Analytics. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (Vol. 2, pp. 447–451). SAGE publication. https://madoc.bib.uni-mannheim.de/38809/

Jensen, J. L., McDaniel, M. A., Kummer, T. A., Godoy, P. D. D. M., & St. Clair, B. (2020). Testing effect on high-level cognitive skills. *CBE—Life Sciences Education*, *19*(3). https://doi.org/10.1187/cbe.19-10-0193

Jost, N. S., Jossen, S. L., Rothen, N., & Martarelli, C. S. (2021). The advantage of distributed practice in a blended learning setting. *Education and Information Technologies*, *26*, 3097–3113. https://doi.org/10.1007/s10639-020-10424-9

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001

Karpicke, J., & Bauernschmidt, A. (2011). Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(5), 1250–1257. https://doi.org/10.1037/a0023436

Karpicke, J., & Roediger, H. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(4), 704–719. https://doi.org/10.1037/0278-7393.33.4.704

Khayi, N. A., & Rus, V. (2019). Clustering students based on their prior knowledge. In F. Collin, A. Merceron, & M. Desmarais (Eds.), *Proceedings of the 12th International Conference on Educational Data Mining.* Université du Québec à Montréal, Polytechnique Montréal (pp. 246–251). https://educationaldatamining.org/edm2019/proceedings/

Kim, D., Yoon, M., Jo, I.-H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education*, *127*, 233–251. https://doi.org/10.1016/j.compedu.2018.08.023

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). Guilford Press.

Kossen, C., & Ooi, C.-Y. (2021). Trialling micro-learning design to increase engagement in online courses. *Asian Association of Open Universities Journal*, *16*(3), 299–310. https://doi.org/10.1108/AAOUJ-09-2021-0107

Kovanovic, V., Gasevic, D., Joksimovic, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, *27*, 74–89. https://doi.org/10.1016/j.iheduc.2015.06.002

Lamotte, M., Izaute, M., & Darnon, C. (2021). Can tests improve learning in real university classrooms? *Journal of Cognitive Psychology*, *33*(8), 974–992. https://doi.org/10.1080/20445911.2021.1956939

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education–A Literature Review. In A. Peña-Ayala (Ed.), *Learning analytics: Fundaments, applications, and trends. Studies in systems, decision and control* (pp. 1–23). Springer. https://doi.org/10.1007/978-3-319-52977-6_1

Li, S., Chen, G., Xing, W., Zheng, J., & Xie, C. (2020). Longitudinal clustering of students' self-regulated learning behaviors in engineering design. *Computers & Education*, *153*, Article 103899. https://doi.org/10.1016/j.compedu.2020.103899

Li, H., Flanagan, B., Konomi, S. & Ogata, H. (2018). Measuring behaviors and identifying indicators of self-regulation in computer-assisted language learning courses. *Research and Practice in Technology Enhanced Learning*, *13*(19), 1–12. https://doi.org/10.1186/s41039-018-0087-7

Mavroudi, A., Giannakos, M., & Krogstie, J. (2018). Supporting adaptive learning pathways through the use of learning analytics: Developments, challenges and future opportunities. *Interactive Learning Environments*, *26*(2), 206–220. https://doi.org/10.1080/10494820.2017.1292531

McDaniel, M. A, Thomas, R. C, Agarwal, P. K, McDermott, K. B, & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology*, *27*(3), 360–372. https://doi.org/10.1002/acp.2914

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology*, *106*(1), 121–131. https://doi.org/10.1037/a0033546

Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., & Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1474–1486. https://doi.org/10.1037/xlm0000486

Moosbrugger, H., & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*. Springer. https://doi.org/10.1007/978-3-642-20072-4

Ning, H. K., & Downing, K. (2015). A latent profile analysis of university students' self-regulated learning strategies. *Studies in Higher Education*, *40*(7), 1328–1346. https://doi.org/10.1080/03075079.2014.880832

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Parpala, A., Mattsson, M., Herrmann, K. J., Bager-Elsborg, A., & Hailikari, T. (2021). Detecting the variability in student learning in different disciplines—A person-oriented approach. *Scandinavian Journal of Educational Research*, *66*(6), 1020–1037. https://doi.org/10.1080/00313831.2021.1958256

Perry, N. E., & Winne, P. H. (2006). Learning from Learning Kits: gStudy traces of students' self-regulated engagements with computerized content. *Educational Psychology Review*, *18*, 211–228. https://doi.org/10.1007/s10648-006-9014-3

Pishgar, F., Greifer, N., Leyrat, C., & Stuart, E. (2021). MatchThem: Matching and weighting after multiple imputation. *The R Journal*, *13*(2), 294–305. https://doi.org/10.32614/RJ-2021-073

R Core Team (2021). *R: A language and environment for statistical computing*. R foundation for statistical computing. Retrieved from https://www.R-project.org

Revelle, W. R. (2022). *psych: Procedures for psychological, psychometric, and personality research. R package version 2.2.5*. Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Robey, A. (2019). The benefits of testing: Individual differences based on student factors. *Journal of Memory and Language*, *108*. https://doi.org/10.1016/j.jml.2019.104029

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.589965

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, *16*(2), 179–196. https://doi.org/10.1177/1475725717695149

Shin, D., & Shim, J. (2021). A systematic review on data mining for mathematics and science education. *International Journal of Science and Mathematics Education*, *19*, 639–659. https://doi.org/10.1007/s10763-020-10085-7

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language*, *73*, 99–115. https://doi.org/10.1016/j.jml.2014.03.003

Sun, Z., & Xie, K. (2020). How do students prepare in the pre-class setting of a flipped undergraduate math course? A latent profile analysis of learning behavior and the impact of achievement goals. *The Internet and Higher Education*, *46*. https://doi.org/10.1016/j.iheduc.2020.100731

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. III. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1392–1399. https://doi.org/10.1037/a0013082

van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2021). Secondary students' online self-regulated learning during flipped learning: A latent profile analysis. *Computers in Human Behavior*, *118*. https://doi.org/10.1016/j.chb.2020.106676

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Vanslambrouck, S., Zhu, C., Pynoo, B., Lombaerts, K., Tondeur, J., & Scherer, R. (2019). A latent profile analysis of adult students' online self-regulation in blended learning environments. *Computers in Human Behavior*, *99*, 126–136. https://doi.org/10.1016/j.chb.2019.05.021

Waspada, I., Bahtiar, N., & Wibowo, A. (2019). Clustering student behavior based on quiz activities on moodle LMS to discover the relation with a final exam score. *Journal of Physics: Conference Series*, *1217*. https://doi.org/10.1088/1742-6596/1217/1/012118

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*, *18*, 1140–1147. https://doi.org/10.3758/s13423-011-0140-7

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Yang, C., Potts, R., & Shanks, D. R. (2017). The forward testing effect on self-regulated study time allocation and metamemory monitoring. *Journal of Experimental Psychology: Applied*, *23*(3), 263–277. https://doi.org/10.1037/xap0000122

Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology & Psychology, 18,* 459–482. https://doi.org/10.1002/cne.920180503

Zheng, J., Xing, W., Zhu, G., Chen, G., Zhao, H., & Xie, C. (2020). Profiling self-regulation behaviors in STEM learning of engineering design. *Computers & Education*, *143*. https://doi.org/10.1016/j.compedu.2019.103669

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-2

GOAL SETTING AND GOAL ORIENTATION AS PREDICTORS OF LEARNING SATISFACTION AND ONLINE LEARNING BEHAVIOR IN HIGHER EDUCATION BLENDED COURSES

Libor Juhaňák[a], Karla Brücknerová[a], Barbora Nekardová[a], Jiří Zounek[a]

[a] Faculty of Arts, Masaryk University, Czech Republic

ABSTRACT

This study investigated how goal setting and goal orientation are related to student learning behavior and engagement in an online learning environment, and how learning behavior, goal setting, and goal orientation are related to student satisfaction with the course they are studying. A total of 882 students from 76 different courses participated in this study, which used both self-reported data from a questionnaire and indicators based on digital traces in an online learning environment. The results of multilevel regression analyses showed that student ability to set learning goals (i.e., goal setting) was positively related to both student learning satisfaction and student learning behavior. Intrinsic goal orientation positively predicted student satisfaction with the course. Extrinsic goal orientation did not show a significant effect in any of the observed relationships. The analyzed indicators of student learning behavior showed no statistically significant association with learning satisfaction. Possible explanations for these findings are discussed, and limitations and directions for future research are suggested.

KEYWORDS

self-regulated learning; goal setting; goal orientation; learning engagement; online learning behavior; course satisfaction

CORRESPONDING AUTHOR

Libor Juhaňák, Department of Educational Sciences, Faculty of Arts, Masaryk University, Arna Nováka 1/1, 602 00 Brno, Czech Republic
e-mail: juhanak@phil.muni.cz

Introduction

Self-regulated learning (SRL) entails processes that empower learners to regulate their cognition, emotions, and behavior as they engage in learning tasks (Pintrich, 2004). Although these processes are employed cyclically throughout the learning process, it is possible to distinguish three phases of SRL: forethought, performance, and self-reflection (Panadero, 2017). In the forethought phase, SRL involves goal setting and strategic planning, incorporating motivational beliefs such as self-efficacy, outcome expectations, goal orientation, and the intrinsic value of the learning task. The performance phase encompasses self-control and self-observation during learning (Zimmerman & Campillo, 2003). In the self-reflection phase, learners engage in self-judgment by comparing their learning outcomes with their expectations, making causal attributions for their results, and responding emotionally to their learning outcomes (Zimmerman, 2002). The extent to which learners engage in SRL during their studies is associated with their academic achievement and satisfaction (Broadbent & Poon, 2015). In our study, we focus on two SRL processes that relate to how learners set their goals in the first phase of their SRL.

Engagement in SRL is essential in online learning environments, including the learning management systems (LMS) that are widely used in higher education, as such environments put high demands on student ability to structure, process, and evaluate their online learning (Wong et al., 2019). In recent years, there has been a shift in the way online learning is studied, with researchers looking for new ways to explore and measure the online learning process, attempting to move away from self-report questionnaires toward new indicators based on capturing digital traces of student learning behavior in online learning environments (Winne, 2010). In this approach, learning engagement in online learning environments may be measured in terms of the number and frequency of visits or learning time spent in the online learning environment (Kim, et al., 2016; Kim et al., 2018; Kovanović et al., 2015a; Kovanović et al., 2015b). Learning analytics, the broader research area that specifically focuses on capturing and investigating online learning processes through the collection and analysis of data available in online learning environments, appears to be a promising approach to studying online learning behavior (Vieira et al., 2018; Winne, 2017). However, in the context of learning in blended courses, learning analytics can provide only a limited insight into student learning processes, as not all learning takes place in online learning environments (Ifenthaler & Schumacher, 2016; Wilson et al., 2017). Digital traces of student learning behavior offer useful insights into the actions that students take while learning; on the other hand, these data can suffer from the ambiguity of interpretation (Gašević et al., 2016; Wise & Shaffer, 2015). This study combines subjective data (i.e., self-reports) and data of an objective nature (i.e., logs) to gain further insight into the relationship between student SRL, online learning behavior, and satisfaction with the course.

We aim to answer the question of how goal setting and goal orientation are related to student engagement within the online learning environment, as demonstrated by the number of visits, regularity of visits, and total time spent. At the same time, we aim to answer the question of how learning behavior, goal setting, and goal orientation are related to student satisfaction with the course.

1 Theoretical background

1.1 Goal setting and goal orientation within self-regulated learning

Setting learning goals for one's own learning is an essential part of SRL processes across the different theoretical approaches to describing the SRL model (Panadero, 2017). In Zimmerman's model, goal setting and goal orientation are processes that fall within the first of three phases of SRL. In this first phase, learners focus on forethought and planning their learning process. The forethought phase consists of task analysis, including goal setting and strategic planning, and self-motivation beliefs consisting of self-efficacy, outcome expectations, intrinsic interest, and goal orientation (Zimmerman, 2002; Zimmerman & Moylan, 2009). Pintrich (2004) argued that the first phase of SRL covers forethought, planning, and activation; it involves cognitive, affective, and behavioral processes, as well as the perception of the learning task and context. In this SRL model, the learner sets goals at the cognitive level and adopts goal orientation at the motivational and affective levels. According to Winne (2013), these processes follow task definition and consist of setting goals and planning how to achieve them, linking goals and tactics before the learner starts working on the task itself. To sum up, goal setting and goal orientation cover the cognitive as well as the motivational and affective aspects of how learners deal with goals in learning.

Goal setting is the process of identifying goals and deciding what outcomes one wants to achieve (Zimmerman & Moylan, 2009). Setting goals is linked to various aspects of student learning at university. Goal setting, along with other SRL behaviors, is associated with the perception of online courses (Barnard et al., 2008) and with the quality of learning resources (Ballouk et al., 2022). Students with higher levels of goal setting are more likely to adopt a deep learning approach to learning (Soyer & Kirikkanat, 2019). Goal setting is also associated with academic achievement (Ballouk et al., 2022; Barnard et al., 2008).

Goal orientations describe the broader purposes of achievement behavior and explain how people behave in achievement situations and why (Kaplan & Maehr, 2007). It is possible to distinguish mastery and performance goal orientations. Students applying mastery goal orientations focus on achieving task-based or intra-personal competence; those applying performance goal orientations focus on being well perceived by others (Miller et al., 2021). Similarly, based on goal content theory, we can distinguish between extrinsic and intrinsic goal orientations (Kasser & Ryan, 1996). Intrinsic goal orientation focuses on personal growth and learning itself; extrinsic goal orientation is associated with the fulfilment of goals such as achievement, recognition from others, and obtaining material benefits (Zhang et al., 2018). Goal orientations influence how students perceive different components of the learning environment, such as course assessment (Kaur et al., 2018), and their behavior, such as task selection (Lindfors, 2021).

1.2 Online learning behavior and course satisfaction

Learning satisfaction is understood as an affective dimension of learning outcomes (Klein et al., 2006). Course satisfaction can then be viewed as the satisfaction arising from studying a particular course. In the context of online and blended learning, learning satisfaction seems to be one of the key factors that determine learning retention and academic success when learning in an online learning environment (Ke & Kwak, 2013; She et al., 2021). Huang (2023) stated that effective goal setting promotes learning motivation and higher learning satisfaction, which in turn leads to better performance and well-being. At the same time, Klein et al. (2006), in their research on blended learning environments, found a significant positive relationship between goal orientation and learning satisfaction. Similar results were found in relation to goal orientation among college students (Sánchez-Cardona et al., 2021). Another study (Ma & She, 2023) found a positive correlation between goal orientation and learning satisfaction, while also focusing on a mediating effect in this relationship, with academic self-efficacy and learning engagement acting as intervening factors. Thus, our first hypothesis is as follows:

- **H1** Goal setting and goal orientation positively correlate with student course satisfaction.

The relationship between student learning behavior and their learning engagement was assumed in Pintrich's SRL model (Pintrich, 2004), in which learning goal orientation was seen as a kind of motivational process within the forethought phase, which lays the foundation for the subsequent performance phase, within which students regulate their learning behavior (i.e., it affects student learning engagement). Several studies have concluded that student online learning behavior is influenced by their affective characteristics. An important implication of a study by Schwam et al. (2021) was that students who lack confidence in their ability to navigate the online learning environment may not use SRL techniques as effectively as their more proficient peers. This discomfort has the potential to impede the learning process, as students not only have to engage with the course content, but also have to invest considerable effort in familiarizing themselves with the intricacies of the online learning environment. Similarly, Zhang & Liu (2019) suggested that student learning behavior and learning engagement are driven by their learning goals and plans. Therefore, we propose the following hypothesis:

- **H2** Goal setting and goal orientation positively correlate with student behavior in the online learning environment.

Pintrich's SRL model (Pintrich, 2004) also proposes a relationship between student learning behavior (in other words, their learning engagement) and learning satisfaction. In the context of online learning environments, learning engagement is manifested by, for example, the frequency of course visits, the regularity of completing assignments, and the number of posts in discussion forums. Learning engagement is a determinant of the learning experience and thus subsequent course satisfaction (Rajabalee & Santally, 2021). Research has suggested that learning behavior and learning engagement are among the strongest predictors of learning satisfaction (Murillo-Zamorano et al., 2019). A study by El-Sayad et al. (2021) focused on online learning during the COVID-19 pandemic found a significant relationship between student behavioral engagement and their learning satisfaction. At the same time, students who are not sufficiently engaged in the learning process tend to experience low levels of learning satisfaction (Gao et al., 2020). Thus, we propose the last hypothesis as follows:

- **H3** Student behavior in the online learning environment positively correlates with student course satisfaction.

2 Methods

2.1 Sample and procedure

This study focuses on goal-related aspects of SRL and their relationship to student satisfaction and behavior in an online learning environment used to support student learning in blended university courses. The study combines data from two different sources: data from a questionnaire survey that the students completed during the semester and data extracted from the database of the online learning environment that the students used during the semester as part of their coursework.

Participants in this study were recruited from 76 different courses taught at the Faculty of Arts, Masaryk University (Czech Republic) during three different semesters. In each of the three semesters, students of different courses were approached to complete a questionnaire focusing on various dimensions of SRL and other relevant factors. After the end of the semester, relevant data on student behavior in the online learning environment were extracted for those students who agreed to participate in the research. All of the research, the distribution of the questionnaire and the extraction of data from the online learning environment, was carried out in cooperation with the teachers of the selected courses.

A total of 882 student responses were analyzed. Only courses with responses from at least five different students were included in the sample. Students from both bachelor's (77.55%) and non-follow-up master's fields (21.77%) were approached to complete the questionnaire. The mean age of the respondents was 21.98 years (med = 21). Regarding gender, 76.87% of the respondents were female and 21.77% were male, which corresponds to the gender distribution of the students at the Faculty of Arts (1.02% of the students chose the option "other"). The vast majority of students in the sample were full-time students (94.56%).

2.2 Measures

Goal setting was measured using a five-item scale developed by Barnard, Lan, To, Paton, & Lai (2009) and used as one of six subscales within the *Online Self-Regulated Learning Questionnaire* (OSLQ). The goal setting scale consisted of five-point Likert scale items such as: "I set standards for my assignments in online courses," "I set short-term (daily or weekly) goals as well as long-term (monthly or for the semester) goals," and "I set goals to help me manage study time for my online courses." The wording of the items was slightly modified to suit the context of this study (e.g., blended courses using online support in an online learning environment to varying degrees in combination with in-class instruction). The Cronbach's alpha in the original study (i.e., Barnard et al., 2009) was 0.95; in our sample, the Cronbach's alpha is 0.763, which can probably be explained by the use of the scale in the context of blended courses.

Student goal orientation was measured using two subscales from the *Motivated Strategies for Learning Questionnaire* (MSLQ) developed by Pintrich et al. (1991). In this questionnaire, the authors distinguish between **intrinsic and extrinsic goal orientation** when measuring student goal orientation. Both the intrinsic and extrinsic goal orientation scales consist of four seven-point Likert scale items. The intrinsic goal orientation scale consisted of items such as: "In a class like this, I prefer course material that really challenges me so I can learn new things" and "The most satisfying thing for me in this course is trying to understand the content as thoroughly as possible." The Cronbach's alpha of the intrinsic goal orientation scale reported in Pintrich's manual for the MSLQ (Pintrich et al., 1991) was 0.74; in our sample, the Cronbach's alpha is 0.815. The extrinsic goal orientation scale consisted of items such as: "I want to do well in this class because it is important to show my ability to my family, friends, employer, or others" and "If I can, I want to get better grades in this class than most of the other students." The wording of the items was slightly modified to fit the context of Czech university studies (for example, the items related to grade point average (GPA) have been modified as GPA is not a highly used indicator in the Czech context compared to other countries). The Cronbach's alpha of Pintrich's manual for the MSLQ was 0.62; in our sample, the Cronbach's alpha is 0.82.

To measure student learning satisfaction when studying a blended course with online support in an online learning environment, we used the five-item **course satisfaction** scale used by Lee, Srinivasan, Trail, Lewis, & Lopez (2011), which included items such as: "This course increased my interest in the subject," "I felt comfortable in this course," and "I would recommend this course to others." Respondents answered on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The Cronbach's alpha in the original study (i.e., Lee et al., 2011) was 0.94; in our sample, the Cronbach's alpha is 0.9.

Table 1
Basic descriptive statistics and Cronbach's alpha for the questionnaire scales

	Min–Max	Mean	Median	SD	Skewness	Kurtosis	Alpha
Goal Setting	1–5	3.28	3.20	0.86	–0.15	–0.25	0.76
Intrinsic goal orientation	1–7	4.91	5.00	1.22	–0.56	0.26	0.82
Extrinsic goal orientation	1–7	4.16	4.25	1.53	–0.12	–0.63	0.82
Course satisfaction	1–5	4.03	4.20	0.94	–1.11	0.61	0.90

The second set of variables used in the analyses were related to online learning behavior. The variables were created based on student log records extracted from the database of the online learning environment that served as the main online learning support for each course. After pre-processing the student log records, three proxy indicators were used to measure and describe student learning behavior in the online learning environment (OLE): 1) number of course visits in OLE, 2) irregularity of visits in OLE, and 3) total time spent in online course support in OLE.

For the **number of course visits** in OLE, a visit was conceptualized as a situation where a student enters a course in OLE, spends some time in the course, and then leaves the course. If the student returns to the course after a period of time (e.g., a week), this is counted as a new visit. In most online learning environments, it is not possible to distinguish exactly between visits, because the end of a visit is usually not explicitly recorded in the system (i.e., the student does not explicitly log out of the system, but simply stops working in the system and closes the browser window). Therefore, an inactivity threshold is usually used to distinguish between individual visits (cf. Kovanović et al., 2015a). For this analysis, an inactivity threshold of 30 minutes was chosen, which means that a 30-minute period of inactivity was used as an indicator of the end of the visit.

The **irregularity of visits** captures another aspect of student learning behavior in OLE. To measure the degree of irregularity of visits, we used the approach suggested by Jo et al., 2015 and Kim et al., 2018, who calculate the irregularity of visits using the standard deviations of the time intervals between individual visits in the course. In general, the lower the value of this variable, the more regularly the student attends the course. For example, a student who regularly attends a course at the same time every week will have a low value for this variable, while a student who attends a course sporadically will have a high value for this variable.

The final variable used to capture student online learning behavior was the **total time spent in the course** in OLE. Similar to the number of visits, a threshold of 30 minutes of inactivity was used to signal the end of a student's visit to the course. The duration of the visit (in minutes) was then calculated as the time from the first log record within that visit to the last log record within that visit (i.e., the last log record preceding the 30-minute period of inactivity). However, for each visit, the estimated time spent on the last activity was added to the time difference between the last and first log within a visit. Following other studies (cf. Kim et. al., 2016; Kovanović et al., 2015b), we estimated the time spent on the last activity as the average time spent on the other activities within the same visit.

As all three variables related to student online learning behavior showed a non-normal distribution, a logarithmic transformation was performed before using the variables in the regression models. The table below shows the descriptive statistics of the variables before and after transformation.

Table 2
Basic descriptive statistics for the student online learning behavior variables

	Min	Max	Mean	Median	SD	Skewness	Kurtosis
Before transformation							
Number of visits in OLE	0	179	37.7	33	27.2	1.39	3.05
Irregularity of visits in OLE	0.02	47.3	4.84	3.14	5.07	3.68	18.9
Total time spent in OLE	0.10	2485.1	402.2	307.5	349.3	2.04	6.00
After transformation							
Number of visits in OLE	0	5.19	3.31	3.5	0.93	–1.16	1.70
Irregularity of visits in OLE	0.03	3.86	1.30	1.15	0.70	0.86	0.55
Total time spent in OLE	0.87	7.82	5.62	5.73	0.97	–0.81	1.10

2.3 Data analysis

Given the hierarchical nature of the data analyzed (i.e., students nested within individual courses), multilevel modelling was used for data analysis (cf. Heck & Thomas, 2015; Hox, 2010; Snijders & Bosker, 2012). The main reason for using multilevel modelling was that we did not control for the learning design of the courses in the study. On the contrary, our aim was not to focus only on a specific learning design of the courses, but to cover a wide range of different types of courses that are commonly used in online learning environments at universities.

First, basic descriptive statistics (see Tables 1 and 2) and correlations between all analyzed variables (see Table 3) were calculated. Then, for each of the dependent variables, a null model was estimated as a basis for calculating the intra-class correlation coefficient (ICC). Finally, separate multilevel models were estimated for each of the hypotheses tested. All preprocessing and analyses were performed using the R statistical software (Posit team, 2023; R Core Team, 2023). The *lme4* library (Bates et al., 2015) was used for multilevel modelling.

3 Results

Initially, descriptive statistics and correlations were calculated (see Tables 1, 2, and 3). Among the descriptive statistics, the higher skewness of course satisfaction is worth mentioning, which showed that in our sample there were rather high values of course satisfaction. As far as correlations are concerned, there was a very strong correlation between all three indicators of student learning behavior. While the strong positive correlation between the number of visits and the total time spent in the course is probably to be expected, the strong negative correlation between the number of visits and the irregularity of visits is perhaps not immediately expected, but it follows from the way the irregularity of visits is calculated. In addition, we calculated the intra-class correlation coefficient for all dependent variables in the following models. The ICC for course satisfaction was 0.241, which means that about 24.1% of the variance in student course satisfaction was due to differences between courses. The ICCs for the indicators of student online learning behavior are as follows: number of visits = 0.533, irregularity of visits = 0.473, total time spent = 0.479. These ICCs can be considered relatively high and further indicate the need to estimate multilevel models.

Table 3
Correlations between all analyzed variables

	1)	**2)**	**3)**	**4)**	**5)**	**6)**
1) Goal setting	1					
2) Intrinsic goal orientation	0.444	1				
3) Extrinsic goal orientation	0.486	0.414	1			
4) Course satisfaction	0.342	0.575	0.267	1		
5) Number of visits in OLE	0.212	0.025	0.105	0.099	1	
6) Irregularity of visits in OLE	−0.196	0.012	−0.061	−0.039	−0.860	1
7) Total time spent in OLE	0.214	0.047	0.080	0.120	0.794	−0.612

To address the first hypothesis, which focused on goal setting and goal orientation and their relationship to student course satisfaction, a model was constructed with goal setting and intrinsic and extrinsic goal orientation as independent variables and course satisfaction as the dependent variable. The resulting model is presented in Table 4, which shows that both goal setting and goal orientation are significant factors in student course satisfaction. However, within goal orientation, only the intrinsic dimension was statistically significant; the extrinsic dimension of goal orientation did not seem to have an effect on student course satisfaction. In both cases, the identified significant relationship was positive: higher goal setting and higher intrinsic goal orientation led to higher student satisfaction with the course.

Table 4
Effects of goal setting and goal orientation on student course satisfaction

	Course satisfaction		
	Coef.	**SE**	**p**
Fixed Effects			
(Intercept)	1.73	0.12	
Goal setting	0.13	0.03	**<0.001**
Intrinsic goal orientation	0.36	0.02	**<0.001**
Extrinsic goal orientation	0.01	0.02	0.584
Random Effects			
Residual variance	0.47		
Intercept variance	0.13		
Fit statistics			
Marginal R^2 / Conditional R^2	0.304 / 0.458		
Deviance	1923.8		
AIC	1957.1		

Subsequently, three models were created in relation to the second hypothesis, which dealt with the relationship between goal-related variables and student behavior in the online learning environment. Thus, a separate model with goal setting, intrinsic goal orientation, and extrinsic goal orientation as independent variables was created for each dependent variable related to student online learning behavior (i.e., number of visits to the course in OLE, irregularity of visits to the course in OLE and total time spent in the course in OLE). The data for all three models are presented in Table 5, which shows that of the three independent variables, only goal setting had a statistically significant effect on student online learning behavior. At the same time, the variable appears to have had a significant effect on all three observed indicators of student behavior. For the number of visits and total time spent, the effect of goal setting was positive, i.e., the better a student was able to set their own goals, the more often they attended the course and the more total time they spent on the course in OLE. For the irregularity of course visits, the observed effect is negative, indicating a positive relationship between goal setting and regularity of course visits. So, similarly to above, the better a student was able to set their learning goals, the more regularly they visited the online learning support of the course being studied.

Table 5
Effects of goal setting and goal orientation on indicators of student online learning behavior

	Number of visits			**Irregularity of visits**			**Total time spent**		
	Coef.	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**
Fixed Effects									
(Intercept)	2.90	0.13		1.59	0.11		5.06	0.14	
Goal Setting	0.10	0.03	**0.002**	–0.11	0.03	**<0.001**	0.11	0.04	**0.003**
Intrinsic goal orientation	–0.01	0.02	0.607	0.03	0.02	0.129	0.01	0.02	0.658
Extrinsic goal orientation	0.03	0.02	0.056	–0.01	0.01	0.455	0.03	0.02	0.164
Random Effects									
Residual variance	0.40			0.27			0.49		
Intercept variance	0.45			0.23			0.45		
Fit statistics									
Marginal R^2 / Conditional R^2	0.015 / 0.536			0.016 / 0.469			0.016 / 0.485		
Deviance	1872.5			1441.7			2022.3		
AIC	1905.2			1476.1			2054.4		

The last hypothesis dealt with the relationship between student online learning behavior and student course satisfaction. To test the first hypothesis, we started from the model presented in Table 4 and added the indicators of student learning behavior as three additional independent variables. Similarly to the first model, student course satisfaction served as the dependent variable. The results for this model are presented in Table 6 and show that none of the three indicators of student behavior had a statistically significant relationship with student satisfaction with the course. At the same time, the fit statistics for this model do not appear to be significantly different from those of the original model (cf. Table 4). This further confirms that the addition of indicators of student online learning behavior does not help explain the variability in student course satisfaction.

Table 6
Effects of goal setting, goal orientation, and indicators of online learning behavior on student course satisfaction

	Course satisfaction		
	Coef.	**SE**	**p**
Fixed Effects			
(Intercept)	1.56	0.36	
Goal Setting	0.12	0.03	**<0.001**
Intrinsic goal orientation	0.37	0.02	**<0.001**
Extrinsic goal orientation	0.00	0.02	0.935
Number of visits	−0.00	0.09	0.965
Irregularity of visits	−0.01	0.07	0.910
Total time spent	0.04	0.05	0.405
Random Effects			
Residual variance	0.46		
Intercept variance	0.11		
Fit statistics			
Marginal R^2 / Conditional R^2	0.313 / 0.445		
Deviance	1802.8		
AIC	1854.3		

4 Discussion and conclusion

The aim of this study was to investigate the relationship between goal-related SRL processes (goal setting and goal orientation) and student learning behavior in the online learning environment, and how goal setting, goal orientation, and student online learning behavior are related to student satisfaction with the course. In order to fulfil these aims, we formulated three hypotheses and tested them with multilevel regression analysis.

The first hypothesis predicted a positive relationship between goal setting and goal orientation and student course satisfaction. Based on the analysis conducted, we can confirm the proposed hypothesis. However, for goal orientation, the analyses carried out in this study distinguished between intrinsic and extrinsic goal orientation, which allowed us to reveal that the positive relationship between goal orientation and student course satisfaction occurs only with intrinsic goal orientation. In contrast, no statistically significant relationship was found between extrinsic goal orientation and student course satisfaction.

These findings are consistent with earlier research showing a positive association with learning satisfaction for both goal setting and goal orientation (Huang, 2023; Klein et al., 2006; Ma & She, 2023; Sánchez-Cardona et al., 2021). On the other hand, our findings highlight the importance of distinguishing between intrinsic and extrinsic goal orientation, as the relationship between these two types of goal orientation and student satisfaction appears to be very different. Existing studies have not always distinguished between these two types of goal orientation and have dealt with goal orientation in general, which may obscure important differences. For example, the study by Sánchez-Cardona et al. (2021), which reported a positive relationship between goal orientation and learning satisfaction, actually focused specifically on intrinsic goal orientation (as can be inferred from a close reading of the methodology and measures used), meaning that their findings are quite consistent with our results. The same is true for a number of other studies that report a positive relationship between goal orientation and student learning satisfaction (Klein et al., 2006; Ma & She, 2023). Many other researchers (Kasser & Ryan, 1996; Lee et al., 2010; Miller et al., 2021; Zhang et al., 2018) have distinguished between intrinsic and extrinsic goal orientation, showing that these are indeed two very different dimensions of goal orientation and that their relationship to student learning satisfaction may be distinct and more complex.

Our second hypothesis focused on the relationship between goal setting and goal orientation and student behavior in an online learning environment, predicting that higher levels of goal setting and goal orientation would lead to higher levels of student activity and engagement in the online learning environment. In dealing with the second hypothesis, we focused our attention on three different indicators of student online learning behavior: the number of student visits in the course, the irregularity of student visits in the course, and the total time spent in the course within the online learning environment. The results of the analysis suggest that only goal setting has a statistically significant positive effect on student learning behavior. That is, a greater ability of students to set goals for their own learning is associated with more frequent course attendance, higher regularity of course attendance, and overall greater time spent in the course. On the other hand, goal orientation was not found to be statistically significantly related to any of the three indicators of student behavior in the online learning environment. This was the case for both intrinsic and extrinsic goal orientation. Thus, the second hypothesis is only partially supported.

The fact that no significant relationship was found between goal orientation and student online learning behavior is surprising, as previous research (Miller et al., 2021; Pintrich, 2004; Zhang & Liu, 2019) suggested that goal orientation and the motivational dimension of SRL in general should influence the subsequent performance phase: the level of student activity in the online learning environment or their engagement in actual learning. At the same time, no significant association with learning satisfaction was found for either intrinsic or extrinsic goal orientation, supporting the explanation that the main reason for these findings may be the way in which student online behavior was measured in our study. In a study by Miller et al. (2021), which found a significant relationship between mastery-approach orientation and a variety of learning engagement indicators, the engagement indicators were designed as self-report measures. On the other hand, a study by Zhang & Liu's (2019), like ours, worked with indicators based on digital traces (e.g., number of logins, assignments submitted, number of posts) and found a significant effect of goal orientation on learning engagement. Thus, it seems that this relationship between goal orientation and student online learning behavior and engagement requires further detailed research to uncover which specific indicators of learning behavior are affected by student goal orientation.

The last tested hypothesis focused on the relationship between online learning behavior of students and their course satisfaction. For this relationship, we predicted a positive association: that higher student activity and learning engagement in the online learning environment would be associated with higher satisfaction with the studied course. However, our results showed no statistically significant relationship between the observed indicators of student learning behavior and their learning satisfaction. That is, the number of visits, the regularity of visits, and the total time spent on the course do not appear to be related to course satisfaction.

This finding was very surprising to us, given that a number of earlier studies (El-Sayad et al., 2021; Gao et al., 2020; Murillo-Zamorano et al., 2019; Rajabalee & Santally, 2021) reported that learning behavior and learning engagement were expected to be significant predictors of learning satisfaction. On the other hand, a closer look at the previous studies on this topic reveals that, in the vast majority of cases, the studies measured student perceived engagement using a self-report method (i.e., a questionnaire) rather than proxy indicators that capture actual student behavior in online courses (i.e., indicators based on digital traces in the online learning environment). Thus, it appears that existing research on the relationship between student learning behavior and learning satisfaction may be largely influenced by the measurement approach used. Therefore, one might be inclined to support the views of some researchers (Winne, 2010, 2017; Zeidner & Stoeger, 2019) who expressed concerns that self-reports and questionnaires capture student learning preferences rather than their actual learning behavior. These views, as well as the results of our study, support the thesis that it is necessary to pay more attention to the use of digital traces in researching student learning behavior and engagement in online learning environments, as indicators based on digital traces seem to measure something different from traditional self-report measures.

4.1 Limitations and future research

The research presented in this study has several limitations that need to be addressed and taken into account when interpreting the findings. The main set of limitations arises from the fact that the data analyzed in this study relate to university courses that are blended by design. That is, only part of the teaching of individual courses takes place in an online learning environment; the other part of the teaching takes place in the "traditional" setting, in the form of either face-to-face lectures or seminars. Relatedly, each course may combine traditional and online teaching to different degrees and in different ways, which can obviously have a significant impact on the resulting student satisfaction with the course and (perhaps most importantly) on student behavior in each course in the online learning environment. While we accounted for this important part of the variability (i.e., course-level variability) in the analyses we conducted by using multilevel modelling, we did not use any additional second-level variables within the individual models that might reveal the influence of course-level differences on the relationships examined between the student-level variables. Future research could enrich the models we present with relevant course-level variables to test whether these variables moderate the relationships examined in this study.

Another limitation of the study is that we only focused on three indicators of learning behavior, which should be understood as proxy indicators that obviously cannot fully capture student learning behavior in an online learning environment. Future research could focus both on a wider range of proxy indicators of student learning behavior and on the development and use of more sophisticated methods of investigating and measuring student online learning behavior.

Last but not least, the study sample can be considered as a limitation of the study. Although we were able to collect a sample of 882 students studying in 76 different courses, the sample we obtained has some limitations. For example, only those students who were willing to complete our questionnaire were included in the sample. This means that we only have data from a subset of students from each course, which may introduce some bias into the analyzed data. The same applies to the courses in our sample, as only those courses whose instructors were willing to cooperate with our research and provide us with access to their courses were included in our sample.

Funding

This work was supported by the Czech Science Foundation [grant number 21-08218S, Multimodal learning analytics to study self-regulated learning processes within learning management systems].

References

Amrita Kaur, M. N., & Awang-Hashim, R. (2018). The role of goal orientations in students' perceptions of classroom assessment in higher education. *Assessment & Evaluation in Higher Education*, *43*(3), 461–472. https://doi.org/10.1080/02602938.2017.1359818

Ballouk, R., Mansour, V., Dalziel, B., McDonald, J., & Hegazi, I. (2022). Medical students' self-regulation of learning in a blended learning environment: A systematic scoping review. *Medical Education Online*, *27*(1), 2029336. https://doi.org/10.1080/10872981.2022.2029336

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

Barnard, L., Paton, V., & Lan, W. (2008). Online self-regulatory learning behaviors as a mediator in the relationship between online course perceptions with achievement. *The International Review of Research in Open and Distributed Learning*, *9*(2). https://doi.org/10.19173/irrodl.v9i2.516

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, *27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007

El-Sayad, G., Saad, N. H. M., & Thurasamy, R. (2021). How higher education students in Egypt perceived online learning engagement and satisfaction during the COVID-19 pandemic. *Journal of Computers in Education*, *8*, 527–550. https://doi.org/10.1007/s40692-021-00191-y

Gao, B. W., Jiang, J., & Tang, Y. (2020). The effect of blended learning platform and engagement on students' satisfaction—the case from the tourism management teaching. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *27*, 100272. https://doi.org/10.1016/j.jhlste.2020.100272

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using mplus*. Taylor & Francis.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Huang, J.-C. (2023). Implementation effect of integrating cooperative inquiry into blended learning: Analysis of students' goal setting, task value, and well-being. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2023.2205896

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, *64*, 923–938. https://doi.org/10.1007/s11423-016-9477-y

Jo, I.-H., Kim, D., & Yoon, M. (2015). Constructing proxy variables to measure adult learners' time management strategies in LMS. *Journal of Educational Technology & Society*, *18*(3), 214–225. Retrieved from http://www.jstor.org/stable/jeductechsoci.18.3.214

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, *19*, 141–184. https://doi.org/10.1007/s10648-006-9012-5

Kasser, T., & Ryan, R. M. (1996). Further examining the american dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, *22*(3), 280–287. https://doi.org/10.1177/0146167296223006

Ke, F., & Kwak, D. (2013). Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction. *Computers & Education*, *61*, 43–51. https://doi.org/10.1016/j.compedu.2012.09.003

Kim, D., Park, Y., Yoon, M., & Jo, I.-H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *The Internet and Higher Education*, *30*, 30–43. https://doi.org/10.1016/j.iheduc.2016.03.002

Kim, D., Yoon, M., Jo, I.-H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in south korea. *Computers & Education*, *127*, 233–251. https://doi.org/10.1016/j.compedu.2018.08.023

Klein, H. J., Noe, R. A., & Wang, C. (2006). Motivation to learn and course outcomes: The impact of delivery mode, learning goal orientation, and perceived barriers and enablers. *Personnel Psychology*, *59*(3), 665–702. https://doi.org/10.1111/j.1744-6570.2006.00050.x

Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015a). Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, *2*(3), 81–110. https://doi.org/10.18608/jla.2015.23.6

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015b). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, *27*, 74–89. https://doi.org/10.1016/j.iheduc.2015.06.002

Lee, J. Q., McInerney, D. M., Liem, G. A. D., & Ortiga, Y. P. (2010). The relationship between future goals and achievement goal orientations: An intrinsic–extrinsic motivation perspective. *Contemporary Educational Psychology*, *35*(4), 264–279. https://doi.org/10.1016/j.cedpsych.2010.04.004

Lee, S. J., Srinivasan, S., Trail, T., Lewis, D., & Lopez, S. (2011). Examining the relationship among student perception of support, course satisfaction, and learning outcomes in online learning. *The Internet and Higher Education*, *14*(3), 158–163. https://doi.org/10.1016/j.iheduc.2011.04.001

Lindfors, E. (2021). Students' goal orientations during a pedagogical innovation process: A study in craft, design and technology teacher education. *Techne Serien – Forskning i Slöjdpedagogik Och Slöjdvetenskap*, *28*(2), 221–232. Retrieved from https://journals.oslomet.no/index.php/techneA/article/view/4381

Ma, L., & She, L. (2023). Self-regulated learning and academic success in online college learning. *The Asia-Pacific Education Researcher*. https://doi.org/10.1007/s40299-023-00748-8

Miller, A. L., Fassett, K. T., & Palmer, D. L. (2021). Achievement goal orientation: A predictor of student engagement in higher education. *Motivation and Emotion*, *45*, 327–344. https://doi.org/10.1007/s11031-021-09881-7

Murillo-Zamorano, L. R., López Sánchez, J. Á., & Godoy-Caballero, A. L. (2019). How the flipped classroom affects knowledge, skills, and engagement in higher education: Effects on students' satisfaction. *Computers & Education*, *141*, 103608. https://doi.org/10.1016/j.compedu.2019.103608

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

Pintrich, P. R. (2000). Chapter 14 - the role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies questionnaire (MSLQ)*. University of Michigan.

Posit team. (2023). *RStudio: Integrated development environment for R*. Posit Software, PBC. Retrieved from http://www.posit.co/

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, *26*, 2623–2656. https://doi.org/10.1007/s10639-020-10375-1

Sánchez-Cardona, I., Ortega-Maldonado, A., Salanova, M., & Martínez, I. M. (2021). Learning goal orientation and psychological capital among students: A pathway to academic satisfaction and performance. *Psychology in the Schools*, *58*(7), 1432–1445. https://doi.org/10.1002/pits.22505

Schwam, D., Greenberg, D., & Li, H. (2021). Individual differences in self-regulated learning of college students enrolled in online college courses. *American Journal of Distance Education*, *35*(2), 133–151. https://doi.org/10.1080/08923647.2020.1829255

She, L., Ma, L., Jan, A., Sharif Nia, H., & Rahmatpour, P. (2021). Online learning satisfaction during COVID-19 pandemic among Chinese university students: The serial mediation model. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.743936

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications.

Soyer, M. K., & Kirikkanat, B. (2019). Undergraduates' achievement goal orientations, academic self-efficacy and hope as the predictors of their learning approaches. *European Journal of Educational Research*, *8*(1), 99–106. https://doi.org/10.12973/eu-jer.8.1.99

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, *122*, 119–135. https://doi.org/10.1016/j.compedu.2018.03.018

Wilson, A., Watson, C., Thompson, T. L., Drew, V., & Doyle, S. (2017). Learning analytics: Challenges and limitations. *Teaching in Higher Education*, *22*(8), 991–1007. https://doi.org/10.1080/13562517.2017.1332026

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, *45*(4), 267–276. https://doi.org/10.1080/00461520.2010.517150

Winne, P. H. (2013). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement* (pp. 145–178). Routledge.

Winne, P. H. (2017). Learning Analytics for Self-Regulated Learning. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The Handbook of Learning Analytics* (1st ed., pp. 241–249). Society for Learning Analytics Research (SoLAR).

Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, *2*(2), 5–13. https://doi.org/10.18608/jla.2015.22.2

Wong, J., Baars, M., Davis, D., Zee, T. V. D., Houben, G.-J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction*, *35*(4-5), 356–373. https://doi.org/10.1080/10447318.2018.1543084

Zeidner, M., & Stoeger, H. (2019). Self-Regulated Learning (SRL): A guide for the perplexed. *High Ability Studies*, *30*(1–2), 9–51. https://doi.org/10.1080/13598139.2019.1589369

Zhang, S., & Liu, Q. (2019). Investigating the relationships among teachers' motivational beliefs, motivational regulation, and their learning engagement in online professional learning communities. *Computers & Education*, *134*, 145–155. https://doi.org/10.1016/j.compedu.2019.02.013

Zhang, Y., Zhang, J., & Li, J. (2018). The effect of intrinsic and extrinsic goals on work performance: Prospective and empirical studies on goal content theory. *Personnel Review*, *47*(4), 900–912. https://doi.org/10.1108/PR-03-2017-0086

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 233–262). Cambridge University Press. https://doi.org/10.1017/CBO9780511615771.009

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge/Taylor & Francis Group.

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-3

DECODING STUDENT SUCCESS IN HIGHER EDUCATION: A COMPARATIVE STUDY ON LEARNING STRATEGIES OF UNDERGRADUATE AND GRADUATE STUDENTS

Ricardo Santos[a], Roberto Henriques[a]

[a] NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal

ABSTRACT

Learning management systems (LMS) provide a rich source of data about the engagement of students with courses and their materials that tends to be underutilized in practice. In this paper, we use data collected from the LMS to uncover learning strategies adopted by students and compare their effectiveness. Starting from a sample of over 11,000 enrollments at a Portuguese information management school, we extracted features indicative of self-regulated learning (SRL) behavior from the associated interactions. Then, we employed an unsupervised machine learning algorithm (k-means) to group students according to the similarity of their patterns of interaction. This process was conducted separately for undergraduate and graduate students. Our analysis uncovered five distinct learning strategy profiles at both the undergraduate and graduate levels: 1) active, prolonged and frequent engagement; 2) mildly frequent and task-focused engagement; 3) mildly frequent, mild activity in short sessions engagement; 4) likely procrastinators; and 5) inactive. Mapping strategies with the students' final grades, we found that students at both levels who accessed the LMS early and frequently had better outcomes. Conversely, students who exhibited procrastinating behavior had worse end-of-course grades. Interestingly, the relative effectiveness of the various learning strategies was consistent across instruction levels. Despite the LMS offering an incomplete and partial view of the learning processes students employ, these findings suggest potentially generalizable relationships between online student behaviors and learning outcomes. While further validation with new data is necessary, these connections between online behaviors and performance could guide the development of personalized, adaptive learning experiences.

KEYWORDS

self-regulated learning; student strategies; learning management systems; higher education; machine learning

CORRESPONDING AUTHOR
Ricardo Santos, NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, 1070-312, Lisbon, Portugal
e-mail: rcsantos@novaims.unl.pt

Introduction

In the evolving landscape of education, the focus has shifted toward individual student progress, significantly altering the dynamics of teaching and learning. This transformation is largely driven by the advent of artificial intelligence tools, leading to substantial investments in personalized learning and intelligent tutoring systems (Holmes & Tuomi, 2022). Despite these advancements, traditional educational methods continue to hold relevance, even as educators grapple with challenges such as larger class sizes and the rise of remote learning, thus reducing the reliability of conventional ways of assessing student progress, such as attendance and in-class behavior (Bellur et al., 2015). These changes challenge educators to identify and support students who require the most assistance.

Consequently, there is a growing emphasis on self-regulated learning (SRL) behaviors, which provide a more comprehensive insight into a student's abilities, motivations, and attitudes toward learning. SRL skills are crucial for students, particularly in higher education where autonomy is expected (Boekaerts, 1997; Broadbent & Poon, 2015) and in the 21st century workplace, where employers prioritize learners who can take charge of their development (Trilling & Fadel, 2009). Thus, gauging and fostering the development of SRL behavior is imperative in both educational and professional settings.

The cyclical model of SRL involves students actively participating in their learning through cycles of forethought, performance control, and self-reflection (Zimmerman, 2000, 2002). Students develop tools to regulate their cognition, behavior, and emotions through repeated engagement in these processes (Zimmerman & Moylan, 2009). A key component of SRL is the development of strategies that enhance students' ability to achieve their learning goals. According to Pintrich et al. (1991), SRL strategies can be divided into three categories: cognitive, metacognitive, and resource management. Both time management and effort regulation are positively correlated with student performance (Broadbent, 2017; Puzziferro, 2008). Conversely, evidence suggests that students with underdeveloped SRL behaviors struggle in contexts where more autonomy is expected, such as in online and blended learning contexts (Broadbent, 2017).

One approach to measure SRL behaviors is direct observation of the students. For example, timing how long it takes for a student to finish a set of tasks provides behavioral evidence of the SRL trait of time management (Winne & Jamieson-Noel, 2002). However, comprehensively observing student learning behaviors through direct means can be difficult in practice, as designing rigorous experiments in controlled settings requires extensive time and resources (Susac et al., 2014). Alternatively, self-report questionnaires, such as the Motivated Strategies for Learning Questionnaire (MSLQ), allow students to self-evaluate SRL traits (Pintrich et al., 1991; Winne & Perry, 2000). These questionnaires are inexpensive and simple to administer, but sole dependence on student self-reports poses risks of bias and only reflects students' perceptions at the time of administration.

In recent years, the widespread adoption of learning management systems (LMS) in higher education institutions has increased the availability of detailed student trace data (Coates et al., 2005). These systems record students' digital interactions within their learning environment. By applying data mining techniques to these logs, researchers can extract variables (from this point onward referred to as features) connected to SRL behaviors (Baker et al., 2020). These features can be used to gain additional insights about learners, the learning processes they engage in, and their academic progress. For example, supervised machine learning algorithms have been successful at flagging students at risk of failing (Bernacki et al., 2020; Macfadyen & Dawson, 2010; Riestra-González et al., 2021). Alternatively, unsupervised machine learning algorithms (also referred to as clustering algorithms) can be used to uncover learner strategy profiles (Cerezo et al., 2016; Riestra-González et al., 2021).

While prior works have utilized unsupervised machine learning to identify learning behaviors from LMS data, a limited number of studies apply these approaches, especially for large, multi-course samples. Moreover, exploring possible differences and effectiveness of learning strategies across different instruction levels is still a relatively unexplored topic. This work aims to address these gaps by leveraging clickstream data to extract course-agnostic features from an LMS, identify learner strategy profiles at the undergraduate and graduate levels, and assess their relative effectiveness for academic success. The research questions are:

1. What course-agnostic learning strategy profiles can be extracted from undergraduate and graduate students' SRL features extracted from LMS data?
2. What is the relationship between the learning strategies uncovered by k-means and end-of-course performance at each instruction level?
3. Are there differences in the effectiveness of the learning strategies between instruction levels?

To answer these questions, Moodle logs were collected from 57 undergraduate and 124 graduate courses taught at a Portuguese information management school during the 2020/2021 academic year. From these logs, 30 SRL features were extracted to build a dataset, which was then split between undergraduate and graduate course enrollments. The *k*-means clustering algorithm was used to identify learner strategy profiles at each instruction level, allowing the comparison of the effectiveness of each strategy.

The remainder of this paper is structured as follows: The next section provides an overview of prior research utilizing unsupervised learning approaches to identify learner strategy profiles from LMS data. The third section presents the study's data and methodology. The fourth section presents the results. The fifth section discusses the results, their alignment with expected outcomes, and key implications. The sixth and final section concludes with a summary of the main findings and a discussion of future research directions.

1 Related work

This section provides an overview of research that uses unsupervised machine learning techniques to identify learning strategies from SRL-related features. The main purpose of this section is to discuss the different existing approaches regarding the adoption of theoretical frameworks, sample size, features extracted, the techniques used and the author's main finding when uncovering learning strategies from data. A literature review table featuring all works covered in this section is provided in Table 1.

The theoretical frameworks most frequently cited include Biggs' 3P model (Biggs, 1987) and the SRL motivational model of Pintrich et al. (1991). These theoretical foundations provide a clear interpretive lens for variables derived from LMS data, a solid rationale for the chosen tools, and a frame of reference for interpreting results. For example, Gašević et al. (2017) used the Study Process Questionnaire (SPQ) instrument to supplement LMS data, which enabled them to distinguish between deep and surface learning indicators among their students. They discovered that students who employ deep learning strategies outperform their peers. Li & Tsai (2017) also reported using the MSLQ to uncover SRL variables from their students to map SRL to academic performance. However, most studies reviewed do not delve extensively into a theoretical SRL framework (Cerezo et al., 2016; Moubayed et al., 2020; Riestra-González et al., 2021). Instead, they merely reference existing frameworks to rationalize how LMS data can reveal learning strategies and the reasons behind selecting specific feature types. This trend could be attributed to a greater focus on using these variables to uncover learning strategy profiles from data (Cerezo et al., 2016; Riestra-González et al., 2021) rather than conducting a thorough discussion of how a specific SRL model explains differences in academic performance or achievement. Another potential reason stems from the nature of the data used.

Table 1

Literature review table research on the use of unsupervised learning to uncover learning strategies

Reference	Grounding	Sample	Source	Variables	Clusters	Key Findings
Hung & Zhang, (2008)	Not specified	98 undergrad students	LMS	5 engagement features (e.g. frequency, materials)	*k*-means (3 clusters)	Active students performed better academically
Cerezo et al. (2016)	Not specified	140 undergrad psychology students	LMS	Interactions with LMS materials	*k*-means (5 clusters)	Better performers procrastinate less while also spending less time on the LMS
Gašević et al. (2017)	Biggs‘ 3P model	144 undergrad students	LMS, survey	Deep/surface learning indicators from LMS	Hierarchical clustering (2 clusters)	Deep learning strategies map to deep learning scales on survey
Li & Tsai (2017)	Pintrich‘s motivational model	59 undergrad computer science students	LMS, survey	Time on materials, SRL features	*k*-means (3 clusters)	Consistent use students had higher motivation and achievement
Çebi & Güyer (2020)	Pintrich‘s motivational model	122 undergrad statistics students	LMS	Time on activities	Unspecified clustering (3 clusters)	Engagement associated with better performance and motivation
Matcha et al. (2020)	Biggs‘ 3P model	~1400 students in a MOOC	LMS	Video, quiz, assignment, forum actions	Hierarchical clustering & Markov model (4 clusters)	Active/highly active had better performance
Moubayed et al. (2020)	None specified	486 undergrad students	LMS	Time on activities, assignments	*k*-means (3 clusters)	Highly engaged students tend to perform better
Yang et al. (2020)	None specified	242 undergrad students	LMS	Homework submission patterns	*k*-means (3 clusters)	Non-procrastinators perform better
Riestra-González et al. (2021)	None specified	~16000 students in 699 courses	LMS	31 features (resource accesses)	*k*-means (6 clusters)	More engaged students perform better

While LMS data is rich, it is essentially a series of timestamped actions. Features like click count can be categorized under Pintrich et al.'s (1991) resource management, but they only offer a partial and indirect insight into crucial constructs such as motivation or emotional state, which are prevalent in popular SRL models (Panadero, 2017).

The sample sizes used in these works also differ greatly. They range from a small group of 59 students in a single course (Li & Tsai, 2017) to a large cohort of nearly 16,000 students spread across 699 different courses (Riestra-González et al., 2021). This significant variation limits the ability to derive insights that can be generalized across different contexts. Moreover, although the LMS is a common data source in the reviewed works, the specific features and contexts for variable usage and extraction vary substantially. Several studies track the frequency of specific student actions or the time spent on the LMS (Cerezo et al., 2016; Matcha et al., 2020; Riestra-González et al., 2021). However, different sets of features have been extracted from the LMS, with Yang et al.'s (2020) approach using the LMS to extract and analyze features related to procrastination behaviors on homework deadlines.

In the process of uncovering learning strategies, the most common method is to group students into clusters using k-means or hierarchical clustering algorithms based on the features extracted from the LMS logs (Cerezo et al., 2016; Hung & Zhang, 2008; Moubayed et al., 2020). For example, Hung & Zhang (2008) extracted five LMS engagement features from 98 students in an online course and used k-means to uncover three clusters that differentiated poor-performing versus above-average students. Similarly, Cerezo et al. (2016) identified four learner strategy profiles in a sample of 140 students using k-means on LMS trace data, finding the cluster with socially-focused and strategic study habits achieved the highest grades. Riestra-González et al. (2021) also found significant differences in four out of the six learning strategies uncovered. Beyond k-means, both Gašević et al. (2017) and Matcha et al. (2020) used hierarchical clustering to group similar sets of students. Finally, Çebi & Güyer (2020) did not mention the specific algorithm used in their work despite also using a clustering technique to uncover three distinct learning strategy profiles from LMS data.

Observations from multiple studies have consistently shown that students who exhibit higher engagement and less procrastination tend to achieve better academic results than their peers (Cerezo et al., 2016; Moubayed et al., 2020; Yang et al., 2020). Tactics such as evenly spacing study time and completing assessments early were positively associated with achievement, while students exhibiting low numbers of clicks, and late and infrequent logins tended to perform worse (Hung & Zhang, 2008; Li & Tsai, 2017; Matcha et al., 2020). These findings align with expectations, as students who demonstrate traits related to the employment of an actual strategy are more likely to have more developed SRL skills. However, only a few studies mapped clusters directly back to established SRL frameworks to confirm theoretical connections between engagement and motivation (Çebi & Güyer, 2020; Li & Tsai, 2017). In terms of implications, the findings of these studies point to the potential of analytics tools that aim to provide adaptive interventions and personalized support starting from the student behaviors (Cerezo et al., 2016).

The research discussed in this section illustrates that using unsupervised machine learning techniques to uncover students' learning strategies with LMS data is an active and growing area of study. A common approach is to use clustering algorithms to group students based on their interactions with course materials and activities. These clusters are then associated with academic performance metrics or self-reported surveys to draw connections between learning strategies, motivation, and achievement.

However, there are notable gaps worth highlighting. Small sample sizes are a common issue, and no studies have explicitly sought to identify and compare learning strategies across different levels of instruction. This limits the generalizability of findings and hinders the development of comprehensive models. Additionally, there are inconsistencies in the features used by different authors, partly due to the absence of a consistent theoretical framework for SRL in most works. This leads to disparate findings and interpretations. While addressing this gap is beyond the scope of this work, adopting a robust theoretical framework could lead to more consistent and comparable findings across studies.

This work aims to address some research opportunities by using larger samples and more courses, contributing to more generalizable models. This could help determine if students' learning strategies can be replicated in a general context and inform the design of personalized learning experiences on the LMS, potentially reducing student dropout rates and improving achievement.

2 Methodology

This work started with the extraction of anonymized institutional Moodle logs and their transformation into a structured dataset indexed by program, course, and student, accompanied by 30 features associated with the resource management construct found in Pintrich's motivational model for SRL (Pintrich et al., 1991). The dataset was split into undergraduate and graduate subsets and given to separate instances of the *k-means* clustering algorithm (Macqueen, 1967). The resulting clusters were characterized and compared. A summary of the adopted approach is depicted in Figure 1. Unless otherwise noted, all data manipulation and analysis procedures were implemented using Python (McKinney, 2017) and Scikit-learn (Pedregosa et al., 2011).

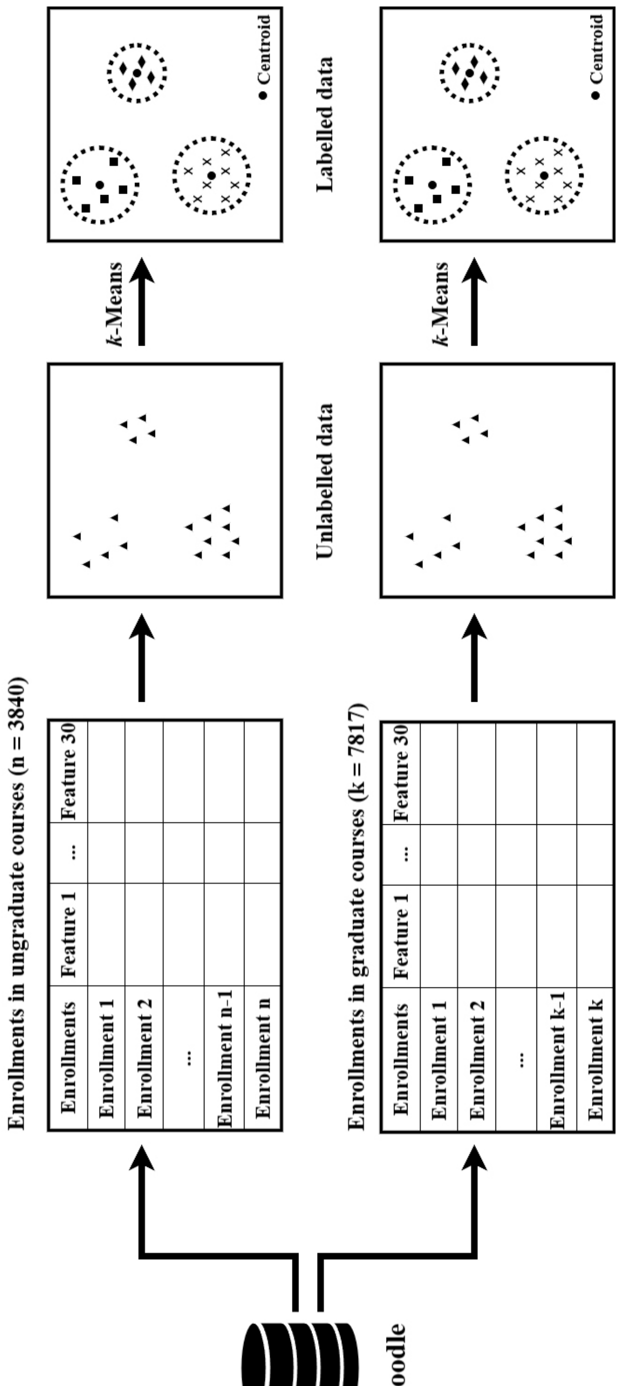


Figure 1
Overview of the approach

2.1 Data

The data was collected from a Portuguese information management school in 2020/2021, which offers graduate and undergraduate programs in data science, information management, and information systems and technologies. The sample includes 1564 graduate and 409 undergraduate students enrolled in 124 and 57 courses, respectively, totaling 11,297 student enrollments. Moodle logs and end-of-course final grades were accessed for each enrollment with no additional data sources being considered. Table 2 presents an overview of the population for each instruction level, including the number of courses, students, enrollments, and average end-of-course performance, which in the Portuguese systems assumes values between 0 and 20, with 10 representing the minimum passing threshold. All student data was anonymized in compliance with GDPR, and the project was approved by the *Ethics Committee and Institutional Review Board* with Code DSCI2022-9-227363.

Table 2
Summary of the characteristics of courses and students per instruction level (grades ranging from 0 to 20)

	Courses	Students	Enrollments	Enrollments per course	Average end-of-course grade (± Standard Deviation)
Undergraduate					
Program A	28	160	1336	47.71	13.48 ± 4.30
Program B	29	249	2144	73.93	14.22 ± 4.02
Sub-total	**57**	**409**	**3480**	**63.87**	**13.94 ± 4.14**
Graduate					
Program 1	10	33	322	32.20	13.68 ± 3.23
Program 2	17	173	755	44.41	13.98 ± 3.52
Program 3	6	31	170	28.33	15.78 ± 2.20
Program 4	6	40	218	36.33	15.10 ± 2.51
Program 5	4	310	120	30.00	16.27 ± 2.36
Program 6	4	27	108	27.00	16.76 ± 1.15
Program 7	9	155	666	74.00	16.50 ± 2.72
Program 8	13	36	391	30.08	15.74 ± 2.68
Program 9	7	82	267	38.14	16.48 ± 1.65
Program 10	2	33	54	27.00	13.07 ± 5.39
Program 11	15	192	1818	121.20	15.73 ± 3.07
Program 12	29	416	2857	98.52	15.67 ± 3.06
Program 13	2	36	71	35.50	15.87 ± 2.65
Sub-total	**124**	**1564**	**7817**	**81.89**	**15.54 ± 3.08**
Total	**181**	**1973**	**11,297**	**72.88**	**15.04 ± 3.52**

2.2 Feature Extraction

The first part of the process involved converting Moodle logs into data structures suitable for statistical analysis. For each course, Moodle keeps a timestamped record of every click made on the LMS, including which student made the click and where it was performed within the LMS. To extract meaningful features from this data, we adopted three perspectives that measure student engagement with the LMS, a critical resource for our students: *Raw activity*, which refers to the number of times a certain action is performed; *Time-on-task*, which refers to the amount of time dedicated to studying on LMS; and *Procrastination*, which measures at which stages of the course the students log into the LMS.

In total, 30 candidate features were extracted and considered for subsequent steps. The reasons for the choice of these specific features are two-fold. First, these features fall under the resource management construct of Pintrich's motivational model for SRL (Panadero, 2017) and measure student interaction with the LMS. Moreover, these features have also been successfully utilized in a plethora of previous learning analytics research (Aljohani et al., 2019; Conijn et al., 2017; Riestra-González et al., 2021; Romero et al., 2013; Santos & Henriques, 2023). Table 3 provides a comprehensive list of the features extracted from the logs and their respective averages and standard deviations for each instruction level.

Table 3
Extracted candidate features

Feature	**N (under-graduate)**	**Mean ± Standard Deviation (undergraduate)**	**N (graduate)**	**Mean ± Standard Deviation (graduate)**
Perspective 1: Raw activity				
Total clicks (n)	3480	279.59 ± 177.69	7817	248.56 ± 168.54
Clicks (% of course total)	3480	1.64 ± 1.01	7817	1.59 ± 1.60
Forum clicks (n)	2222	8.63 ± 13.85	6390	19.62 ± 25.31
Forum posts (n)	27	1.52 ± 1.08	375	1.76 ± 1.13
Discussions viewed (n)	1283	6.47 ± 8.98	5277	11.38 ± 14.47
Folder clicks (n)	1825	20.20 ± 28.40	63.68	20.29 ± 23.64
Resources viewed (n)	3460	64.16 ± 55.39	6928	46.96 ± 38.85
URLs viewed (n)	2419	25.14 ± 16.96	5514	17.65 ± 13.67
Course clicks (n)	3480	110.48 ± 80.62	7816	94.41 ± 66.97
Assessments started (n)	1688	3.02 ± 2.56	3763	2.56 ± 3.16
Assignments viewed (n)	974	17.40 ± 24.88	3618	12.98 ± 17.51

Assignments submitted (n)	893	5.55 ± 5.81	3113	4.90 ± 4.83
Submissions (% of course total)	893	3.36 ± 2.56	3113	2.60 ± 5.14
Perspective 2: Time-on-task				
Online sessions (n)	3480	59.96 ± 41.77	7817	47.92 ± 29.82
Clicks/session (n)	3480	4.96 ± 2.33	7817	5.40 ± 2.80
Clicks/day (n)	3480	1.85 ± 1.18	7817	2.02 ± 1.43
Total time online (min)	3480	491.71 ± 394.82	7817	396.98 ± 331.03
Aver. duration of online sessions (min)	3480	8.06 ± 3.85	7817	8.20 ± 5.64
Perspective 3: Procrastination				
Largest period of inactivity (h)	3480	463.88 ± 283.71	7817	415.60 ± 269.13
Days with 0 clicks (% of period)	3480	62.95 ± 11.87	7817	63.92 ± 11.88
PercCourse_1Login	3480	7.06 ± 9.32	7817	0.61 ± 8.52
PercCourse_NLogin ($n \in [2, 9]$)	…	…	…	…
PercCourse_10Login	3387	22.16 ± 15.10	7538	22.44 ± 19.45

2.3 Data analysis

The data for graduate and undergraduate students were processed separately but followed similar pipelines for preprocessing, feature selection, and clustering. The preprocessing stage involved three main steps. In the first step, the Jarque-Bera normality test (Jarque & Bera, 1980) was used to assess how reasonable it would be to assume the normal distribution of the data. This test measures the skewness and kurtosis of a feature and determines if it deviates significantly from those of a normal distribution (skewness of 0 and kurtosis of 3). In the second step, all features that could not be reasonably assumed to follow a normal distribution were transformed using the Yeo-Johnson power transformation (Yeo & Johnson, 2000). This method aims to transform non-normally distributed data into a shape resembling a normal distribution by raising the data to an appropriate power. The transformed variables were then standardized, which is the final step of the preprocessing stage. This rigorous preprocessing ensures that the data is appropriately conditioned for the subsequent stages of feature selection and clustering.

The feature selection process aimed to eliminate any variable that could be considered irrelevant or redundant for cluster construction from each perspective. This was achieved through a two-step strategy. The first step involved setting an absolute value of 0.8 on the Spearman correlation index to flag potentially redundant variables. In the second step, *k*-means was used to create clustering solutions for each perspective, and the explained variance of each feature toward that solution was measured. Variables with very low explained variance (i.e. irrelevant variables) were removed, as were redundant features that exhibited the lowest explained variance. This process was repeated until a satisfactory clustering solution was achieved for each perspective. The resulting variables were then combined into a final dataset. Consequently, at the end of this stage, there were two preprocessed datasets: one containing the features necessary to build clusters on undergraduate enrollments, and another containing the features deemed relevant for clustering graduate enrollments.

In the third stage, each dataset was used as input to a separate instance of the *k*-means clustering algorithm. *k*-means is an iterative algorithm that groups data points based on distance, minimizing within-group distance while maximizing between-group distances. A key component of *k*-means is the concept of a centroid, which can be understood as a data point representing the coordinates of the center a group. By comparing the positions of these centroids, it is possible to understand the differences and similarities between the groups. Despite its simplicity, *k*-means enjoys widespread adoption when partitioning data into different groups (Wu et al., 2008). However, a limitation of *k*-means is that the number of resulting groups must be set *a priori*. In this implementation, the optimal number of groups (each referring to a learning strategy) was determined using the elbow method (Cerezo et al., 2016; Riestra-González et al., 2021) and found to be five for both instruction levels.

Once the groups were formed, they were analyzed to answer the research questions. To answer the first research question, the different learning strategies were characterized. This involved comparing the strategies adopted by students at the same instruction level to ensure there were no overlaps. The differences between learning strategies were measured by comparing the coordinates of the centroids determined by *k*-means. Between-group comparisons were performed at the feature level but interpreted at the perspective level. Two learning strategies were considered significantly different in one perspective if there were statistically significant differences in most variables belonging to that perspective. Due to the differences in scale, these comparisons were performed using standardized scores (0 mean and unit variance).

To answer the second research question, the average end-of-course grade associated with each learning strategy was calculated. This was followed by a comparison of the end-of-course grade of the various learning strategies at the same instruction level using Welch's t-test.

To answer the third and final research question, we performed a qualitative comparison of the learning strategies adopted by undergraduate students with those adopted by graduate students. The aim was to identify whether there were unique undergraduate or graduate-level strategies that did not exist at the other level of instruction. Moreover, the comparison also aimed to identify whether the relative effectiveness of strategies varied between the two instruction levels.

3 Results

3.1 Learning strategies in undergraduate and graduate students

The centroid coordinates presented in Table 4 show that all five resulting learning strategies differ significantly from one another regarding the *Raw activity* and *Time-on-task* perspectives, with strategies B and E not being significantly different when it comes to *Procrastination*.

From the perspective of *Raw activity*, students adopting different strategies exhibited varying levels of engagement with Moodle. Strategy D students exhibited the highest overall levels of engagement, with the highest number of clicks, both overall and across multiple pages, including resources, external links, and course page visits. Strategy C students had the second highest average engagement across most raw activity features, ranking highest in folder clicks and assessments started. In contrast, Strategy E students displayed the lowest raw activity engagement, with the least clicks across all features measured. Strategy A engagement was also relatively low, with all raw activity metrics falling below or slightly above average. Finally, while generally a low activity strategy, Strategy B students completed a relatively high number of assessment starts compared to other low engagement strategies.

Table 4

K-means standardized mean ± standard deviation for all variables in clustering for undergraduate enrollments (values with statistically significant (p-value < 0.05 on t-test) differences against all other strategies are in bold)

Feature	Strategy A (n =883)	Strategy B (n=615)	Strategy C (n=735)	Strategy D (n=702)	Strategy E (n=545)
Perspective 1: Raw activity					
Total clicks (n)	**−0.28 ± 0.39**	**−0.41 ± 0.49**	**0.79 ± 0.62**	**1.34 ± 0.84**	**−1.60 ± 0.53**
Folder clicks (n)	−0.29 ± 0.92	**0.05 ± 0.92**	**0.93 ± 0.77**	−0.25 ± 0.94	**−0.51 ± 0.71**
Resources viewed (n)	**0.08 ± 0.76**	**−0.47 ± 0.78**	**−0.07 ± 0.86**	**1.23 ± 0.84**	**−0.98 ± 0.79**
URLs viewed (n)	−0.04 ± 0.97	−0.21 ± 0.86	−0.12 ± 1.05	**0.85 ± 0.74**	**−0.68 ± 0.68**
Course clicks (n)	**0.02 ± 0.51**	**−0.67 ± 0.57**	**0.38 ± 0.57**	**1.45 ± 0.92**	**−1.49 ± 0.68**
Assessments started (n)	**−0.70 ± 0.53**	0.29 ± 0.84	**0.96 ± 0.65**	0.38 ± 1.10	**−0.88 ± 0.24**
Perspective 2: Time-on-task					
Online sessions (n)	**0.15 ± 0.48**	**−0.87 ± 0.55**	**0.31 ± 0.47**	**1.47 ± 0.97**	**−1.42 ± 0.72**
Clicks/session (n)	**−0.60 ± 0.66**	**0.85 ± 0.82**	**0.74 ± 0.74**	**−0.05 ± 0.88**	**−0.86 ± 1.02**
Clicks/day (n)	**−0.29 ± 0.41**	**−0.42 ± 0.50**	**0.80 ± 0.60**	**1.32 ± 0.78**	**−1.58 ± 0.46**
Total time online (min)	**−0.19 ± 0.53**	**−0.44 ± 0.48**	**0.56 ± 0.51**	**1.44 ± 0.87**	**−1.57 ± 0.53**
Average duration of online sessions (min)	**−0.39 ± 0.72**	0.58 ± 1.06	0.54 ± 0.74	**0.43 ± 0.83**	**−1.19 ± 0.90**
Perspective 3: Procrastination					
Largest period of inactivity (h)	**−0.22 ± 0.76**	0.85 ± 1.08	**−0.14 ± 0.81**	**−0.95 ± 0.96**	0.67 ± 2.16
Days with 0 clicks (% of period)	**−0.28 ± 0.64**	1.07 ± 0.62	**−0.13 ± 0.62**	**−1.24 ± 0.76**	1.01 ± 0.82
PercCourse_1Login	0.09 ± 1.05	0.09 ± 0.97	−0.19 ± 0.81	−0.20 ± 1.03	**0.26 ± 1.13**

Similar trends were observable for the *Time-on-task* and *Procrastination* perspectives, with some key exceptions. Aligned with their raw activity totals, Strategy D students spent the most time on the LMS, logged the highest number of sessions, and started accessing the system as early as possible, displaying low procrastination tendencies. Mirroring their overall inactivity, Strategy E students spent the least amount of time on the LMS, had the fewest sessions, and tended to start accessing the system later than the others. Strategies A and C again fell in between. Finally, the behavior displayed by students who adopted Strategy B was somewhat different. Their values on features related to *Procrastination* showed that they displayed values that were statistically similar to the highly inactive Strategy E students. However, there were some divergences between the *Raw activity* and *Time-on-task* perspectives as these students exhibited long sessions and the highest number of clicks per session of all learning strategies uncovered for undergraduate enrollments.

To facilitate interpretation, the strategies were labeled based on these engagement characteristics. Strategy A was termed *mildly frequent, mild activity in short sessions*, Strategy B *likely procrastinators*, Strategy C *mildly frequent and task-focused*, Strategy D *active, prolonged and frequent* and Strategy E *inactive*.

Table 5 presents the centroid coordinates for the five learning strategies uncovered for graduate students. A key difference between undergraduate and graduate enrollments is that forum clicks and assessments viewed impacted cluster construction for graduate students when they had provided little explanatory power for undergraduates. All five graduate learning strategies show significant differences across all perspectives.

From the perspective of *Raw activity,* students adopting Strategy 5 were the most engaged with Moodle materials, presenting the highest values for total clicks, clicks on course-related and resource pages, and assessments viewed. In contrast, students adopting Strategy 1 had the lowest levels of engagement across most features. The remaining strategies presented engagement values somewhere in between: Strategy 2 tended toward higher levels of engagement on most features; Strategy 4 tended toward lower values for total clicks but had high values for clicks on resources, external URLs, and assessment views; and Strategy 3 had close to average total clicks with high values for folder clicks and assessments started.

Table 5

K-means standardized mean ± standard deviation for all variables used in clustering of graduate enrollments (values with statistically significant (p-value < 0.05 on t-test) differences against all other groups are in bold)

Feature	**Strategy 1 (n =1381)**	**Strategy 2 (n=1697)**	**Strategy 3 (n=1503)**	**Strategy 4 (n=2093)**	**Strategy 5 (n=1143)**
Perspective 1: Raw activity					
Total clicks (n)	**−1.55 ± 0.55**	**1.02 ± 0.73**	**−0.02 ± 0.58**	**−0.31 ± 0.41**	**1.18 ± 0.93**
Folder clicks (n)	**−0.63 ± 0.80**	**0.66 ± 0.84**	**0.18 ± 0.88**	**−0.28 ± 0.97**	**0.04 ± 1.04**
Resources viewed (n)	**−0.94 ± 0.69**	0.16 ± 1.07	**−0.36 ± 0.80**	0.13 ± 0.75	**1.06 ± 0.85**
URLs viewed (n)	**−0.80 ± 0.36**	0.23 ± 0.73	**−0.35 ± 0.69**	0.23 ± 0.61	0.49 ± 0.99
Course clicks (n)	**−1.48 ± 0.69**	**0.79 ± 0.60**	**−0.41 ± 0.53**	**−0.11 ± 0.52**	**1.37 ± 0.92**
Assessments started (n)	**−0.84 ± 0.36**	0.87 ± 0.74	0.87 ± 0.69	**−0.67 ± 0.61**	**−0.08 ± 0.99**
Assessments viewed (n)	−0.51 ± 0.72	**−0.26 ± 0.85**	−0.52 ± 0.67	**0.43 ± 1.02**	**0.83 ± 0.93**
Forum clicks	**−0.41 ± 1.01**	0.54 ± 0.89	**−0.15 ± 0.97**	**−0.32 ± 0.87**	0.49 ± 0.91
Perspective 2: Time-on-task					
Online sessions (n)	**−1.35 ± 0.79**	**0.65 ± 0.56**	**−0.59 ± 0.55**	**−0.01 ± 0.55**	**1.40 ± 0.87**
Clicks/session (n)	**−1.01 ± 1.09**	**0.59 ± 0.82**	**0.99 ± 0.87**	**−0.38 ± 0.61**	**−0.12 ± 0.75**
Clicks/day (n)	**−1.51 ± 0.56**	**0.74 ± 0.63**	**−0.08 ± 0.68**	**−0.23 ± 0.50**	**1.33 ± 0.78**
Total time online (min)	**−1.48 ± 0.60**	**0.89 ± 0.72**	−0.21 ± 0.54	−0.18 ± 0.55	**1.24 ± 0.95**
Aver. duration of online sessions (min)	**−1.04 ± 1.16**	0.56 ± 1.01	0.56 ± 0.99	**−0.19 ± 0.73**	**0.17 ± 0.79**
Perspective 3: Procrastination					
Largest period of inactivity (h)	**0.31 ± 1.84**	**0.05 ± 0.74**	**0.67 ± 1.00**	**−0.12 ± 0.92**	**−1.02 ± 0.81**
Days with 0 clicks (% of period)	**0.48 ± 1.16**	**−0.15 ± 0.58**	**0.79 ± 0.76**	**−0.04 ± 0.79**	**−1.36 ± 0.68**
PercCourse_1Login	**0.55 ± 1.27**	−0.10 ± 0.68	**0.34 ± 0.91**	−0.06 ± **0.89**	**−0.68 ± 1.07**

As for the remaining perspectives, most of the results are consistent with the observations for undergraduate students for most strategies. Students with the highest level of activity (Strategy 5) presented the highest values for the *Time-on-task* perspective and the lowest for the *Procrastination* perspective. Likewise, the least engaged students (Strategy 1) consistently had the lowest values concerning *Time-on-task* and relatively high values in features in *Procrastination*. In learning Strategy 3, students adopting it were characterized by high levels in *Procrastination*, having the longest periods of inactivity and the greatest number of days without any activity. Although these students accessed Moodle infrequently, when they did, they tended to have long and click-intensive sessions. Despite having long sessions, they had a low number of sessions overall and spent less total time on Moodle.

Again, to facilitate interpretation, the strategies were labeled based on these engagement characteristics in a manner similar to the labels attributed to the undergraduate students. Strategy 1 was labelled *inactive*, Strategy 2 *mildly frequent and task-focused*, Strategy 3 *likely procrastinators*, Strategy 4 *mildly frequent, mild activity in short sessions* and Strategy 5 *active, prolonged and frequent*.

3.2 End-of-course performance for undergraduate and graduate students

The main focus of this second section was the exploration of the relationship between various learning strategies and student performance. A Welch's t-test was employed to compare the average end-of-course performance of each learning strategy against all others within the same level of instruction (Table 6). The analysis revealed significant differences in performance among the learning strategies identified by *k*-means clustering.

Specifically, three out of the five strategies showed a significant difference from all others in undergraduate enrollments. Strategies A (characterized by moderate frequency and activity in short sessions) and C (moderate frequency and task-focused) were not significantly distinct from each other, but they were significantly different from all other strategies (p-value = 0.14).

Table 6

*Pairwise comparison of the statistics and p-values obtained for the Welch's t-tests comparing the end of course grades obtained by each learning strategy (cells with p-value < 0.05 identified with *)*

Undergraduate learning strategies								
	Strategy A		Strategy B		Strategy C		Strategy D	
	t-stat	p-value	t-stat	p-value	t-stat	p-value	t-stat	p-value
Strategy A (n = 883)								
Strategy B (n = 665)	6.50	1.21^{-10}*						
Strategy C (n = 735)	1.46	0.14	−4.65	$3.75e^{-6}$*				
Strategy D (n = 702)	−2.41	0.02*	−8.17	$7.96e^{-16}$*	−3.48	$5.11e^{-4}$*		
Strategy E (n = 545)	7.45	$2.60e^{-13}$*	2.50	0.01*	6.11	$1.44e^{-9}$*	8.71	$1.68e^{-17}$*
Graduate learning strategies								
	Strategy 1		Strategy 2		Strategy 3		Strategy 4	
	t-stat	p-value	t-stat	p-value	t-stat	p-value	t-stat	p-value
Strategy 1 (n = 1381)								
Strategy 2 (n = 1697)	−4.36	$1.34e^{-5}$*						
Strategy 3 (n = 1503)	0.18	0.86	5.34	$1.02e^{-5}$*				
Strategy 4 (n = 2093)	−9.07	$2.49e^{-19}$*	−6.10	$1.17e^{-09}$*	−11.00	$1.26e^{-27}$*		
Strategy 5 (n = 1143)	−10.68	$4.29e^{-36}$*	−8.26	$2.26e^{-26}$*	−12.46	$1.16e^{-34}$*	−3.22	1.29e-^{34}*

A closer look at the performance of students who adopted each strategy (Figure 2) provides more insights. Students who adopted Strategy D (*active, prolonged, and frequent engagement*) achieved the highest average grade of 14.85 (± 3.29). They were closely followed by students employing Strategies A and C, with average grades of 14.46 (± 3.26) and 14.19 (± 3.94), respectively. On the other hand, students using Strategy B (*likely procrastinators*) had the second-lowest average grades (13.17 ± 4.09). Notably, students who adopted Strategy E (*inactive*), despite some exceptions indicated by the high standard deviation, generally achieved lower grades (12.41 ± 5.85) than their peers using other strategies.

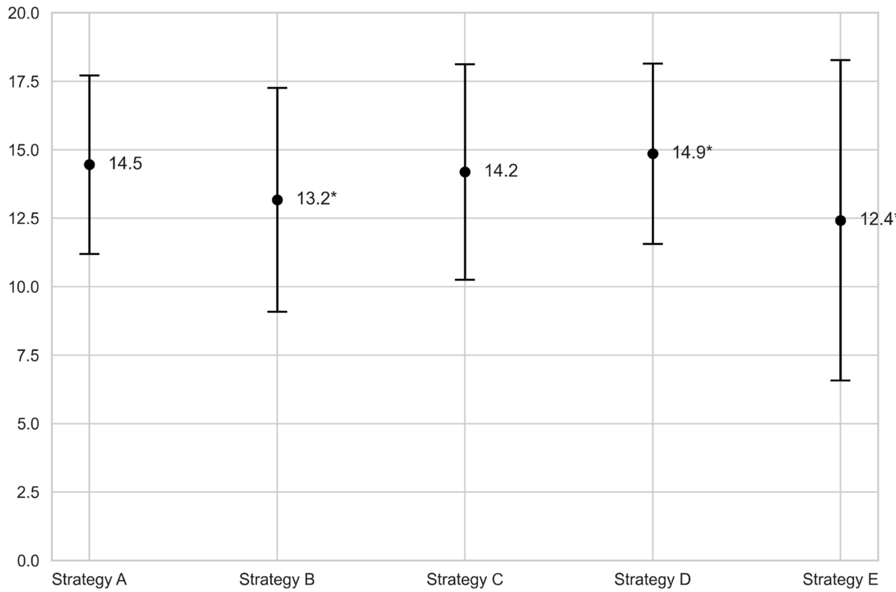


Figure 2
Average and standard deviation of the end-of-course grade for undergraduate learning strategies (values that exhibit statistically significant differences against all other groups are identified with an asterisk)

In the case of graduate students, three of the learning strategies were found to be significantly different from all others, with strategies 1 (*inactive*) and 3 (*likely procrastinators)* not showing significant distinction from each other (p-value = 0.86) while being significantly different from the remaining strategies. Figure 3 displays the average and standard deviation of the end-of-course grades for the graduate learning strategies identified by the *k*-means algorithm. Students who adopted learning Strategy 5 (*active, prolonged, and frequent engagement*) achieved the highest average grades (16.33 ± 2.74), followed by those adopting learning Strategy 4 (*mildly frequent, mild activity in short sessions*) with an average grade of 16.01 (± 2.77) and Strategy 2 (*mildly frequent and task-focused*) with an average grade of 15.46 (± 2.75). Strategies 1 and 3 were associated with the lowest average grades among all learning strategies used by graduate students, with average grades of 14.92 (± 3.85) and 14.90 (± 3.15), respectively.

This section has provided a detailed analysis of the relationship between various learning strategies and student performance. Significant differences in performance among the learning strategies were observed at both undergraduate and graduate levels. The data suggests that the choice of learning strategy can significantly impact academic performance.

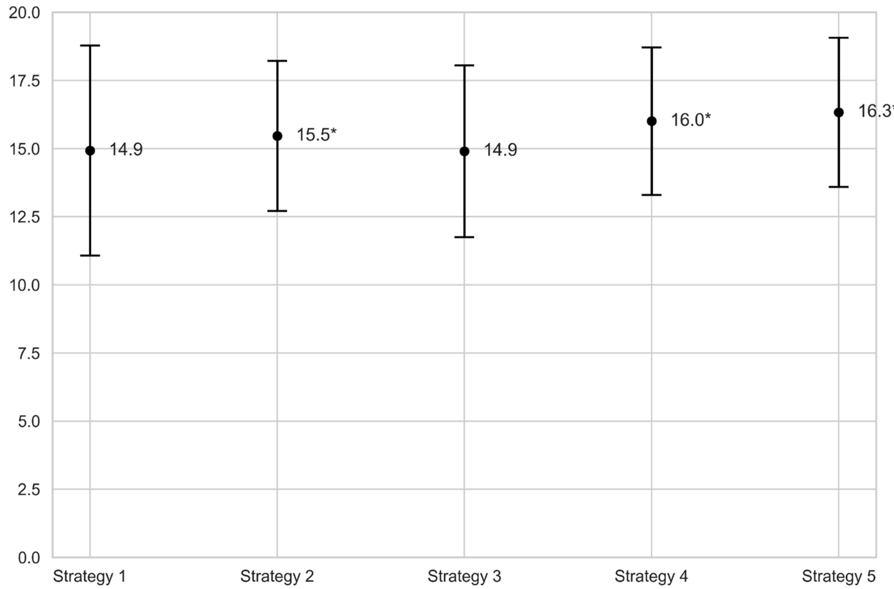


Figure 3
Average and standard deviation of the end-of-course grade for graduate learning strategies (values that exhibit statistically significant differences against all other groups are identified with an asterisk)

4 Discussion

4.1 Research question 1: What course-agnostic learner strategy profiles can be extracted from undergraduate and graduate students' SRL features extracted from LMS data?

The first research question in this study aimed to uncover course-agnostic learning strategy profiles from undergraduate and graduate students based on SRL features extracted from LMS data. The analysis identified five distinct profiles at each instruction level with varying levels of engagement, activity, and procrastination tendencies. The strategies identified were relatively similar for both the graduate and undergraduate levels.

The first learning strategy, *active, prolonged and frequent*, refers to students who were generally the most engaged across all perspectives. This learning strategy suggests that these students consistently devote time and effort to accessing the LMS and the materials contained therein, thus suggesting well developed SRL resource management skills (Pintrich et al., 1991). More specifically, regular and prolonged accesses hint at the students' awareness of the materials available and their ability to schedule the necessary time to study (*Time and study environment*). Moreover, the frequent accessing also suggests discipline to continue studying over the entire semester, suggesting elevated *effort regulation*.

The second strategy, *mildly frequent, mild activity in short sessions*, is associated with students who logged into the LMS somewhat regularly but had short sessions with average levels of activity. The regular accesses also point to a certain degree of development in skills associated with *effort regulation* and *time and study environment*. While additional data would be needed to confirm this, the behavior exhibited by these students suggests that their main focus would be having the discipline to access specific materials deemed relevant, and logging out of the LMS afterwards, suggesting the existence of a more strategic approach, which was something observed in Cerezo et al.'s (2016) *Task-oriented and socially focused group*.

Students adopting the third learning strategy, *mildly frequent and task-focused*, showed average values for most activity metrics but specifically concentrated their efforts on completing assessments. This group shares certain similarities in learning strategy with the second group, with the main difference being the types of resources accessed by the students, which suggests some degree of development in skills associated with *effort regulation* and *time and study environment*. However, due to the partial nature of LMS data, it is impossible to draw meaningful distinctions between these two groups regarding SRL traits.

The fourth learning strategy, *likely procrastinators*, consisted of students who started interacting with course materials later, indicating procrastination. However, once logged in, they had long and intensive sessions, which aligns with conventional procrastination behavior, indicative of poor resource management skills, and has been shown to be a marker for poorer academic performance (Cerezo et al., 2016; Riestra-González et al., 2021; Yang et al., 2020).

The fifth and final learning strategy, termed *inactive*, is associated with students who exhibited the lowest LMS activity and engagement levels across all metrics. These students may be facing challenges that prevent them from engaging with the course materials or rely on resources outside of the LMS for their learning. Future research could focus on identifying the reasons behind such low engagement levels, in order to provide appropriate support and resources to better understand and address their needs.

Considering the results and the information presented in Table 1, it is possible to see differences in how students at the undergraduate and graduate levels behave on Moodle in absolute terms. However, in relative terms, the learning strategies they followed share similarities that do not warrant a meaningful distinction in their description. Thus, Research Question 1 can be answered by stating that *k*-means uncovered five distinct patterns of interaction for learning strategies that were similar for both instruction levels: *active, prolonged and frequent engagement; mildly frequent and task-focused engagement*; *mildly frequent, mild activity in short sessions engagement*; *procrastinators;* and *inactive*.

4.2 Research question 2: What is the relationship between the learning strategies uncovered by k-means and end-of-course performance at each instruction level?

Baker et al. (2020) noted that clickstream data from LMS logs provide only a noisy and partial view of student behavior and learning. However, when the average end-of-course performance of students was mapped to their Moodle learning strategies, similar patterns were found for both undergraduate and graduate instruction levels.

Students who adopted the *inactive* learning strategy achieved the lowest grades, with an average of 12.41 for undergraduates and 14.90 for graduates. They were followed by those who adopted the *likely procrastinators* approach, with an average of 13.17 for undergraduates and 14.92 for graduates. These grades, in conjunction with the observed behavior on the LMS, suggest that some students in these groups either lacked a learning strategy with Moodle or had an inefficient approach to learning, both indicative of poor resource management skills development. These findings are consistent with other studies that have found lower levels of engagement to be associated with lower academic achievement (Cerezo et al., 2016; Hung & Zhang, 2008; Riestra-González et al., 2021; Yang et al., 2020). However, it is important to interpret these results with caution, as some students who did not engage with Moodle still obtained remarkable grades, possibly due to having a learning strategy that did not include active engagement with the LMS.

On the other hand, students who followed the *active, prolonged and frequent engagement* strategy achieved the highest overall grades. They were followed by those who adopted the *mildly frequent, mild activity in short sessions engagement* strategy, and those who followed the *mildly frequent and task-focused engagement* strategy. The evidence suggests that starting early and logging in frequently is an important factor in achieving better outcomes than the other strategies discussed previously. Although additional data would be needed for a more comprehensive assessment of these students, the behavior exhibited at least hints at the existence of a baseline learning strategy in place for the students' interactions with Moodle. An additional factor that may differentiate between grades are the types of actions performed on Moodle and the time spent on it. While it is true that the most successful students were also the most active, there is evidence that the types of interaction, rather than total activity, also play a relevant role in determining academic success. The results show that the two most successful strategies focused more on consulting theoretical content such as resources or external URLs. This is particularly interesting because other studies (Cerezo et al., 2016; Riestra-González et al., 2021) found that students with a theoretical focus were surpassed by those who were equally engaged but followed a task-oriented approach, which was not the case for the present data. It is also important to note that not all time spent studying is equal, as noted by Cerezo et al. (2016). The second-most successful students clicked less and spent considerably less time on Moodle than their peers following the first and third-most successful approaches. This suggests that these students may have adopted a more strategic approach to their learning, resulting in a more efficient and higher quality use of their study time.

The findings from this study provide an answer to Research Question 2: A generally positive relationship was observed between the levels of engagement in learning strategies, as uncovered by *k*-means, and end-of-course performance across both instruction levels. Students who adopted *inactive* or *likely procrastinator* approaches to learning tended to have the lowest grades, while those who engaged in *active, prolonged, and frequent* interactions with Moodle achieved the highest overall grades. Early and frequent access to Moodle emerged as a key factor in achieving better outcomes. However, while this relationship was clear at the extreme ends of the spectrum, it became less distinct in the middle. Here, other factors such as the types of actions performed on Moodle and the time spent on it began to influence academic success in ways that were not always immediately apparent. Moreover, it is crucial to remember that Moodle logs represent only a portion of the learning process. This approach does not measure other potentially impactful factors, such as intrinsic motivation. Therefore, while Moodle logs provide valuable insights, they should be viewed in a broader context when evaluating student learning strategies and academic performance.

4.3 Research question 3: Are there differences in the effectiveness of the learning strategies between instruction levels?

When examining the clustering analysis results, there appear to be only minor differences between the learning strategies adopted by undergraduate and graduate students, as the same five general strategies emerged at both instruction levels. The primary difference was that, despite starting to access Moodle much later, undergraduate students exhibited higher overall levels of engagement in comparison to their graduate counterparts. From Table 1, we know that, on average, undergraduate students had higher amounts of clicks, sessions, and time spent on Moodle. These findings are also supported by the differences in prevalence of the different strategies at both levels. Approximately 25.01% of the undergraduate students adopted the *mildly frequent and task-focused engagement* strategy (against 21.71% in graduate students), while the most common learning strategy among graduate students is the *mildly frequent, mild activity in short sessions* (26.78% compared to 20.82% of undergraduates). Graduate students also have a lower prevalence of *active, prolonged and frequent engagement* than their undergraduate counterparts (14.62% to 19.89%). These results align with expectations, as graduate students are generally older and are expected to have more developed resource management SRL skills, thus being more likely to efficiently manage their time and resources, and not needing to spend as much time logged in to fulfil their study objectives.

However, when examining the relative effectiveness of strategies at each instruction level, the patterns were remarkably similar. Across both groups, the ranking of learning strategies relative to their end-of-course grades followed the same order, with the strategies involving the most frequent accesses leading to the highest grades and procrastination and inactivity being associated with the lowest student performance. The consistency of these findings suggests that the core relationships between LMS engagement patterns and course outcomes are potentially generalizable across undergraduate and graduate contexts. While undergraduate students may utilize online platforms more extensively overall, the basic connections between behavior and performance appear to hold steady at both instruction levels.

Therefore, the answer to Research Question 3 is that no major differences were observed in the relative effectiveness of learning strategies between instruction levels. The key factors leading to positive outcomes remained important for both undergraduates and graduate students.

4.4 Implications

The findings presented herein provide relevant implications for both research and practice. On the research front, this work contributes to a growing body of literature aimed at uncovering learning strategies from trace data through unsupervised machine learning techniques. The results showcase both the potential and limitations of using LMS logs to categorize students based on their engagement patterns. In particular, the consistency of the relationships between strategy and performance across undergraduate and graduate contexts points to opportunities for developing more generalized models. Exploring the reasons behind students' choice of strategies is another area for future work, as the motivations and challenges faced by different learners, especially the less active ones, are still unclear. Qualitative or survey data collected alongside the logs may reveal additional insights into which motivational and emotional factors contribute to the understanding of some of the performance differences between strategies.

In practice, categorizing students into strategy profiles could inform the design of personalized interventions to improve resource management skills. Students following less successful approaches could receive prompts or tutorials for developing better time management habits or content pacing. These adaptive supports would not be a one-size-fits-all solution; they would target the specific gaps exhibited through the engagement patterns. Moreover, course designers could use this knowledge to design programs and courses to promote forms of engagement that are more conducive to developing SRL skills and, more importantly, student success. Additionally, the presented methodology for extracting and analyzing variables from LMS data could be packaged into a reusable toolkit for institutions with accessible analytics dashboards that automatically cluster students based on trace behaviors, providing educators with actionable insights to refine their instructional practices and better support learners.

4.5 Limitations

This study has several limitations that must be acknowledged. The data source consists exclusively of LMS logs from a single institution over one academic year. While the sample size is large, incorporating multiple schools over longer periods could improve generalizability. Reliance on a unique data source also provides an incomplete picture of the learning process, as offline behaviors and other contextual variables are unavailable. Future research on this topic could complement data from the LMS with other instruments to develop a more comprehensive understanding of the learning strategy profiles.

Another relevant limitation concerns the SRL theoretical grounding of this approach. While theoretical connections are drawn between strategies, features, and SRL skills, all of them are indirect measurements of engagement with a single platform, and no direct observations of SRL constructs were performed. These connections, while suggested by empirical relationships, are not definitively confirmed. Future studies could incorporate established SRL instruments, such as the MSLQ, or use open-ended surveys or interviews. This could reveal individual motivations, challenges, and decision-making processes, providing a richer explanation for observed engagement patterns and performance differences. Such an approach could strengthen the theoretical basis of the analysis and offer nuanced insights into how students' SRL processes manifest in their online behaviors. Moreover, it could guide the development of interventions that target specific phases of the SRL process, thereby offering more targeted and effective support for students.

Conclusion

This work presented an analysis of uncovering learning strategies from Moodle log data through an unsupervised machine learning approach to assess learning strategy effectiveness across undergraduate and graduate contexts.

Clustering algorithms were leveraged to categorize over 11,000 student enrollments into distinct profiles based on their LMS engagement patterns. The findings revealed five similar strategies at both instruction levels: *active, prolonged and frequent engagement; mildly frequent and task-focused engagement*; *mildly frequent, mild activity in short sessions engagement*; *likely procrastinators*; and *inactive*.

Clear relationships emerged between engagement behaviors and student outcomes by mapping academic performance to these strategies. Across contexts, prolonged activity and early access were reliable markers of success, while procrastination and disengagement corresponded to lower achievement. However, success factors were more complex for some groups, involving strategic use of time and choice of activities. Still, the core patterns translating engagement to performance were strikingly consistent between undergraduates and graduates.

Nonetheless, this research makes valuable contributions. It demonstrates the feasibility of extracting meaningful learning strategy profiles from LMS data at scale across courses and instruction levels. The findings illustrate connections between online behaviors and performance. The findings also inform design principles for personalized interventions that target the development of successful learning strategies.

However, some limitations should be acknowledged. The study relied solely on clickstream data, providing an incomplete view of learning processes. Additional data on student demographics, prior achievement, and psychological factors like motivation could enrich the analysis. Adding this data would allow for a more comprehensive incorporation of the results presented herein into one of the existing SRL models (Panadero, 2017), which would not only provide a clearer interpretation of the results but would also contribute to an increased understanding of the motivational and emotional processes that lead students to adopt specific learning strategies. Moreover, the specific courses, instructors, and institutional contexts likely influenced the results. The sample was collected from an information management school, and replicating this approach across more diverse settings would strengthen conclusions about the potential generalizability of a course-agnostic approach.

There are several promising avenues for future work building on this research. One direction involves applying similar techniques to datasets across multiple institutions over longer timeframes. This could evaluate the consistency of findings and further establish generalizability of the relationships between online behaviors, strategy profiles, and achievement. Additionally, incorporating supplementary data sources beyond Moodle logs, whether institutional datasets or direct SRL measurements, holds potential for constructing more comprehensive learner models. Methodologically, exploring alternatives beyond *k*-means clustering, and developing personalized feedback mechanisms tailored to strategy profiles may unlock new possibilities. These next steps emphasize the importance of understanding the factors influencing learning strategies and academic performance and, hopefully, translate analytics into positive pedagogical impact through interventions that develop effective self-regulated learning strategies among students.

In conclusion, this work contributes both methodologically and empirically to the growing body of literature on mining learner strategies from trace data. The findings provide a foundation for personalized interventions while highlighting opportunities for future research. Supplementing logs with additional data sources and perspectives would lead to more robust, generalizable, and actionable models.

References

Aljohani, N. R., Fayoumi, A., & Hassan, S.-U. (2019). Predicting at-risk students using clickstream data in the virtual learning environment. *Sustainability*, *11*(24), Article 7238. https://doi.org/10.3390/su11247238

Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, C., Rodriguez, F., Warschauer, M., & Smyth, P. (2020). The benefits and caveats of using clickstream data to understand student self-regulatory behaviors: Opening the black box of learning processes. *International Journal of Educational Technology in Higher Education*, *17*(1), Article 13. https://doi.org/10.1186/s41239-020-00187-1

Bellur, S., Nowak, K. L., & Hull, K. S. (2015). Make it our time: In class multitaskers have lower academic performance. *Computers in Human Behavior*, *53*, 63–70. https://doi.org/10.1016/j.chb.2015.06.027

Bernacki, M. L., Chavez, M. M., & Uesbeck, P. M. (2020). Predicting achievement and providing support before STEM majors begin to fail. *Computers & Education*, *158*, Article 103999. https://doi.org/10.1016/j.compedu.2020.103999

Biggs, J. B. (1987). *Student approaches to learning and studying. Study process questionnaire manual.* Australian Council for Educational Research.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, *7*(2), 161–186. https://doi.org/10.1016/S0959-4752(96)00015-1

Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, *33*, 24–32. https://doi.org/10.1016/j.iheduc.2017.01.004

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A Systematic Review. *The Internet and Higher Education*, *27*, 1–13. http://dx.doi.org/10.1016/j.iheduc.2015.04.007

Çebi, A., & Güyer, T. (2020). Students' interaction patterns in different online learning activities and their relationship with motivation, self-regulated learning strategy and learning performance. *Education and Information Technologies*, *25*(5), 3975–3993. https://doi.org/10.1007/s10639-020-10151-1

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education*, *96*, 42–54. https://doi.org/10.1016/j.compedu.2016.02.006

Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary Education and Management*, *11*(1), 19–36. https://doi.org/10.1007/s11233-004-3567-9

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics*, *4*(2), 113–128. http://dx.doi.org/10.18608/jla.2017.42.10

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Hung, J.-L., & Zhang, K. (2008). Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching. *Journal of Online Learning and Teaching*, *4*(4), 426–437.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255–259.

Li, L.-Y., & Tsai, C.-C. (2017). Accessing online learning material: Quantitative behavior patterns and their effects on motivation and learning performance. *Computers & Education*, *114*, 286–297. https://doi.org/10.1016/j.compedu.2017.07.007

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam, & J. Neyman (Eds.), *Berkeley Symposium on Mathematical Statistics and Probability* (s. 281–297). University of California Press. https://www.cs.cmu.edu/~bhiksha/courses/mlsp.fall2010/class14/macqueen.pdf

Matcha, W., Gašević, D., Jovanović, J., Uzir, N. A., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of learning strategies: The association with the personality traits. *LAK '20: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 151–160. https://doi.org/10.1145/3375462.3375534

McKinney, W. (2017). *Python for data analysis: Data wrangling with pandas, NumPy, and IPython* (2nd Ed.). O'Reilly Media.

Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2020). Student engagement level in an e-learning environment: Clustering using k-means. *American Journal of Distance Education*, *34*(2), 137–156. https://doi.org/10.1080/08923647.2020.1696140

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ).* The Regents of The University of Michigan.

Puzziferro, M. (2008). Online technologies self-efficacy and self-regulated learning as predictors of final grade and satisfaction in college-level online courses. *American Journal of Distance Education*, *22*(2), 72–89. https://doi.org/10.1080/08923640802039024

Riestra-González, M., Paule-Ruíz, M. del P., & Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education*, *163*, Article 104108. https://doi.org/10.1016/j.compedu.2020.104108

Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2013). Web usage mining for predicting final marks of students that use Moodle courses. *Computer Applications in Engineering Education*, *21*(1), 135–146. https://doi.org/10.1002/cae.20456

Santos, R. M., & Henriques, R. (2023). Accurate, timely, and portable: Course-agnostic early prediction of student performance from LMS logs. *Computers and Education: Artificial Intelligence*, *5*, Article 100175. https://doi.org/10.1016/j.caeai.2023.100175

Susac, A., Bubic, A., Kaponja, J., Planinic, M., & Palmovic, M. (2014). Eye movements reveal students' strategies in simple equation solving. *International Journal of Science and Mathematics Education*, *12*(3), 555–577. https://doi.org/10.1007/s10763-014-9514-4

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, *27*(4), 551–572.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). Academic Press. https://doi.org/10.1016/B978-012109890-2/50045-7

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. https://doi.org/10.1007/s10115-007-0114-2

Yang, Y., Hooshyar, D., Pedaste, M., Wang, M., Huang, Y.-M., & Lim, H. (2020). Predicting course achievement of university students based on their procrastination behaviour on Moodle. *Soft Computing*, *24*(24), 18777–18793. https://doi.org/10.1007/s00500-020-05110-4

Yeo, I.-K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, *87*(4), 954–959. https://doi.org/10.1093/biomet/87.4.954

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An Overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge.

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-4

"I SHOULD, BUT I DON'T FEEL LIKE IT": OVERCOMING OBSTACLES IN UPPER SECONDARY STUDENTS' SELF-REGULATION USING LEARNING ANALYTICS

Mattias Wickberg Hugerth[a], Jalal Nouri[a], Anna Åkerfeldt[b]

[a] Department of Computer and Systems Sciences, Stockholm University, Sweden
[b] Department of Teaching and Learning, Stockholm University, Sweden

ABSTRACT

While research has been conducted on self-regulated learning in relation to learning analytics, there remains a knowledge gap regarding the obstacles secondary education students face in regulating their learning and how learning analytics can support their self-regulation. This paper investigates two questions: 1) What challenges do secondary education students experience in the process of regulating their own learning?, and 2) What information and data do secondary education students need to better regulate their own learning? We conducted a study at a mid-sized upper secondary school in middle Sweden, to better understand how these issues manifest among students. We analyzed data collected by the school twice annually between 2015 and 2022, and administered a questionnaire to 224 students to answer the research questions. Through descriptive statistics and a thematic analysis, we identify prevalent problems that students encounter, as well as the necessary information that is essential for scaffolding self-regulated learning. We discuss the implications of our findings for the design of systems that provide students with relevant data to enhance their learning experiences.

KEYWORDS

self-regulated learning; obstacles; learning analytics; scaffolding; secondary education

CORRESPONDING AUTHOR
Mattias Wickberg Hugerth, Department of Computer and Systems Sciences, Stockholm University, Borgarfjordsgatan 12, 164 55 Kista, Sweden
e-mail: mattias.wickberg@dsv.su.se

Introduction

This story begins, as many do, as a tale of frustration. As teachers, we have been frustrated, at times exasperated, by the never-decreasing number of students who can't seem to get the job done on time. Missed deadlines, cramming just before tests, writing a whole essay the night of the deadline, extensions, procrastination, and bad planning – these are staples of our lives. The experiences led us to wonder – can Self-Regulated Learning (SRL) and Learning Analytics (LA) be part of the solution to this problem?

The importance of planning related to self-regulation is well known (Gollwitzer, 1999), and the problem with students being unable to plan, or focus on their work, is not a story unique to our experience, but something that permeates education everywhere. As one researcher notes: "The problems of distracted learning, as well as the associated solutions, are far deeper than meets the eye" (Schmidt, 2020, p. 286). It has been noted that while education often creates learning situations where SRL skills are needed (Bolhuis & Voeten, 2001; Dignath & Veenman, 2021), and that there are ways to teach SRL that affects learning and motivation positively (Dignath et al., 2008), schools rarely teach SRL skills in an effective manner (Dignath & Veenman, 2021). If schools are failing to teach students SRL skills, while still requiring it of them, there is a need to both further understand what problems students encounter, and to investigate other means of supporting SRL.

One such possible means of support is through the use of Learning Analytics, which offers ways to analyze and present data insights for students that may help them self-regulate (Lodge et al., 2018; Winne, 2022). While the research around SRL and LA is vast, most of it has been done in higher education (Heikkinen et al., 2023; Schwendimann et al., 2016), so there is a still a need for further investigation of these areas in primary and secondary education.

This study aims to understand where students in upper secondary school encounter problems in regulating their learning, which areas may be suitable for scaffolding using learning analytics, and what data is needed for such scaffolding. Thus we ask the following research questions:

- Research Question 1: What challenges do secondary education students experience in the process of regulating their own learning?
- Research Question 2: What information and data do secondary education students need to better regulate their own learning?

1 Self-regulated learning and learning analytics

Self-Regulated Learning (SRL) has over the last decades become important to the field of educational research (Schunk & Greene, 2017). According to Zimmerman, "self-regulated learners are persons who plan, organise, self-instruct, self-monitor, and self-evaluate at various stages during the learning process" (Zimmerman, 1986, p. 308).

There are several different models of SRL (Panadero, 2017; Puustinen & Pulkkinen, 2001), six of which were analyzed by Panadero (2017), who concludes that five of the six models can be said to include three phases of SRL, although their names and structure may differ. The three common phases are: 1. Preparatory phase, 2. Performance phase, and 3. Appraisal phase. For the purpose of this paper, Zimmerman's cyclical phase model of SRL (Zimmerman, 2000) will be used as a primary model for reference and analysis, and guides both the data collection and subsequent analysis. In this model the three phases mentioned above are called 1. Forethought, 2. Performance and 3. Self-reflection, where each phase has two subcategories, as seen in Figure 1 (Zimmerman & Moylan, 2009).

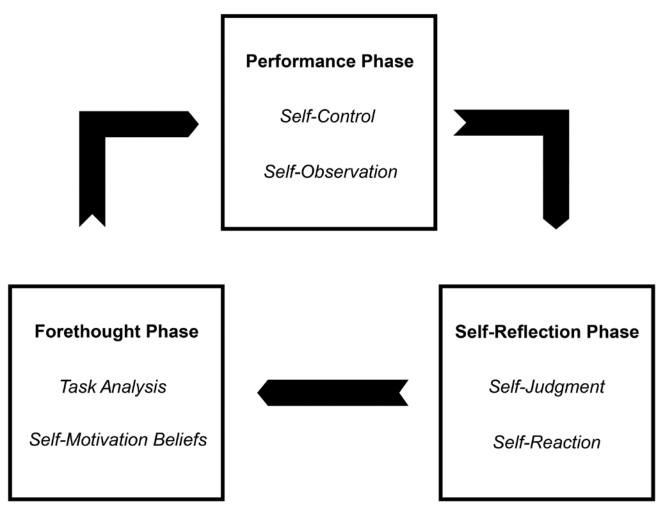


Figure 1
Cyclical phase model based on Zimmerman & Moylan (2009)

Zimmerman (1986) describes how one can become a self-regulated learner, implying that not everyone is, while Winne instead argues that SRL is ubiquitous (Winne, 1995) but not all learners regulate in ways that are suited for the task, or they regulate at non-optimal times (Winne, 2005). Winne (2005) also suggests that scaffolding can help mediate this problem, and that learners need information (Winne, 2005, 2022), particularly process feedback, but they also need tools to become better self-regulated learners.

In attempting to scaffold students' SRL, it is imperative to know at what stage in the process the student needs scaffolding, which means looking at what problems the students encounter while regulating their learning. There are several examples of problems in the literature. Boekaerts (1999) points out that there are differences in students' abilities to handle goals, with some students seemingly unable to handle multiple goals and instead focusing on one goal at a time, meaning some goals may be postponed, and sometimes never even make it into the student's focus.

One way of scaffolding learners' SRL is to provide information. As students are developing their regulatory skills, they are often hindered by a lack of information and feedback that could guide their efforts (Winne, 2022), which is one central argument for the importance of using Learning Analytics for SRL development. The function of LA in this context can be, for instance, to provide students with feedback on their efforts, and suggestions for action based on analysis of previous students' actions and performance (Afzaal et al., 2021a, 2023).

The field of LA has grown alongside digitalization in education, and is most commonly defined as "[...] the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing learning and the environments in which it occurs." (Siemens, 2013, p. 1382), a definition that has been standing since the first International Conference of Learning Analytics in 2011.

In LA, a fair amount of work has been done on the intersection of learning analytics and self-regulated learning (Álvarez et al., 2022; Heikkinen et al., 2023; Matcha et al., 2020). In a systematic review Matcha et al. (2020) look at existing Learning Analytics Dashboards (LADs), and conclude that there are problems relating to user-centered design, that there is a lack of knowledge of, among other things, study strategies and tactics used by students, and that information provided to students as feedback should be presented in multiple forms. Álvarez et al. (2022) made another systematic review of LADs that points out that the links between functionality in dashboards and the processes that they are supposed to support are lacking. Heikkinen et al. (2023) point out that while there has been a recent trend toward a broader view, most studies of learning analytics interventions in support of SRL have been focused on a single course.

It has also been pointed out that while scaffolding may be important for many students, it can, in fact, even be detrimental to students who already have a high level of intrinsic motivation (Duffy & Azevedo, 2015), and that scaffolding systems have to give the user a certain amount of control over the amount of support they receive. To make sure that the LA-based interventions do not take over regulation completely and thus act detrimentally on the students' skill progression, hybrid methods have been proposed that share responsibility for regulation between system and human (Molenaar, 2022), an approach that is still in the preliminary stages.

However, most of the research done has been in higher education rather than primary or secondary. The systematic review by Heikkinen (2023) looked at studies of learning analytics interventions to support SRL, and out of 56 studies, only one looked at primary or secondary education. Another systematic review shows that most learning analytics dashboards aimed at supporting SRL seem to be focused on the reflection phase of SRL and provide little or no support for the other two (Jivet et al., 2017).

2 Methodology

This study is the first part of a larger study, where Design-Based Research (DBR) is used for designing support for self-regulated learning in upper secondary education. A very short description of DBR, rephrasing that from a 2012 paper (Anderson & Shattuck, 2012, pp. 16–17) describes it as follows: A study situated in a real educational context, often using mixed methods, that involves iteratively designing and testing an intervention in close partnership between researchers and practitioners. This study identifies what needs pertaining to SRL are suitable for intervention in the form of a digital interface using learning analytics to scaffold students' SRL development, and will act as a basis for the next study, which will concern the iterative process of prototyping that interface.

2.1 Data Collection

The school where this study was conducted is a mid-sized (378 enrolled students at the start of data collection) upper secondary school, called gymnasium in Sweden (three years, starting the year a student turns 16) in the middle of Sweden. The school offers a technological program with a profile of information technology. Most of the students are male, with about 10% female enrollment at the time of the study. The first author of this paper was a part-time teacher at the school in question, which may have increased the response rate. All data was collected anonymously.

Data collection for this study was done in two parts. The first part is data that the school itself collected over 7.5 years from 2015 to 2022, and which consisted of students identifying which skills they most needed to improve. The second part was a survey created and sent by the authors of this paper to students at the school. The first set of data was collected and analyzed to identify what challenges students find in their studies, and whether those challenges were stable over time. The fact that the existing data had been collected at regular intervals, in the same context and with exactly the same wording over time, provided an opportunity to investigate trends over time. The results of the analysis of the first data then provided the basis for deciding which areas should be focused on in the second data collection, aimed at better understanding the nuances of the problems identified in the first.

The second part of the data collection consisted of a survey, based on the results of the analysis of the data from the school in combination with Zimmerman's model for SRL (2000). The survey consisted of 29 questions, focusing on the most important skills as identified by the students in the first dataset, that are also important aspects of SRL. Four of these questions were about basic information, asking which year they were in, which of the two programs they took, a self-reported average grade, and a question where the students were asked which of three descriptions fit them best. This last question aimed to group the students by how much they perceive themselves to struggle in school, to see if there are patterns in students' answers relating to this perception of themselves. This section was followed by 21 questions divided into the sections *focus*, *planning*, *information*, *engagement*, and *motivation*, where each section had both multiple-choice questions on a scale of 1 (*Never/Almost never)* to 5 (*Always/Almost always*) and open questions.

Table 1

1	To what extent can you focus on your schoolwork during classes?
1a	When you can't focus during classes, what are the reasons for this?
2	To what extent can you focus on your schoolwork outside of class time?
2a	When you can't focus outside of class time, what are the reasons for this?
3	To what extent do you plan your studies?
3a	How do you plan? What information do you use?
3b	If you do not plan, why not?
3c	What support and information do you need to plan your studies better?
4	To what extent do you follow your plans?
4a	When you fail to follow your plan, what do you think is the cause/causes?
4b	Is there any information that you think would help you follow your plans better?
5	To what extent do you believe you have the information you need to develop in the various courses you are taking?
5a	What information do you have today that is useful to you?
5b	What information do you currently lack?

6	To what extent can you feel engaged in schoolwork?
6a	In which situations is it easy for you to feel engaged?
6b	In which situations is it difficult?
7	Do you feel motivated to learn?
7a	When do you feel the most motivated?
7b	When do you lose motivation?
7c	What information could increase your motivation?

Table 1 shows the central questions in the questionnaire, translated to English. Questions with only a number are closed questions. Questions with a letter are open. Questions that were not analyzed for this study have been left out of this table.

2.2 Participants

The first set of data consisted of 9655 answers from a total of 973 students, over 15 semesters (7.5 years).

The questionnaire was sent in early October 2022, with two reminders one and two weeks later. The students could fill in the questionnaire at any time they preferred during the three weeks of data collection, but were also given time during the weekly scheduled class council. Out of the 378 students that received the questionnaire, 224 answered, for a response rate of 59%. The respondents were relatively evenly spread among the years, with 41.1% in year 10, 29.0% in year 11 and 29.1% in year 12. Question number 4 asked the students which of three descriptions best describes them: Persona A "the struggling student", Persona B "the student who does ok but could work more," and Persona C "the student who does well."

2.3 Data analysis

The first data set and the second set's multiple-choice questions were analyzed using descriptive statistics. The open-ended questions underwent thematic analysis using Braun and Clarke's method (Braun & Clarke, 2006), involving six phases: familiarization with data, code generation, theme identification, theme review, theme definition, and report production. The process involved reading the data set, noting initial impressions, coding each answer to capture its essence, and discussing findings among researchers. Codes were revised, grouped into themes with examples, and continuously checked for relevance to research questions. Themes were then mapped, named, and described, ensuring alignment with the underlying data. A heat map was made to visualize theme frequencies (Figure 3) Finally, the findings were discussed in relation to Zimmerman's Self-Regulated Learning (SRL) model, focusing on students' challenges and support needs within the model.

3 Results

3.1 Analysis of first survey

From the analysis of the first survey, two factors stand out that students see as obstacles for their studies, and these two seem stable over time: Planning their studies, and focusing on their school work (figure 2). This result can be seen in both datasets, which will be shown in greater detail below. The reasons for these obstacles vary (figure 3), but again some stand out. The obstacles to focusing are primarily external factors like noise or other people disturbing them, along with electronic devices, and internal factors like tiredness and lack of motivation. More details about these factors will be presented in the sections below. The latter two categories are the most commonly occurring answers overall in the data. Another commonly reported obstacle is lack of knowledge about the planning process itself.

As for the second research question we posed, the major factors seem to be lack of information, lack of clarity in the information they do have available, and lack of information regarding how they should progress in their studies.

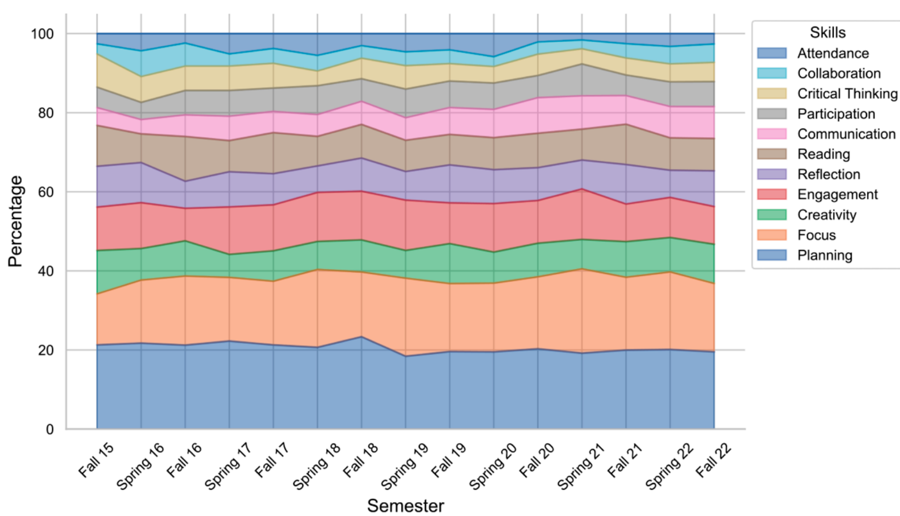


Figure 2
Chosen skills as percentages of all answers, by semester

Skills chosen illustrated with an area graph. Note that each student chose three skills. Approximately 20 % of choices were planning, which comes out to roughly 60% of students having chosen planning as one of their three each term. Y-axis: Share of the total number of answers given, adding up to 1 (100%).

Theme	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6
1. Internal distractions	1.1 Tiredness 161	1.2 Lack of motivation 121	1.3 Difficulties focusing 81	1.4 Hunger 5	1.5 Psychological reasons 76	
2. External distractions	2.1 Other people 70	2.2 Sound or volume 47	2.3 Electronics 92	2.4 Other priorities 81	2.5 Other distractions 21	
3. Information available	3.1 Learning platform 36	3.2 Temporal information 98	3.3 Other people 19	3.4 Formal information 33	3.5 Personal knowledge related information 13	3.6 Planning related information 5
4. Information needs	4.1 Information needs 76	4.2 Need for clarity in information 83	4.3 Personal knowledge-related information 22	4.4 Personal development-related information 59	4.5 Planning-related information 10	4.6 Information organisation 17
5. Support for planning	5.1 Tools 39	5.2 Needs for planning 70	5.3 Needs for execution 21			
6. Motivational needs	6.1 Utilization of knowledge 34	6.2 Other motivational factors 6				

Figure 3

Themes and categories identified in the survey data, with number of occurrences of each category in the lower right corner of each cell. Cells are formatted on a red-yellow-blue scale. with red for highest number of occurrences and blue for lowest.

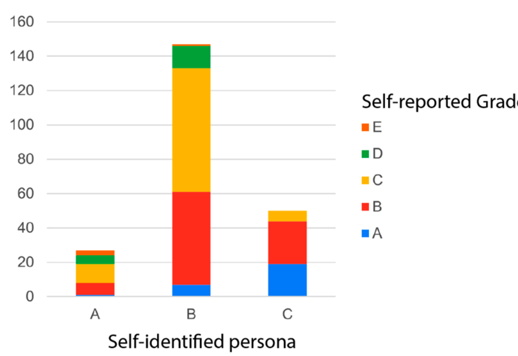


Figure 4

Students by self-identified persona, stacked by self-reported approximate grade where A is the highest grade, and E is the lowest passing grade. It can be noted that there are few low-performing students in this data set.

3.2 RQ1: What challenges do students experience in the process of regulating their own learning?

Looking at this in more detail, we can see that both the first set of data, which the school itself has been collecting over the years, and parts of the data collected specifically for this study provide answers for the first research question.

The data from the school ranges from the second half of 2015 until the second half of 2022 for 15 terms of data in total; there is a very stable trend (figure 2). At every point in the time covered, the skills planning and focus are the two most commonly chosen skills that students declare a need to develop. The fractions for the different skills are mostly the same over the three years of the students attending the school (figure 5), with little change in chosen skills at group level from starting the school to graduating three years later. There is a sharp decline in how many students reply to the questions in the system in the sixth and last semester, from 1464 data points for the fifth semester to 564 in the sixth and last. Among the eleven skills that are presented to the student, three clearly map to the three phases of SRL: Planning as part of the forethought phase, focusing as part of the performance phase, and reflection as part of the self-reflection phase, which means that for the students attending this school, there are clear problems relating to the forethought and performance phases. Based on this analysis, the questionnaire developed for the second part of data collection was focused primarily on planning, focus and engagement as the main components that stood out.

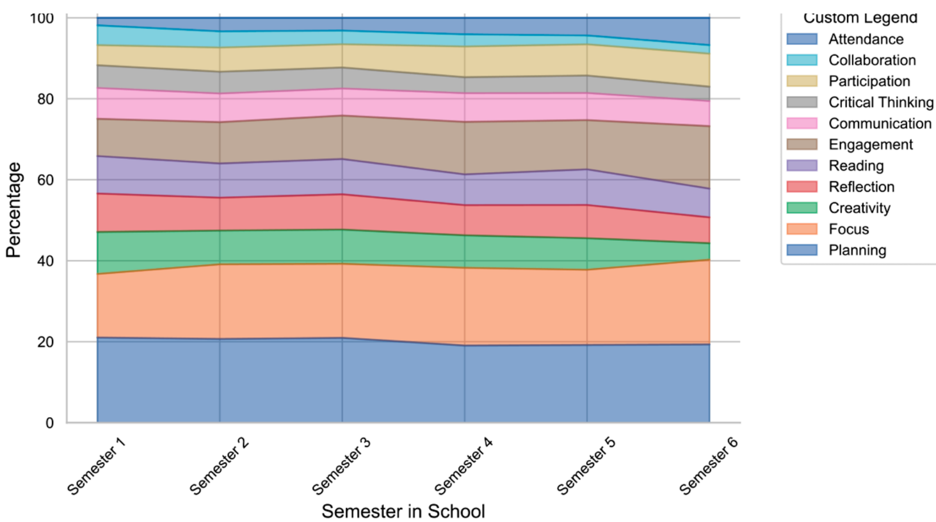


Figure 5

Chosen skills as percentages of all answers, by semester in school

Like Figure 1, but by term 1–6 of the pupils' three-year education, i.e. 1 is for the answers from the first semester of year 10, 6 is for the second semester of year 12. Y-axis: Share of the total number of answers given, adding up to 1 (100 %).

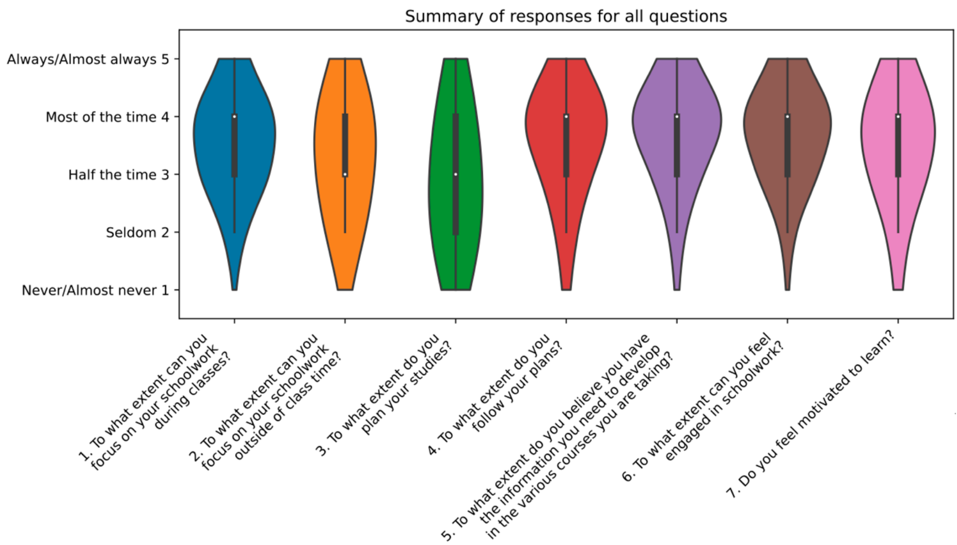


Figure 6a

Answers to multiple-choice questions. Question number 4 was optional. X-axis: Questions 1–7, see Table 1. Y-axis: Answers on Likert scale from 1 (Never/Almost never) to 5 (Always/Almost always).

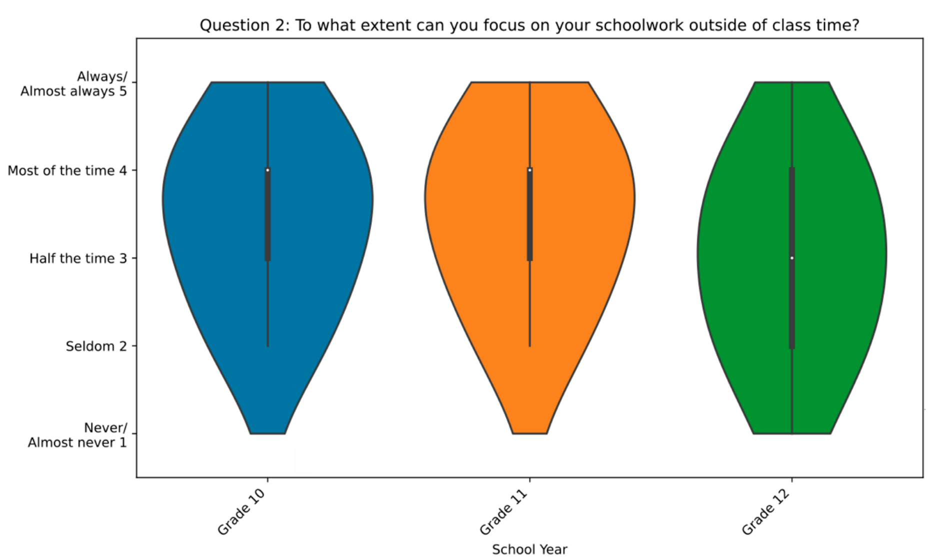


Figure 6b

Question 2 according to school year 1–3 on X-axis.

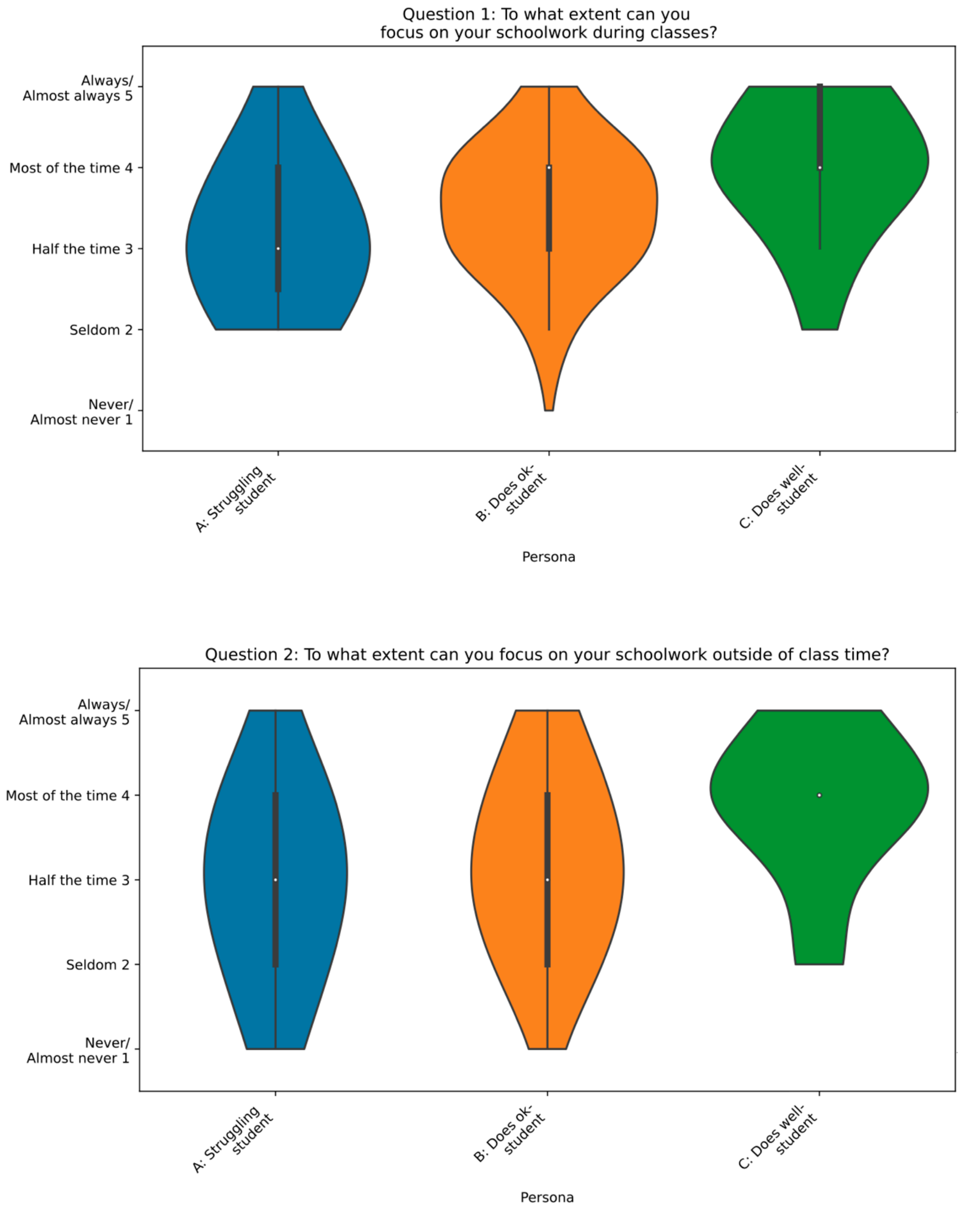


Figure 7
Answers to question 1 (top) and question 2 (below) according to persona (X-axis). Y-axis is the same Likert scale as Figure 6.

There were seven multiple-choice questions on a Likert scale asked in the questionnaire. The distribution of answers can be seen in Figure 6, where we can see that while many students answer 4 or 5, there is a large group of students that have trouble in these areas. Questions 2 and 3, about focus outside of class and planning, respectively, have the lowest score, which corroborates the results from the retrospective data analysis. The difference between the answers to questions 1 and 2 are noteworthy, as it means students have an easier time focusing on schoolwork in class than at other times. In the diagram on the right, we can also see that focus out of class is similar for years ten and eleven, but that year twelve has a harder time. Looking at Figure 6 we can see the question about focus as it relates to the different personas we asked the students to identify with. Persona C students clearly have an easier time focusing in either setting than the others, and there is not much difference in their ability to focus in class or out, while personas A and B have a harder time focusing, especially outside of class.

Theme	Category 1	Category 2	Category 3	Category 4	Category 5
1. Internal distractions	**1.1 Tiredness** 161	**1.2 Lack of motivation** 121	**1.3 Difficulties focusing** 81	**1.4 Hunger** 5	**1.5 Psychological reasons** 76
2. External distractions	2.1 Other people 70	2.2 Sound or volume 47	2.3 Electronics 92	2.4 Other priorities 81	2.5 Other distractions 21
5. Support for planning	5.1 Tools 39	5.2 Needs for planning 70	5.3 Needs for execution 21		
6. Motivational needs	6.1 Utilization of knowledge 34	6.2 Other motivational factors 6			

Figure 8
Themes and categories related to RQ1

For better insight into these problems we can look at the result of the thematic analysis, where themes 1, 2, 5 and 6 provide some answers to our first research question.

Theme 1: Internal distractions

By internal distractions we are referring here to what causes the student to become distracted from their current task, where the students answer in ways that place the source mainly within themselves. Four separate categories were identified within this theme. Category 1.1 is **Tiredness** (161 occurrences), and consists mostly of answers stating just that, either in just the one word, or with some context, for example "*I'm too tired*" (question 4a), "*when I'm too tired,*" or "*haven't slept enough*" (question 1a), or "*I usually procrastinate because I don't have the energy to do it. It also takes longer to do it (even though I would probably save time by planning).*"

Category 1.2 is **Lack of motivation or interest** (121 occurrences), which typically consists of answers like "*it's boring*" or "*I'm not interested in the assignment*" for question 1a, but like tiredness, these kinds of answers appear in response to several questions. In reference to lack of planning, we find answers like "*I should, but I don't feel like it,*" and "*it never works to plan more than a few hours ahead, because suddenly you get an assignment in the last class that is due earlier than the assignment you had planned to work on.*" 1.3 **Difficulties focusing** (81 occurrences) is mainly about questions 1a and 2a, and relates to different ways that the students encounter problems in focusing on the task at hand: "*zone out,*" "*my brain wanders,*" "*I procrastinate*" are typical answers. 1.4 **Hunger** (5 occurrences) consists of a few answers to question 1a, about lack of focus during lessons. Category 1.5 **Psychological reasons** (76 occurrences) is the final category in the first theme, and consists of answers relating to psychological reasons other than the ones covered by other categories. This means for example stress, depression or simply forgetfulness.

Theme 2: External distractions

With external distractions we mean those which distract the students from their tasks where the student identifies a source other than themselves. The first category here is 2.1 **Other people** (70 occurrences). In the school (question 1a) this tends to be about other students, from simply "*classmates*" or "*friends who are unfocused,*" to the more specific "*when students from other classes come into our lesson and disturb us.*" This category also appears in relation to question 2a, where the answers center around friends, but are also about family "*I need to help out at home,*" or "*people (mostly family) need me to do stuff every 20 minutes.*" The next category, 2.2 **Sound or volume** (47 occurrences) are mostly about disturbing sounds, or about the noise level being too high in the classroom. "*People talk too much," "it's too loud."* A category that doesn't come up as much as a source of distraction inside the classroom, compared to focusing outside the classroom is 2.3 **Electronics** (92 occurrences). By this we mean answers about mobile phones, computers, games or the like. Most answers in this category are simply those: "*Games,*" "*YouTube,*" "*my phone*". In category 2.4 **Other priorities** (81 occurrences) we find answers where the student makes clear that they make a choice to focus on other things, like "*I want to do other things,*" "*going to the gym,*" "*exercising,*" or "*work.*" Finally, the last category, 2.5 **Other distractions** (21 occurrences) is a category for those answers that are too unspecific to fit in the other categories. Most of these simply state "*distractions*" or "*other stuff*".

Theme 5: Support for planning

The fifth theme we call Support for planning, and here we collect answers that relate to the active task of planning. The first category in this theme is 5.1 **Tools** (39 occurrences), which consists of answers that are about which tools students use, or would like to use, for their planning. Among these we mostly find digital calendars, the learning platform, other digital tools or in a few cases physical calendars or paper. The second, and largest category is 5.2 **Needs for planning** (70 occurrences), which relates to those answers pointing out needs to improve the planning process, such as "*some information about the most effective ways of planning, facts about circadian rhythms and times of day when you work efficiently, etc.*," "*that someone explains how one should prioritize and plan,*" and "*how to study best in different subjects.*" In this category we also find a group of answers relating to goal setting, for example "*a clear goal and where I am in relation to that goal*" and "*how to create goals that suit you*". Finally, the category 5.3 **Needs for execution** (21 occurrences) consists of those answers that call for support in the execution of the plan, such as "*a system that forces me to work,*" or "*automatic notifications that remind me of what happens next week and that reminds me to do school work at home.*"

Theme 6: Motivational needs

The sixth theme is motivational needs. By this we mean answers where students point out things that they feel that they would need in order to be better motivated for doing school work. This consists of two categories, where the first one is 6.1 **Utilization of knowledge** (34 occurrences), which consists of answers that ask for the knowledge they are meant to acquire in school to better be related to real world applications. "*Use what you do for something practical that I can actually have use for,*" "*areas of use for what I learn and connections between all the things I learn.*" "*Things that make the world better.*" Finall, 6.2 **Other motivational factors** (6 occurrences) is a category consisting of a few answers in the vein of competitions, or more challenges.

3.3 RQ2: What information and data do the students need in order to better regulate their own learning?

For RQ2, we mostly look to the thematic analysis, where the main result is that while students do report that they have access to and use information, mostly information about test times and deadlines, they also report a need for this same kind of information, suggesting that the temporal information they have access to is incomplete or inconsistent. They also ask for clearer information, and for information about the planning and study process itself.

Theme	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6
3. Information available	3.1 Learning platform 36	3.2 Temporal information 98	3.3 Other people 19	3.4 Formal information 33	3.5 Personal knowledge related information 13	3.6 Planning related information 5
4. Information needs	4.1 Temporal information 76	4.2 Need of clarity in information 83	4.3 Personal knowledge-related information 22	4.4 Personal development-related information 59	4.5 Planning-related information	4.6 Information organization 17

Figure 9
Themes and categories related to RQ2

Theme 3: Information available

The third theme is about answers related to what information the students have and utilize today for planning and for their own development. The answers fall into six different categories. 3.1 **Learning platform** (36 occurrences). Several students point to the learning platform used at the school (Canvas) as a source of information, but the main sources of information seem to be the ones falling into category 3.2 **Temporal information** (98 occurrences), where students point to the school calendar. "*Schedule,*" "*school calendar,*" "*I use a digital calendar*" are among the answers here. Some students answer that they get information from "*other students*" or a "*teacher that reminds us when we have a deadline,*" which we categorize as 3.3 **Other people** (19 occurrences). As 3.4 **Formal information** (33 occurrences) we have categorized answers that bring up formal documents like grading criteria, course plans or course books as sources of information. In category 3.5 **Personal knowledge related information** (13 occurrences) we have some answers that mention that the students use information such as preliminary grades and feedback from teachers for their personal development, and finally we have a few answers that fall into the 3.6 **Planning related information** (5 occurrences), which is about tips for studying, or the school's work with study technique.

Theme 4: Information needs

The fourth theme is about what the students themselves identify as information they need, either for planning, for their own development, or for motivational reasons. The first category here is 4.1 **Temporal information** (39 occurrences), which identifies the need for information about deadlines and tests. Some of these only say "*deadlines,*" while others are more specific, such as "*telling us what to hand in and when and do it in advance and not a few days before deadline.*" The largest category in this theme is 4.2 **Need of clarity in information** (70 occurrences). "*Examples of solutions,*" "*clearer assignments,*" "*that the teacher is clear about what should be studied and the amount of studying timewise so that you can plan accordingly. Also remind that you should study,*" and "*much clearer guidelines for what one should be able to do for attaining a specific grade,*" are some of the answers fitting into this category.

We then have two categories that are closely related, which we call 4.3 **Personal knowledge-related information** (22 occurrences) and 4.4 **Personal development-related information** (59 occurrences). In these two categories we have student answers that bring up the need for information about what knowledge level or grade level the student is currently at (4.3), and information about what they need to do to advance their knowledge and skills (4.4). Some variation of "*how am I doing?*"[1] is the most common need pointed out here. This information can also be related to Figure 4.9, where we can see that it is primarily persona A students who feel that they don't have enough of this kind of information. 4.5 **Planning-related information** (10 occurrences), consists of answers that point out the need for information that is vital to planning, such as the scope of the assignment, or how important that assignment is for the course grade. Finally, we have the category 4.6 **Information organization** (17 occurrences), which are requests for information to be

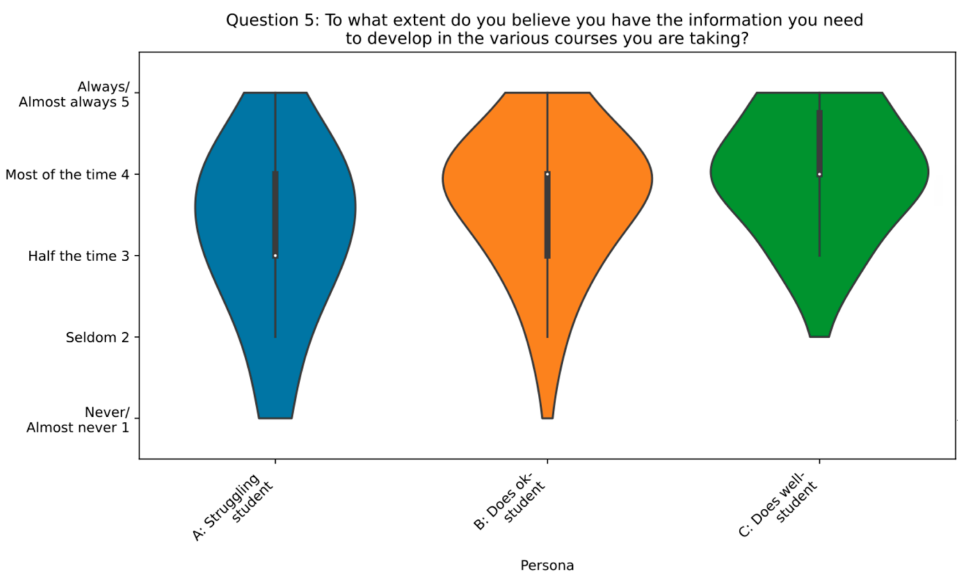


Figure 10
Question 5 (To what extent do you believe you have the information you need to develop in the various courses you are taking?) arranged by persona.

[1] In Swedish: "Hur jag ligger till". The Swedish expression is more clearly referring to a result, in this case a grade, than may be clear from the English translation.

organized in a better way than currently, "*some easy digital tool where everything is collected in the same place,*" "*one place where all the week's homework and tests are written. In one place, not spread out on a webpage where I have to search for things.*"

4 Discussion

4.1 Challenges that students experience in the process of regulating their own learning

The analysis shows that students face a number of different challenges when regulating their own learning, especially in the planning and performance phases, stressing the need for supporting these phases, which may be lacking proper support in existing learning analytics platforms (Jivet et al., 2017). The major factors can be found in the two themes internal and external factors, where the internal factors like fatigue, lack of motivation, and psychological health are the most common obstacles the students face. In relation to psychological health, this is a pressing issue in need of addressing, but it is outside the scope of this study, and as such will not be addressed in detail. Among external factors, it is clear that the two main challenges are the actions of other students in the classroom, and electronic devices, especially mobile phones, outside the classroom. There are, of course, synergies between these to take into account. It is naturally more difficult to be motivated when you are very tired, and all of the internal factors can exacerbate the distractive nature of electronic devices. It has also been suggested that learning analytics in supporting learning can also increase motivation (Aguilar et al., 2021) Finally, one challenge mentioned by several students (see Fig. 9), is that the students simply don't feel that they know how to plan, or how to study, which is in line with Winne's research about students' needs for better tools and knowledge about how to regulate their learning (Winne, 1995, 2005, 2022).

When looking at how the results differ by persona, the most striking result is that persona C is quite stable across all the different questions, while personas A and B have clear problems with both planning and focus, suggesting that the issues that students face are not limited to a small group of students, but that it is rather a small group of students who are doing well when it comes to regulating their learning, while the majority struggle.

4.2 Information students need to support their learning

The data that the students need is primarily 1) information about tests and deadlines, 2) clarity about information and assignments, and 3) how to improve in their school subjects. While the students report that they do have access to information of type 1 that they use for planning, it is currently incomplete, inconsistent, and spread out across different channels and platforms. Even more than temporal information, students report a need for increased clarity in information, saying that they often don't understand the task nor what is required or expected of them. The third type of information needed is information about what they as students need to improve upon and how to improve. This is all information that is generally provided by teachers. Further, the data the students acquire could be scaffolded through technologies.

It has been shown that learning analytics can provide feedback that helps students with what they need to focus on to improve (Afzaal et al., 2021b), and it should be quite possible to build systems that can ease the process of communicating both information and the lack thereof. As for the clarity of information, it should be quite possible to take advantage of the recent growth in generative AI to accommodate the need for clarification and examples that students signal, as exemplified recently by Mollick & Mollick (2023). Students also suggest that they need prompts, or notifications, to help remind them and push them into action, another area that should be perfectly suitable for learning analytics.

4.3 Balancing needs and wants

Some students' answers (see e.g. Theme 4: information needs above) seem to place the full weight of improving their existing situation on the teacher, seemingly abdicating from regulating their learning and preferring the teacher to do the regulating. This is a dangerous road to take if the goal is to improve the student's self-regulation. Others put the full responsibility on the students themselves which, considering the number of students who have problems with the current situation, seems an unproductive stance to take. Winne (2022, p. 775) claims that "learners need to be able to perform learning tactics and strategies without undue effort. Otherwise, excessive cognitive load or inept execution of those skills would worsen rather than enhance progress on academic tasks." This, combined with the earlier mentioned suggestion that too much scaffolding harms the students' regulatory skills (Duffy & Azevedo, 2015), suggests that addressing the needs identified by the student has to be balanced so that they can scaffold the students in what they need without taking over the task of regulation.

4.4 Limitations of the study

This study was conducted at a single school, with a limited number of students who had all chosen the same programme. This was due partly to the existing data, which had been collected at this particular school, and unfortunately not in a broader context. It would be valuable to see if the long-standing trends at this one school can also be found at other schools and within other programs. This limitation means that several groups of students are underrepresented, or not represented at all, such as the lowest-achieving students who don't have the grades necessary to get accepted into that school. Further, the school has an IT focus, and many such schools have an underrepresentation of female students. The methodology of analyzing anonymous survey data also has its limitations, in that we are not able to ask for clarifications or follow-up questions when we get answers that lead us to further questions. The chosen method of looking at long-term trends in existing data to provide the focus for further investigation was a rewarding choice in this study and put the research more in line with the needs of the teaching profession, but it also means we may have missed important aspects that were not part of that long-term data.

With these limitations, it is imperative that these results are not taken as generalizable to secondary students at large, but it is our hope that what we have seen here can provide insight and knowledge that can be applicable in similar contexts.

Conclusions and implications

The implications of this study can be divided into two main areas: Implications for educational practice, and implications for systems design. The practical implications of these results are that secondary education does need to improve its working environment to minimize distractions, in accordance with previous research (Schmidt, 2020), and that there is a lot of work that needs to be done about the psychological health of secondary education students. There is also clear support for more explicit instruction on how to regulate learning, in line with the findings of Dignath & Veenman (2021). Lastly, students find that information is fragmented, inconsistent and unclear, which is an area where systems design can help, but which is at its core also something that educational practitioners have to consider.

When it comes to the implications for systems design, there is a pressing need for systems that not only communicate information from teachers to students, but also augment that information with analytics to support students where the existing information flow is lacking. This is in line with previous research. However, findings from this study highlight that while trace data of student activities and regular learning analytics methods often used (e.g. Heikkinen et al., 2023; Winne, 2022) are important, there are other areas that may need other techniques to address the need for organizing and clarification of existing information. This is an area where recent advances in AI, specifically large language models, may be helpful. This is a direction for research that would be important to investigate further. Systems need to be designed with the students' regulatory processes in mind, and with built-in support for scaffolding self-regulation, since our results suggest it may be quite difficult not only for some, but for the majority of students. Support for the planning and performance phases seem to be the most crucial, where planning, adhering to the plan and adapting plans to changing circumstances stand out as the most challenging parts. In previous research, Jivet et al. (2017) found that few interventions targeted the planning phase, while Heikkinen et al. (2023) found 45% of the studies they look at supporting this phase, suggesting that there has been improvement in the attention to planning, but that it may still need more attention considering its importance to students, as shown in our results.

As for further research implications, we see that there is a need to look at how systems for secondary education can be designed in an integrated manner, in line with Jivet et al. (2017), to support the full range of needs found in this study, as well as studying the effects of such a system to make sure the scaffolding is designed in such a way that it supports the development of students' regulatory processes and does not replace them. Most of the research in this field is limited to higher education (Heikkinen et al., 2023; Schwendimann et al., 2016), calling for more research looking at what aspects of learning analytics support for students also translate to secondary education.

References

Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., Li, X., & Weegar, R. (2021a). Automatic and intelligent recommendations to support students' self-regulation. In M. Chang, N.-S. Chen, D. G. Sampson, & A. Tlili (Eds.), *2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 336–338). IEEE. https://doi.org/10.1109/ICALT52272.2021.00107

Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., Li, X., & Weegar, R. (2021b). Explainable AI for data-driven feedback and intelligent action recommendations to support students self-regulation. *Frontiers in Artificial Intelligence*, *4*, Article 723447. https://doi.org/10.3389/frai.2021.723447

Afzaal, M., Zia, A., Nouri, J., & Fors, U. (2023). Informative feedback and explainable AI-based recommendations to support students' self-regulation. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-023-09650-0

Aguilar, S. J., Karabenick, S. A., Teasley, S. D., & Baek, C. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, *162*, Article 104085. https://doi.org/10.1016/j.compedu.2020.104085

Álvarez, R. P., Jivet, I., Pérez-Sanagustín, M., Scheffel, M., & Verbert, K. (2022). Tools designed to support self-regulated learning in online learning environments: A systematic review. *IEEE Transactions on Learning Technologies*, *15*(4), 508–522. https://doi.org/10.1109/TLT.2022.3193271

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189X11428813

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, *31*(6), 445–457. https://doi.org/10.1016/S0883-0355(99)00014-2

Bolhuis, S., & Voeten, M. J. M. (2001). Toward self-directed learning in secondary schools: What do teachers do? *Teaching and Teacher Education*, *17*(7), 837–855. https://doi.org/10.1016/S0742-051X(01)00034-8

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Dignath, C., & Veenman, M. V. J. (2021). The role of direct strategy instruction and indirect activation of self-regulated learning—evidence from classroom observation studies. *Educational Psychology Review*, *33*(2), 489–533. https://doi.org/10.1007/s10648-020-09534-0

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348. https://doi.org/10.1016/j.chb.2015.05.041

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503. https://doi.org/10.1037/0003-066X.54.7.493

Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions – a systematic literature review. *Education and Information Technologies*, *28*(3), 3059–3088. https://doi.org/10.1007/s10639-022-11281-4

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning* (pp. 82–96). Springer. https://doi.org/10.1007/978-3-319-66610-5_7

Lodge, J. M., Panadero, E., Broadbent, J., & de Barba, P. G. (2018). Supporting self-regulated learning with learning analytics. In J. Lodge, J. Horvath, & L. Corrin (Eds.), *Learning Analytics in the Classroom* (pp. 45–55). Routledge. http://dx.doi.org/10.4324/9781351113038-4

Matcha, W., Uzir, N. A., Gašević, D., & Pardo, A. (2020). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence*, *3*, Article 100070. https://doi.org/10.1016/j.caeai.2022.100070

Mollick, E. R., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts* (SSRN Scholarly Paper 4475995). https://papers.ssrn.com/abstract=4475995

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research*, *45*(3), 269–286. https://doi.org/10.1080/00313830120074206

Schmidt, S. J. (2020). Distracted learning: Big problem and golden opportunity. *Journal of Food Science Education*, *19*(4), 278–291. https://doi.org/10.1111/1541-4329.12206

Schunk, D. H., & Greene, J. A. (2017). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk, & J. A. Greene (Eds.), *Handbook of Self-Regulation of Learning and Performance* (2nd Ed.) (pp. 1–15). Routledge. https://doi.org/10.4324/9781315697048-1

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2016). Understanding learning at a glance: An overview of learning dashboard studies. In D. Gašević, G. Lynch, S. Dawson, H. Drachsler, & C. P. Rosé (Chairs), *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge – LAK '16* (pp. 532–533). Association for Computing Machinery. https://doi.org/10.1145/2883851.2883930

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Winne, P. H. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. *Educational Psychologist*, *30*(4), 223–228. https://doi.org/10.1207/s15326985ep3004_9

Winne, P. H. (2005). A perspective on state-of-the-art research on self-regulated learning. *Instructional Science*, *33*(5–6), 559–565. https://doi.org/10.1007/s11251-005-1280-9

Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning*, *17*(3), 773–791. https://doi.org/10.1007/s11409-022-09305-y

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5

Zimmerman, B. J. (2000). Attaining self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In J. Douglas, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299–315). Routledge.

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-5

A SCOPING REVIEW OF WEBCAM EYE TRACKING IN LEARNING AND EDUCATION

Nicol Dostálová[a], Lukáš Plch[a]

[a] Faculty of Arts, Masaryk University, Czech Republic

ABSTRACT

The use of eye tracking in educational research has shown great potential in recent decades. There are various approaches to the usage of eye tracking technology in this area, including investigation of self-regulated learning from different types of learning environments. Nonetheless, the majority of published research studies have one tremendous limitation: using expensive remote or tower-based eye trackers to provide high-quality data in laboratory conditions. Now, new webcam eye trackers may offer an easily affordable approach allowing eye tracking measurements in the real environment, such as the investigation of learning behavior in online learning environments. The main aim of this scoping review is to explore the use of webcam eye tracking technology in the field of learning and education. We established three specific purposes: 1) to introduce educational topics being explored using webcam eye tracking, 2) to discuss the methodological aspects when exploring educational topics with webcam eye tracking, and 3) to investigate the eye tracking aspects used for the analysis. To do this, we analyzed 16 studies that used webcam eye tracking. The results of the scoping review show that 1) selected studies focus mainly on students' behavior in online learning environments, such as engagement, lack of attention, cheating and others; 2) a wide range of studies aimed at the development of automatized detection tools; and 3) studies are mainly focused on extracting raw and event data features using them mostly for automatized detection purposes.

KEYWORDS

self-regulated learning; online learning environments; eye tracking; webcam eye tracking

CORRESPONDING AUTHOR

Nicol Dostálová, Department of Educational Sciences, Faculty of Arts, Masaryk University, Arna Nováka 1, 602 00 Brno, Czech Republic
e-mail: nicol.dostalova@mail.muni.cz

Introduction

The use of eye movement tracking has been on the rise in the field of education and learning in recent decades. Eye tracking technology is able to provide researchers with a wide range of information about the metacognitive, and behavioral processes of learners (Antonietti et al., 2014). Eye tracking technology contributes to the development of educational processes and the individuals themselves. Eye tracking is used in a large number of sectors within education, e.g., processes in the classroom, teaching and learning in virtual reality, reading research (Rayner, 1998), and attention, perception, and language learning (Šmideková, 2018; Lai et al., 2013). Lai et al. (2013) also mention topics of metacognition and learning strategies. Another broad area of educational research in which eye tracking and analysis of eye movement data can be used extensively is in the context of multimodal learning analytics (MLA). In general, MLA refers to the integration or the elicitation of different data from multiple sources (e.g., audio, video, eye tracking, biosensors and more; see Blikstein & Worsley, 2016 or Worsley, 2018) that can subsequently provide crucial insights regarding an individual's learning process from a variety of perspectives (Ochoa, 2017). Such analyses can provide assessments of students' knowledge, behavior, intentions, or even physiological characteristics (Blikstein & Worsley, 2016), which can help to increase the holistic understanding of an individual. Alemdag and Catilgay (2018) then explore in more depth the area of eye tracking research (particularly with remote and webcam eye tracking devices) in multimedia learning, which can be closely related to the online learning environments and the learner's behavior with learning materials, and their metacognitive and self-regulatory abilities, which are crucial for successful and effective learning from e-learning materials.

Self-regulated learning has become crucial in the last two decades with the expansion of learning with online learning environments where students must demonstrate sufficient self-regulation skills, such as motivation, strategic planning, responsibility, and time management to learn effectively (Panadero, 2017). Self-regulated learning is described as a learning process that is built on cognitive strategies, motivation, and metacognitive skills. At the same time, responsibility and autonomy are also necessary for successful self-regulated learning (Carneiro et al., 2011). Zimmerman (2000) describes self-regulated learning as a cyclical process that is composed of three phases. These phases include the preparatory phase, the performance phase, and the reflective phase. However, exploring self-regulated learning in online learning environments is relatively challenging. Researchers often focus on self-reports and questionnaires, which can be subjective (Dostálová et al., 2022). For this reason, the focus in research on self-regulated learning has begun to shift simultaneously to the use of new technologies and to the collection of psychophysiological data that can point to behavioral patterns in individuals' learning that were not previously apparent. The method of tracking eye movements is included among such technologies (e.g., Antonietti et al., 2014).

Tracking eye movements can contribute greatly to uncovering self-regulated processes during learning (e.g., Taub & Azevedo, 2018) in the sense of detecting what areas a student has looked at during the learning process, for how long, and possibly in what sequence. However, eye tracking research in the area of online learning environments presents major challenges in the form of lab-based data gathering. Currently, some of the eye tracking devices that are commonly used for research are most often based on the principle of pupil and corneal reflection and varying levels of sampling frequency, precision and accuracy (Holmqvist et al., 2011). However, these devices can be relatively sensitive to the conditions in which the measurements are conducted. These in-lab eye tracking devices can also be relatively expensive (Semmelmann & Weigelt, 2017), and so the methodological aspects and the data collection are usually rather more limited and simplified, which may reduce the ecological validity of the research since the participant is not measured under natural conditions (Papoutsaki et al., 2016).

For such reasons, it is necessary to pay attention to the new technological possibilities regarding both the availability of eye tracking devices and data gathering in ecological settings (Wisiecka et al., 2022). A new approach can be provided by webcam eye tracking. Webcam eye tracking is based on the principle of face landmark detection and a machine learning approach for gaze position prediction (Wisiecka et al., 2022). Recent studies have sought to compare the accuracy and precision of the webcam eye tracking solution with commercial eye trackers showing reliable results depending on, e.g., experimental stimuli (for detailed information about the eye tracking device parameters and experimental settings see Burton et al., 2014; Skovsgaard et al., 2011; Wisiecka et al., 2022). Nonetheless, Wisiecka et al. (2022) suggest that more replications are needed in this field to properly consider all aspects of the webcam eye tracking solution (e.g., research topic, procedure, experimental stimuli, and more).

However, the main advantage of a webcam eye tracking device is the possibility of its use by nearly anyone, since only a device with a webcam is needed to conduct the research, and therefore a higher ecological validity of the measurements might be enabled, which can be significantly useful, e.g., in research on self-regulated learning from online learning materials and also to support the multimodal learning analytics during learning processes. This research aims to investigate whether and how eye tracking can be used in education and learning and thus contribute to the further use and development of this technological approach.

1 Methods

This scoping review study aims to provide and analyze an insight into the current state of knowledge on the use of webcam eye tracking in the field of learning and education. To this end, a main research question and then two specific research questions were established as follows:

RQ1: How is webcam eye tracking technology used in the field of education and learning?

RQ1.a: What fields of education and learning are explored with webcam eye tracking technology?

RQ1.b: How is webcam eye tracking used from the methodological perspective and what aspects of eye tracking are considered in selected studies?

In order to answer the research questions, which are exploratory in nature, we chose to do a scoping review. We followed the methodology set out by Tricco et al. (2018) and Munn (2018).

1.1 Inclusion and exclusion criteria

Inclusion and exclusion criteria determine which documents will or will not be included in the search to meet the study objective set by the research questions. In our case, the inclusion criteria are eye tracking, eye movement, webcams, and school or university settings. For the summary of all inclusion and exclusion criteria, see Table 1. For the review, we concentrated on studies published between 2010 and 2023, and we also decided to work in a broader scope in this review, not only with studies of the "article" type but also with "conference papers." This is mainly because webcam eye tracking is a relatively new technology that has developed more widely in the last two decades. Furthermore, conference proceedings papers are a common type of document in this field.

Table 1
Table summarizing the inclusion and exclusion criteria.

Inclusion criteria	**Exclusion criteria**
• Webcam eye tracking • Learning and education • "Peer-reviewed article" and "conference paper"	• Language other than English • Documents published outside the 2010-2023 period • Document type other than "peer-reviewed article" or "conference paper" • Absence of webcam eye tracking • Absence of learning and education field • Description of methodological or technological aspects missing

1.2 Literature search

Literature searches were conducted on April 30, 2023 in the databases Scopus, Web of Science, MEDLINE/PubMed, Embase, PsycINFO/PsyARTICLES by EBSCOHost, Education Resources Information Center (ERIC), Academic Search Ultimate by EBSCOHost and IEEE. We used the following search string, which we tailored to each search platform. For the purposes of the literature search, we established the keywords listed in Table 2.

Table 2
Key concepts and search terms used for the literature search.

Key concept	**Search term**
Eye tracking	((eye* OR ocular*) AND (track* OR movement*)) OR ("eye movement*" OR "movement of the eye*" OR "ocular movement*" OR gaze-track*) AND
Webcam eye tracking	(webcam* OR webcast* OR digicam*) OR ((web OR digital) AND (camera*)) AND
Education and learning	educat* OR learn* OR teach* OR study* OR student* OR instruct* OR pupil* OR school* OR universit* OR college*

1.3 Data extraction

Data extraction was carried out according to the Joanna Brigss Institute scoping review methodology guidelines (*11.2.7 Data Extraction*, 2022) in order to extract data on the authors of the publication, year of publication, origin, research objectives, population studied, research methodology, and findings. See Table 1 for the extraction results.

1.4 Critical appraisal

This is an optional part of the scoping review, and was not conducted in our case.

1.5 Data synthesis

The objectives of the data synthesis were descriptive qualitative content analysis. The tool used to conduct it was open coding. A simple frequency count was used for descriptive statistical analysis of quantitative data (Aromataris & Munn, 2020).

1.6 Data management and screening

Data were processed in the online software Rayyan https://www.rayyan.ai/, where they were deduplicated and screened independently by both authors of the study. First, studies were screened by reading titles and abstracts and in the second stage by full-text analysis. The key measures of whether to include or exclude studies were inclusion and exclusion criteria. If there were any disagreements, these were resolved in online meetings via MS Teams.

1.7 Ethical considerations

This scoping review does not require ethical approval. All data are gathered from publicly available sources, either licensed or open access.

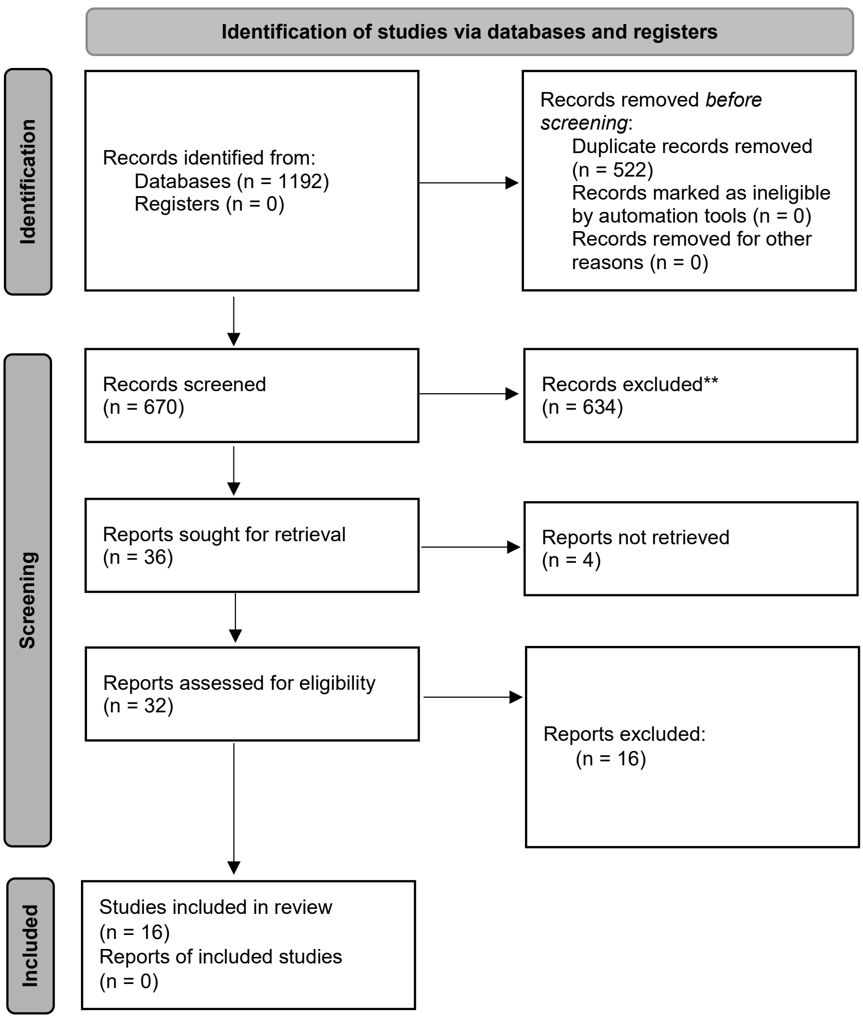


Figure 1
Prisma diagram showing the process of studies selection

1.8 Data analysis

The analysis of the selected articles commenced by scrutinizing the research objectives, inquiries, methodologies, and study outcomes. We systematically extracted pertinent information from the articles, which encompassed metadata such as title, authors, DOI and name of a journal or conference. We also extracted the aim of a study, sample, research method and eye tracking approach.

2 Results

This scoping review presents a summary of 16 articles and conference papers that combine two key topics: the use of webcam eye tracking and the thematic areas of education and learning. We decided to divide the results into two main categories according to the focus of the specific research questions. The following two categories are: 1) the field of interest regarding education and learning, and 2) methodological and eye tracking aspects of selected studies.

2.1 Exploring educational horizons: Webcam eye trackers and learning domains

The main aim of this section is to present the possibilities for using webcam eye tracking in teaching and learning that have already been explored in various contexts.

From a general perspective, eye tracking technology is used to investigate cognitive functional processes. This has been addressed by Lin et al. (2022), who focused on testing the feasibility of a webcam eye tracking interface for commonly used cognitive tasks. They subsequently tested the webcam interface on Chinese reading tasks. Reading is a complex cognitive ability, and the reading tasks were also applied by Guan et al. (2022) pointing out the importance of low-cost eye tracking data gathering in natural conditions, especially in the era of digital and online learning, using webcam eye tracking to explore the relationship between reading behavior and reading performance. Cognitive processes related to reading comprehension and mind wandering during an online reading comprehension task are also discussed by Hutt et al. (2022). In the case of a specific reading disorder, learning how to read and the reading itself can be a very challenging task. Calabrich et al. (2021a, 2021b) explored cross-modal bindings and episodic memory in dyslexic and intact adult readers. The reader's attention is also closely related to working with the text and reading. A reading task for a group of neurodivergent students (ADHD, autism, and learning disability) was also used by Wong et al. (2023) who explored the use of webcam eye tracking to support these students during learning. Li et al. (2016) took a different approach to reading

Table 1
Resumé of extracted data

Authors	**DOI/URL**	**Title**	**Journal/Conference**
Alkabbany et al. (2023)	10.3390/s23031614	An Experimental Platform for Real-Time Students Engagement Measurements from Video in STEM Classrooms	Sensors
Behera et al. (2020)	10.1007/s40593-020-00195-2	Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems	International Journal of Artificial Intelligence in Education
Calabrich et al. (2021)	https://escholarship.org/uc/item/76b3c54t	Episodic Memory Cues in Acquisition of Novel Visual-Phonological Associations: a Webcam-Based Eye tracking Study	43rd Annual Meeting of the Cognitive Science Society
Calabrich et al. (2021)	10.3389/fpsyg.2021.754610	Audiovisual Learning in Dyslexic and Typical Adults: Modulating Influences of Location and Context Consistency	Frontiers in Psychology
Dilini et al. (2021)	10.1109/ICITR 54349.2021.9657277	Cheating Detection in Browser-based Online Exams through Eye Gaze Tracking	6th International Conference on Information Technology Research (ICITR)
Guan et al. (2022)	https://ceur-ws.org/Vol-3120/paper5.pdf	An analysis of reading process based on real-time eye tracking data with webcamera—Focus on English reading at higher education level	4th Workshop on Predicting Performance Based on the Analysis of Reading Behavior
Hutt et al. (2023)	10.3758/s13428-022-02040-x	Webcam-based eye tracking to detect mind wandering and comprehension errors	Behavior Research Methods

Key topic	**Population sample**	**Research method**	**ET approach and device**
Real-time automatic student engagement measure based on behavioral and emotional engagement	10 students of 300-level STEM classes	In-wild data collection (lecture); each lecture is 75 min in length and was divided into windows (1360 samples)	Webcam ET system (2–3 frames per second); features: head pose, eye gaze, action units;
Automatic detection of body, gaze movements and gestures while computer-mediated learning	9 male undergraduate students	In-wild data collection; deep learning approach for automatic detection; two separate sessions: (1) five students in one session and (2) four in another session	Webcam IntraFace tool (15 Hz sampling rate); the eye gaze is recorded for both eyes
Spatial and contextual cues to aid learning of novel cross-modal bindings	14 university students (1 participant excluded)	Experiment; two-phase measurement (recognition task and cued recall)	Webgazer.js (60 Hz); eye movements not recorded during the second phase
Analyzing the contribution of episodic memory cues to cross-modal bindings (readers with and without dyslexia during learning process)	70 university students (35 dyslexic and 35 non-dyslexic readers)	Experiment; 3 tasks: (1) a training recognition task with interspersed cued-recall trials; (2) a post-training cued-recall test; (3) a post-training recognition test	Webgazer.js (60 Hz/depending on the refresh rate of the monitor); features: fixations in AOIs
Browser-based cheating detection approach in online examinations through eye gaze tracking	5 undergraduate participants; evaluation on 10 undergraduate students	Step 1: creating the ET dataset (cheating and non-cheating), step 2: verifying predictive models	Features: gaze in/out of the screen
Developing platform using web camera eye tracking to get physiological indicators based on eye tracking data to analyze the reading process	35 higher education university students	Experiment: pre-test, English reading test, interview	Webgazer.js features: frequency and rate of fixation on page; frequency of regressions (RS) among pages
Goal 1: detecting comprehension and task-unrelated thought in real-time using webcam eye tracking Goal 2: using eye tracking system to reproduce and build quantified relationships between these learning-related cognitive constructs and eye gaze	study 1: 105 university students study 2: 173 participants	Experiment: between-participants manipulation with participants randomly assigned to one of two conditions; narrative anticipation task	Webgazer.js; features: global gaze features (number of samples, dispersion of gaze points), local gaze features (patterns inside of AOI)

Khan et al. (2022)	10.1109/EDUCON 52537.2022.9766506	EXECUTE: Exploring Eye Tracking to Support E-learning	2022 IEEE Global Engineering Education Conference (EDUCON)
Khosravi et al. (2022)	10.1109/EDUCON 52537.2022.9766468	Self-Directed Learning using Eye Tracking: A Comparison between Wearable Head-worn and Webcam-based Technologies	2022 IEEE Global Engineering Education Conference (EDUCON)
Li et al. (2016)	10.1145/3015297.3015301	Multimodal Human Attention Detection for Reading from Facial Expression, Eye Gaze, and Mouse Dynamics	ACM SIGAPP Applied Computing Review
Lin et al. (2022)	10.1016/j.bspc.2022.103521	An eye tracker based on webcam and its preliminary application evaluation in Chinese reading tests	Biomedical Signal Processing and Control
Madsen et al. (2021)	10.1073/pnas.2016980118	Synchronized eye movements predict test scores in online video education	Proceedings of the National Academy of Sciences (PNAS)
Robal et al. (2018)	10.1145/3172944.3172987	Webcam-based Attention Tracking in Online Learning: A Feasibility Study	IUI '18: 23rd International Conference on Intelligent User Interfaces
Wong et al. (2023)	10.1145/3576050.3576115	Using a Webcam-Based Eye tracker to Understand Students' Thought Patterns and Reading Behaviors in Neurodivergent Classrooms	13th International Learning Analytics and Knowledge Conference (LAK23)
Yi et al. (2015)	10.1109/ CADGRAPHICS.2015.13	Real Time Learning Evaluation Based on Gaze Tracking	14th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics)
Zhao et al. (2017)	10.1007/978-3-319-66610-5_24	Scalable Mind-Wandering Detection for MOOCs: A Webcam-Based Approach	12th European Conference on Technology Enhanced Learning

Developing an e-learning framework for capturing and analyzing the students' attention during remote teaching sessions	25 participants	Measurement for raw data; robust machine learning approach	Webgazer.js (29 Hz); features: raw gaze points, fixation, saccade, eye blinks in pyGaze; number of fixation and fixation duration in AOI, duration of largest fixation backtracks fixation dispersion, first fixation duration and entry time of the first fixation in AOI
Using a low-cost webcam eye tracking solution for e-learning materials and its comparison to the head-mounted ET	8 learners	Measurement: 4-minute lecture; qualitative analysis	PupilCore and Webgazer.js; features: fixation data, gaze plots, heatmaps, comparison on visualizations
Detecting human attention when reading	6 subjects	Reading articles measurement; machine learning approach for identification of attention level	Unspecified webcam ET; features: blink rate, fixation rate and duration, saccadic rate and duration
Using webcamera to capture eye movements for cognitive function assessment	Comparative experiment: 62 subjects application experiment: 72 subjects	Comparative experiment: visual latency task application experiment: reading task	SMI RED250 + Microsoft® LifeCam Studio; features: total reading time, first-pass reading time, re-reading and key press reaction time, scanpath length
Analyzing students' attention in home online education	88 subjects	Series of measurements (6 short videos)	–
Detecting a loss of focus in the online learning setting	20 regular MOOC learners	Pilot study; benchmark set of tasks (50 tasks)	Webgazer.js, tracking.js, HW based solution (Tobii ET); features: face-hit, face-miss, likely-face-miss
Examining of validity and applicability of using webcam-based eye tracking to study neurodivergent students in educational settings	43 university students	Between-subject quasi-experimental design; reading task	Webgazer.js features: AOI proportions on paragraphs
Presenting a system for extraction of eye movements information to analyze learners' behavior	4 participants (152 samples of idle, 341 samples of seeking, 530 samples of scanning)	HMM classifier training on three basic patterns: scanning, seeking and idle	Webcam ET (not specified); description of eye movement detection
Proposing an automatic detection in MOOC of learners' mind-wandering through webcam eye tracking	13 participants	Machine learning approach for detection (Logistic Regression, Linear SVM and Naive Bayes classifiers)	High-quality ET (sampling rate 30 Hz) + Webgazer.js (5 Hz); features: 58 features (parameters of fixation and saccade)

research using webcam eye tracking and concentrated on developing a system for detecting attention during reading in an e-learning environment. For this purpose, they adopted a multimodal approach, thus using information related to facial expression, eye tracking, and mouse dynamics.

Students' attention during learning from online learning materials is of great interest to Robal et al. (2018), Khan et al. (2022) and Madsen et al. (2021). Khan et al. (2022) reacted to the pandemic situation and the subsequent conversion of in-person lectures to an e-learning environment. They proposed an e-learning framework that would be able to determine student's attention levels during online learning sessions, in part to address the problem of a potential lack of self-regulation among students during online lessons. Robal et al. (2018) follow up on the issue of students not being able to adequately regulate their learning process in Massive Open Online Courses (MOOCs) and propose a tool that would be able, based on a face-capturing webcam eye tracker, to detect the loss of attention. Koshravi et al. (2022) also directed their attention to self-directed learning in the online environment, using both a webcam eye tracker and commercial eye tracking glasses to collect psychophysiological data (gaze position), and therefore, to potentially improve the quality of e-learning materials. Madsen et al. (2021) then used webcam eye tracking for an experiment addressing students' attention while watching online tutorial videos. Lack of attention and mind wandering during the learning process has also been investigated by Zhao et al. (2017), who focused on detecting mind wandering for MOOCs based on webcam gaze data. Furthermore, behavioral and emotional engagement and its automatized detection in the educational environment were explored by Alkabbany et al. (2023). The analysis of learners' general behavior in e-learning was the focus of Yi et al. (2015), who used machine learning to analyze eye movements captured by a webcam for the purpose of learning patterns classification, which can lead to a better understanding of students' learning intentions and their overall behavior when studying in online environments.

A completely different area has been explored by Dilini et al. (2021), who have concentrated on the area of remote online exams and have developed eye-movement-based cheating detection for this purpose.

Based on the thematic analysis of the selected studies in the scoping review, it is apparent that the use of webcam eye tracking in the field of learning and education is relatively broad but focuses almost exclusively on the area of online or e-learning environments. With the help of simple and accessible webcam eye tracking technology, it is possible to observe learners' behavior in online learning environments (attention, concentration, engagement, etc.), and at the same time, eye movement data can be used to create classification and detection tools that could lead to the improvement of online learning environments.

2.2 Unraveling the gaze: Essential eye tracking aspects and methods

In this section, we focus on the types of webcam eye tracking devices employed in selected studies, the eye tracking metrics, and the way of working with the eye tracking outputs used in the research. Given the inconsistency of webcam eye tracking usage, we decided to proceed with this section according to the selected eye tracking device and then the aspects of eye tracking data analyzed.

The most used webcam eye tracking device in the selected studies was Webgazer.js (see Papoutsaki et al., 2016). Superficial eye tracking metrics were used for subsequent analyses. A webcam eye tracking system was used by Alkabbany et al. (2023) whose main goal was to develop an automated measurement of behavioral engagement in students. For this purpose, head position, eye gaze, and action units were considered. Eye gaze was focused where the student was looking. The webcam system collected data at a frequency of 2–3 seconds and achieved a total of 240 feature vectors (considering head pose, eye gaze, and action units), which were then processed using a support vector machine (SVM) classifier. Robal et al. (2018) worked on self-regulation in MOOCs and the detection of attention in the online learning environment. For these purposes, they used both a commercial eye tracker and Webgazer.js to track eye movements and tracking.js (see tracking.js, n.d.) for face tracking. Nonetheless, only the accuracy and reaction times were discussed for the detection development. As Dilini et al. (2021) focused on cheating detection, they used WebGazer.js eye tracking to gather a set of raw eye tracking data containing the estimated *x* and *y* positions and corresponding timestamps. These data were processed and divided into two main categories: "looking at the screen" and "looking outside of the screen." Khosravi et al. (2022) compared head-mounted eye tracking (Pupil Core eye tracking glasses) and webcam eye tracking (Webgazer.js). In this study, recorded eye tracking data was used to visualize its performance and to compare visualizations from both devices, showing similar accuracy.

Nonetheless, the next wide group of research studies focused on the use of Webgazer.js for various thematic purposes. These authors already used a broader set of eye tracking features for the subsequent analyses. For example, Hutt et al. (2022) used Webgazer.js and converted the total gaze raw data into global gaze features (general eye movement data independent of the presented stimuli, e.g., number of gaze samples, number of unique gaze samples, and variance of gaze points), and local gaze features (dependent on predefined AOIs). The AOIs approach was also chosen by Wong et al. (2023) to examine the validity and usability of webcam eye tracking as an aid tool for neurodivergent learners. Gaze data were recorded with Webgazer.js and processed in proportion to each AOI (AOI corresponds to a paragraph on the stimuli). Calabrich et al. (2021b) used Webgazer.js operating at a sampling rate of 60 Hz (depending on the screen refresh rate) to investigate audiovisual learning in dyslexic adult readers and intact adults. An algorithm to detect fixations from the raw data was used for the analysis. The position of these fixations was related to the regions of interest (ROIs) generated. Webgazer.js was also used in the research of Khan et al. (2022), who first worked with raw data and then moved on to selecting individual eye tracking features that related to predefined areas of interest, primarily related to fixations (e.g., number, variance, duration and ratio of fixations etc.). These metrics were processed through machine learning (logistic regression, SVM and polynomial regression) to create a framework to capture attention loss and engagement in e-learning environments. Guan et al. (2022) used this webcam eye tracking with a focus on analyzing reading performance. The eye tracking features selected were related to more detailed fixation parameters (e.g., frequency of fixation) or frequency of regressions on pages. These data were then statistically processed. Aiming for mind-wandering detection, Zhao et al. (2017) compared eye tracking data from a high-quality commercial eye tracker with a sampling rate of 30 Hz with a Webgazer.js sampling rate of 5 Hz. For these purposes, they selected 58 eye tracking features based on detailed parameters of fixations and saccades. Calabrich et al. (2021a) also focused on intact adult readers and tracked their eye movements while reading pseudowords using Webgazer.js. For this study, Webgazer.js was set to a frame rate of 60 Hz and focused on fixation location and consistency.

Nonetheless, several studies used a different webcam eye tracking tool to perform the measurement. Rather superficial eye tracking aspects have been considered by Behera et al. (2020) who performed their research using webcam hand-over-face gestures, head and eye movements, and facial emotions. Focusing on the eye movements themselves, the researchers used the IntraFace tool (see De la Torre et al., 2015) to directly record the left and right eye gaze data. On the other hand, Li et al. (2016) investigated attention detection in an online learning environment using a multimodal approach that includes facial expression, mouse dynamics, and eye gaze patterns. Using an unspecified webcam, they analyzed eye blinks, fixations (fixation rate and duration), and saccades (saccade rate and duration). These features are further used for machine learning analysis. Yi et al. (2015) used unspecified webcam eye tracking and described eye-movement detection for the purposes of real-time learning evaluation.

Based on the eye tracking aspects used in the selected studies, it can be evident that the most used webcam eye tracking framework is probably Webgazer.js, which the researchers used in different thematic contexts. If we focus on the selection and further analysis of the eye tracking data, we notice a relatively high diversity. In some cases, the authors focus their attention only on the superficial detection of faces and possible on-screen and off-screen gaze; in other cases, the authors work with raw data, which are processed into standard eye tracking event metrics, including fixations and saccades and their detailed characteristics as well (e.g, count or duration of fixations and saccades). In some cases, these feature metrics are also analyzed in the context of their location and delineated areas of interest (AOIs). Eye tracking metrics are further used for statistical analysis or alternatively analyzed at the level of several machine learning approaches.

3 Conclusion

The present scoping review was devoted to the current state of knowledge in the area of webcam eye tracking in the field of education and learning. In our study, we found that research in this area is still somewhat in its early stages and a large majority of the research is from a computer science background and focuses mainly on automatized detection systems that can be potentially used for education and learning in various learning environments. Nonetheless, there is a huge opportunity to expand research in the field of education, in terms of proper investigation of educational and learning processes, and both thematic focus and the actual processing of the gaze data (as an example of such an approach see Calabrich et al., 2021).

In the first section of our review, we concentrated on the possibilities and areas of using webcam eye tracking in the field of learning and education. Webcam eye tracking is mainly used in e-learning environments, often in the context of observing the learning process, various aspects of an individual's behavior, or enhancing the quality and functionality of such systems and environments. At the same time, webcam eye tracking measures can be applied to a variety of detection functions (e.g. proctoring).

From the methodological perspective, our results show that the field of webcam eye tracking is developed primarily in the field of computer science in the form of designing detection tools, and only a few studies aimed to experimentally explore the cognitive processes (e.g., reading patterns in neurodivergent students, see Wong et al., 2023). The actual use of webcam eye tracking in educational settings, in the sense of replacing conventional eye tracking, has been rather sporadic. The subsequent manner of working with eye tracking data is then rather extensive, ranging from basic gaze tracking on- and off-screen, to analyzing detailed metrics of saccades and fixations, also at the level of selected areas of interest. The choice of eye tracking metrics was determined by the main objectives of the research concerned.

Nonetheless, this scoping review study provides a summary of current trends in the field of webcam eye tracking in the context of learning and education. Within such a context, the authors of the selected studies primarily focus on the area of learning in online learning environments and student behavior while working with them. Based on this, they also concentrate predominantly on the development of automated detection tools in the field of learner attention or engagement.

The results of our study may offer a new perspective and new challenges for educational researchers considering the use of eye tracking for investigative purposes. At the same time, however, more research is also needed on the quality of webcam measurements, even though the selected studies that focused on comparing commercial eye tracking with webcam eye tracking showed reliable outcomes. In any case, this information offers a direction for further research in this area that may lead to broadening and deepening the research.

Acknowledgements

This study is an outcome of the project "Multimodal learning analytics to study self-regulated learning processes within learning management systems" (21-08218S) funded by the Czech Science Foundation.

References

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, *125*, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023

Alkabbany, I., Ali, A. M., Foreman, Ch., Tretter, T., Hindy, N., & Farag, A. (2023). An experimental platform for real-time students engagement measurements from video in STEM classrooms. *Sensors, 23*(3), 1614. https://doi.org/10.3390/s23031614

Antonietti, A., Colombo, B., & Di Nuzzo, C. (2014). Metacognition in self-regulated multimedia learning: Integrating behavioural, psychophysiological and introspective measures. *Learning, Media and Technology*, *40*(2), 187–209. https://doi.org/10.1080/17439884.2014.933112

Aromataris, E., & Munn, Z. (Eds.). (2020). *JBI Manual for Evidence Synthesis*. JBI. https://doi.org/10.46658/JBIMES-20-01

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, *3*(2), 220–238. https://doi.org/10.18608/jla.2016.32.11

Burton, L., Albert, W., & Flynn, M. (2014). A comparison of the performance of webcam vs. infrared eye tracking technology. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 1437–1441. https://doi.org/10.1177/1541931214581300

Calabrich, S. L., Oppenheim, G. M., & Jones, M. W. (2021a). Episodic memory cues in acquisition of novel visual-phonological associations: A webcam-based eye tracking study. *Proceedings of the Annual Meeting of the Cognitive Science Society, 43*. https://escholarship.org/uc/item/76b3c54t

Calabrich, S. L., Oppenheim, G. M., & Jones, M. W. (2021b). Audiovisual learning in dyslexic and typical adults: Modulating influences of location and context consistency. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.754610

Carneiro, R., Lefrere, P., Steffens, K., & Underwood, J. (2011). *Self-regulated learning in technology enhanced learning environments: A European perspective*. Sense Publishers.

Catrysse, L., Gijbels, D., Donche, V., De Maeyer, S., Lesterhuis, M., & Van den Bossche, P. (2018). How are learning strategies reflected in the eyes? Combining results from self-reports and eye tracking. *British Journal of Educational Psychology*, *88*(1), 118–137. https://doi.org/10.1111/bjep.12181

De la Torre, F., Chu, W.S., Xiong, X., Vicente, F., Ding, X., Cohn, J. F. (2015). *IntraFace*. In IEEE International Conference on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia.

Dilini, N., Senaratne, A., Yasarathna, T., Warnajith, N., & Senevirante, L. (2021). Cheating detection in browser-based online examsthrough eye gaze tracking. In *6th International Conference on Information Technology Research (ICITR)*. https://doi.org/10.1109/ICITR54349.2021.9657277

Dostálová, N., Juhaňák, L., & Plch, L. (2022). A Narrative Review of Eye tracking Research on Self-Regulated Learning from Multimedia Learning Materials in Higher Education. *Studia Paedagogica*, *27*(2), 77–96. https://dx.doi.org/10.5817/SP2022-2-4

Guan, X., Lei, Ch., Huang, Y., Chen, Y., Du, H., Zhang, S., & Feng, X. (2022). An analysis of reading process based on real-time eye tracking data with web-camera—Focus on English reading at higher education level. In *Proceedings of the 4th Workshop on Predicting Performance Based on the Analysis of Reading Behavior, March 21–22, 2022.*

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Hutt, S., Wong, A., Papoutsaki, A., Baker, R. S., Gold, J. I., & Mills, C. (2022). Webcam-based eye tracking to detect mind wandering and comprehension errors. *Behavior Research Methods*, *56*. https://doi.org/10.3758/s13428-022-02040-x

Khan, A. R., Khorsravi, S., Hussain, S., Ghannam, R., Zoha, A., & Imran, M. A. (2022). EXECUTE: Exploring Eye Tracking to Support E-learning. In *Global Engineering Education Conference (EDUCON)*. https://doi.org/10.1109/EDUCON52537.2022.9766506

Koshravi, S., Khan, A. R., Zoha, A., & Ghannam, R. (2022). Self-Directed Learning using Eye-Tracking: A Comparison between Wearable Head-worn and Webcam-based Technologies. In *Global Engineering Education Conference (EDUCON)*. https://doi.org/10.1109/EDUCON52537.2022.9766468

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, Ch. Y., Liu, T. Ch., Lee, S. W. Y., Lee, M. H., Chiou, G. L., Liang, J. Ch., Tsai, & Ch. Ch. (2013). A review of using eye tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Li, J., Ngai, G., Leong, H. V., & Chan, S. C. F. (2016). *Multimodal Human Attention Detection for Reading from Facial Expression, Eye Gaze, and Mouse Dynamics*. *ACM SIGAPP Applied Computing Review*, *16*(3), 37–49. https://doi.org/10.1145/3015297.3015301

Lin, Z., Liu, Y., Wang, H., Liu, Z., Cai, S., Zheng, Z., Zhou, Y., & Zhang, X. (2022). An eye tracker based on webcam and its preliminary application evaluation in Chinese reading tests. *Biomedical Signal Processing and Control*, *74,* 103521. https://doi.org/10.1016/j.bspc.2022.103521

Madsen, J., Júlio, S. U., Gucik, P. J., Steinberg, R., & Parra, L. C. (2021). Synchronized eye movements predict test scores in online video education. *PNAS*, *118*(5). https://doi.org/10.1073/pnas.2016980118

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1), 143. https://doi.org/10.1186/s12874-018-0611-x

Ochoa, X. (2017). *Multimodal Learning Analytics*. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook of Learning Analytics* (pp. 129–141). Society for Learning Analytics Research.

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*(422), 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). WebGazer: Scalable Webcam Eye Tracking Using User Interactions. In *Proceedings of the Twenty-Fifth International Joint Conference on Artifcial Intelligence*, 3839–3845.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Robal, T., Zhao, Y., Lofi, Ch., & Hauff, C. (2018). Webcam-based Attention Tracking in Online Learning: A Feasibility Study. In *IUI '18: 23rd International Conference on Intelligent User Interfaces*. https://doi.org/10.1145/3172944.3172987

Scheiter, K., Schubert, C., & Schüler, A. (2018). Self-regulated learning from illustrated text: Eye movement modelling to support use and regulation of cognitive processes during learning from multimedia. *British Journal of Educational Psychology*, *88*(1), 80–94. https://doi.org/10.1111/bjep.12175

Semmelmann, K., & Weigelt, S. (2017). Online webcam-based eye tracking in cognitive science: A first look. *Behavior Research Methods, 50*(2), 451–465. https://doi.org/10.3758/s13428-017-0913-7

Skovsgaard, H., Agustin, J. S., Johansen, S. A., Hansen, J. P., Tall, M. (2011). Evaluation of a remote webcam-based eye tracker. In *NGCA '11: Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*. https://doi.org/10.1145/1983302.1983309

Šmideková, Z. (2018). Eye tracking v prostredí školskej triedy. Pedagogika, *68*(1), 25–50. https://doi.org/10.14712/23362189.2017.773

Taub, M., & Azevedo, R. (2018). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an intelligent tutoring system? *International Journal of Artificial Intelligence in Education*, *29*, 1–28. https://doi.org/10.1007/s40593-018-0165-4

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., Lewin, S., Godfrey Ch., M., Macdonald, M., T., Langlois, E., V., Soares-Weiser, K., Moriaty, J., Clifford, T., Tunçalp, Ö., & Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Wisiecka, K., Krejtz, I., Sromek, D., Cellary, A., Lewandowska, B., & Duchowski, A. (2022). Comparison of Webcam and Remote Eye Tracking. *ETRA '22*, *32*, 1–7. https://doi.org/10.1145/3517031.3529615

Wong, A. Y., Bryck, R. L., Baker, R. S., Hutt, S., & Mills, C. (2023). Using a Webcam Based Eye-tracker to Understand Students' Thought Patterns and Reading Behaviors in Neurodivergent. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA, ACM, New York, USA*. https://doi.org/10.1145/3576050.3576115

Worsley, M. (2018). Multimodal Learning Analytics' Past, Present, and Potential Futures. In A. Pardo, K. Bartimote, G. Lynch, S. Buckingham Shum, R. Ferguson, A. Merceron, & X. Ochoa (Ed.), *Companion Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK'18)* (pp. 404–419). Society for Learning Analytics Research.

Yi, J., Sheng, B., Shen, R., Lin, W., & Wu, E. (2015). Real Time Learning Evaluation Based on Gaze Tracking. In *14th International Conference on Computer-Aided Design and Computer Graphics*. https://doi.org/10.1109/CADGRAPHICS.2015.13

Zhao, Y., Lofi, Ch., & Hauff, C. (2017). Scalable Mind-Wandering Detection for MOOCs: A Webcam-Based Approach. In *European Conference on Technology Enhanced Learning 2017*. https://doi.org/10.1007/978-3-319-66610-5_24

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

11.2.7 Data extraction. (2022). JBI Manual for Evidence Synthesis. https://jbi-global-wiki.refined.site/space/MANUAL/4687700/11.2.7+Data+extraction

STUDIA PAEDAGOGICA

vol. 28, n. 3, 2023
https://doi.org/10.5817/SP2023-3-6

VOCABULARY LEARNING STRATEGIES, SELF-REGULATED LEARNING, AND LEARNERS' OUTCOMES IN PRIMARY SCHOOL PAIR WORK

Barbora Al Ajeilat Kousalová[a]

[a] Faculty of Education, Masaryk University, Czech Republic

ABSTRACT

This study investigates vocabulary learning strategies (VLS) used among ten primary school learners. Through video recordings, the research explores specific VLS utilized during pair work and their influence on learning outcomes, analyzed with qualitative content analysis. The research questions address the identification and utilization of VLS, the relationship between VLS usage and the ability to infer word meanings, and learner engagement in VLS usage. Findings indicate a notable co-occurrence of some strategies. Moreover, the broader the learners' prior knowledge, the more successful they were with inferring word meanings. The study also emphasizes the need for balanced VLS engagement to optimize outcomes in pair work. This research aims to create new impulses for learning/teaching vocabulary within a foreign language classroom through the targeted practice of vocabulary learning strategies. Such practice aims to facilitate students' learning processes in promoting their self-regulated learning.

KEYWORDS

vocabulary learning; pair work; strategies; foreign language; self-regulated learning; primary school

CORRESPONDING AUTHOR

Barbora Al Ajeilat Kousalová, Faculty of Education, Masaryk University, Poříčí 9, 603 00 Brno, Czech Republic
e-mail: bkousalova@mail.muni.cz

Introduction

Vocabulary is acknowledged as the core of (foreign) language learning, without which successful communication cannot occur (Schmitt, 2010). Moreover, vocabulary acquisition is one of the biggest problems in learning a foreign language (Alqahtani, 2015). Since the 1980s, there has been a growing emphasis on student-centered approaches and their learning processes in learning and teaching (Nunan, 1990). This shift is accompanied by a general interest in learning strategies, particularly within the context of acquiring a second language (Oxford, 2013).

The self-regulated learning (SRL) concept is an umbrella term involving diverse techniques and modalities to foster students' self-directed learning. Models of SRL consistently incorporate learning strategies through explicit or implicit instruction (Oxford, 2013). A learning strategy is an action plan to achieve a learning objective, a technique aiming to facilitate the active learning process (Oxford, 1990). It involves a wide range of approaches for acquiring and applying knowledge and skills to solve problems and achieve success. Vocabulary learning strategies (VLS) put the focus on techniques employed to learn and expand one's vocabulary, specifically facilitating vocabulary learning in foreign languages (Schmitt, 2000). Positioned within the broader context of SRL, VLS can be recognized as a vital component, strategically leveraging the power of SRL through collaborative pair work activities. This deliberate integration fosters optimal vocabulary learning, where the dynamic interplay between SRL and VLS becomes not only more observable but also connects these two concepts over an important social form of learning, in which learners' metacognitive engagement is combined with help-seeking strategies through their peers (Karabenick & Berger, 2013).

1 Vocabulary learning strategies

Vocabulary learning strategies (VLS) can be defined as actions that learners take to (a) determine the meaning of unknown words, (b) retain them in long-term memory, (c) recall them at will, and (d) use them in oral or written mode (Catalán, 2003, p. 56.) In this study, VLS refer to the techniques learners employ to discover the meaning of a new word. VLS are rooted in the theoretical framework of language learning strategies and constitute integral components in their taxonomies. However, most of the VLS taxonomies omit the aspect of discovering the meaning of a new word, concentrating solely on vocabulary learning and retention (e.g., Cohen, 1990; Gu & Johnson, 1996; Rubin & Thompson, 1994; Stoffer, 1995) and have therefore been excluded from the theoretical framework of this study. Schmitt (1997), on the other hand, proposed a comprehensive VLS list based on Oxford's taxonomy (1990), adopting four strategy groups (social, memory, cognitive, metacognitive) and expanding them to include the group of discovery strategies for inferring the meanings of new words. For this reason, Schmitt's taxonomy (1997) was selected as the primary theoretical framework for this study.

Schmitt (1997) categorizes VLS into two main groups: discovery strategies for uncovering the meanings of new words and consolidation strategies for solidifying the meanings of such words. For this study and following the definition of the VLS as stated above, only the first category is described in this section. Discovery strategies can be further divided into determination strategies, which assist a learner in determining a new word's meaning without the help of a qualified person, and social strategies, which involve another person in discovering a new word's meaning.

Table 1
Schmitt's Taxonomy of Vocabulary Learning Strategies: Discovery Strategies (Schmitt, 1997)

Discovery strategies	**Determination**	Analyze part of speech
		Analyze affixes and roots
		Check for L1 cognate
		Analyze any available pictures or gestures
		Guess from textual context
		Bilingual dictionary
		Monolingual dictionary
		Word lists
		Flashcards
	Social	Ask teacher for an L1 translation
		Ask teacher for paraphrase or synonym of new word
		Ask classmates for meaning
		Discover new meaning through group work activity

Discovery strategies aid learners in uncovering the meanings of new words and can be grouped into determination and social strategies. Determination strategies involve analyzing part of speech helping learners identify a word's word class. Examining a word's roots or suffixes can also provide valuable hints regarding its meaning. Another strategy involves checking for L1 cognates, which allows learners to estimate word meanings based on shared origins, such as words derived from the same parent word, e.g., "Mutter" in German and "mother" in English. Visual cues are also helpful; analyzing available pictures or accompanying gestures and intonation in oral discussions can assist learners in guessing meanings. Furthermore, learners can estimate a word's meaning by considering its textual context and cues. To further support vocabulary learning, reference material, including bilingual or monolingual dictionaries, word lists, and flashcards, can be provided to learners.

Social strategies come into play when learners seek assistance from others, such as asking the teacher for a translation into their mother tongue or requesting a paraphrase or a synonym of the new word. Learners can also ask a classmate about a word's meaning or engage in group activities to acquaint themselves with new words collaboratively.[1]

Using strategies shifts the focus from the teacher to the learners and their learning. In this sense, the learner, not the teacher, controls the learning process (Hsu & Malkin, 2011). Strategy use is part of a larger concept called self-regulated learning (SRL), which involves systematically activating behavior, cognition, and motivation toward one's goals (Schunk & Greene, 2017). A student who successfully engages in SRL uses multiple strategies to support their learning, such as seeking assistance or using all available resources (Alvi & Gillies, 2021). SRL within learning analytics (LA) refers to understanding students and their learning in different environments. It is "the measurement, collection, analysis, and reporting of data about learners and their contexts for understanding and optimizing learning and the environments in which it occurs" (Siemens, 2013, p. 1382). Although the scientific research in LA mainly focuses on virtual environments, Long and Siemens (2014) intentionally avoid restricting LA solely to the online education space and digital technologies because of the increasing need to apply LA to face-to-face interactions in physical classrooms. This field of research is called Multimodal Learning Analytics (MLA). It engages different sources of learning data, targeting the understanding of learning and attempting to optimize it without the mediation of digital technology (Ochoa, 2017). In the present study, the VLS are defined as a constituent phase of the SRL structure. By using different modalities, i.e., video and audio recordings, a comprehensive view of the learning processes and actions of learners is provided.

Because the conceptualizations of VLS have been imprecise, and there is no unanimous consensus on the criteria for its definition, it remains undetermined whether they should be classified as observable behaviors, internal mental processes, or a combination of both (Schmitt, 2010). In the past, the assessment of VLS use has primarily relied on self-report questionnaires (e.g., Soureshjani, 2011; Yaacob et al., 2019), since strategic learning is influenced by cognitive processes that are typically not directly observable.

---

1 For a more detailed description of specific strategies see appendix A.

Another approach to investigating the use of VLS was experimental research (e.g., Kaplan-Rakowski, 2019; Maheswari & Sultana, 2019). However, no study has directly observed the use of these strategies in foreign language lessons. In contrast, a qualitative approach is needed to examine the complex strategies and record any relevant learners' behavior contributing to vocabulary learning. The qualitative approach constantly compares and expands existing models with emerging categories from the recordings, focusing on causality (Oxford, 2013). Additionally, observational records afford a higher level of objectivity than questionnaires, in which the learners often provide answers they believe are socially acceptable (Cohen, 2011).

To further enhance the observability of VLS, the learners may engage in pair work activities. Learners working in pairs actively employ metacognitive strategies, allowing them to reflect on and control their learning processes, such as establishing learning goals, connecting new with previous knowledge, gathering and organizing material, monitoring mistakes, or making any required modifications (Oxford, 1990). Additionally, learners are encouraged to ask their peers for clarifications on words or concepts they are unfamiliar with, linking metacognitive and help-seeking strategies, both central concepts of SRL (Karabenick & Berger, 2013).

In conclusion, most VLS studies have been conducted in the quantitative research tradition, lacking lesson observations and relying on students' reports. Furthermore, the target group in most of the research was secondary or university students. However, the mapped research is beneficial in establishing the theoretical-methodological framework for the current research. While the studies mentioned above, which investigated the use of VLS, lacked lesson observations and relied on reported strategies from students, this study's primary objective is to observe the VLS utilization during pair work and their connection to inferring word meanings and the learners' engagement.

2 Methodology

It was initially planned to carry out the pilot study in the spring of 2019, but due to the COVID pandemic and the closing of schools it was postponed until autumn of 2020. The main study was carried out in the spring of 2021. Nevertheless, the schools were open only for a month, so the study's time frame had to be adjusted.

2.1 Sample

The sample consisted of ten primary school learners (n=10) in their ninth (final) grade of primary school. Of the ten learners, seven were female, and three were male. The sample was selected purposely as I was a teacher of this group, which allowed me to grasp the learners' interactions as naturally as possible without disturbing their attention during the data collection. Because of the learners' age, I collected informed consent for recording from the legal representatives, stating that all the data would be anonymous and only the learners' pseudonyms would be used. Moreover, the results would be published only in connection with the study. From a group of twenty learners, ten (and their legal guardians) agreed to be recorded for scientific purposes.

The learners worked in pairs and, if possible, with their preferred choice of partner to ensure a pleasant atmosphere and working environment. A critical factor in the composition of the pairs was that one of the learners had taken part in the pilot study and so had previous experience with a similar task and could provide the other learner with an explanation of the working procedure and steps needed to be taken to perform the assigned task. Table 1 informs about the composition of the pairs according to their pseudonyms (only the beginning letter of their given name was preserved), age[2], grade[3], and participation in the pilot study. Every pair was assigned a working number, later used in the result section for clarity.

Table 2
Description of the study sample

Pseudonym	**Age**	**Grade**	**Pilot study participation**	**Assigned number**
František	15	2	Yes[4]	1
Kryštof	15	2	No	1
Viktorie	15	1	Yes	2
Kateřina	15	1	No	2
Andrea	15	1	Yes	3
Vlasta	14	1	No	3
Erika	15	1	Yes	4
Jaromír	15	2	No	4
Tamara	14	1	Yes	5
Lenka	14	1	No	5

2 Age of a learner on the day of data collection for the main study

3 Grade from the German language course from the first term of the school year 2020/2021

4 The learner took part in executing the task itself, however he was not recorded and therefore not included in the pilot study sample. He fulfilled the assigned task with the other included pairs.

The learners learned German for their obligatory second foreign language instruction, which in 2013 was made part of the primary school curriculum in the Czech Republic (MŠMT, 2017). The selected primary school does not offer a choice of languages, and German is the obligatory second foreign language (L3). German lessons take place twice a week in a forty-five-minute session. I have chosen ninth-grade learners because they are studying German in their third year and can use German vocabulary at a basic level. By the end of the year, they achieve an A1 level as defined by the Common European Framework of Reference for Languages (2012). The learners' interactions were implemented in the Czech language; all excerpts included here were translated into English by the author of this study.

2.2 Research aims and questions

The aim is to determine the specific VLS employed by Czech primary school learners during pair work and to examine how the learners implement the identified VLS. Another aim is to investigate whether applying these VLS contributes to the learners' ability to infer the meanings of new words. Finally, the last aim is to explore the extent of engagement in utilizing VLS during pair work. To address these aims, the following research questions were formulated:

1 Which VLS are employed during pair work?
2 How are the identified VLS used during pair work?
3 Do the employed VLS lead to inferring the meanings of words?
4 How are the learners engaged in the VLS usage?

2.3 Research design

To address the research questions, I employed a qualitative study approach. The qualitative approach addresses the gap identified in previous research that lacked a qualitative perspective of VLS used by learners in the foreign language classroom. The previous research relied mainly on reported strategies, and lacked the quality of observing the learner's behavior directly in the lessons. Learners were divided into pairs to make the behavior more observable and allow the linkage between strategies used and self-regulated learning (see literature review). I utilized audio and video recordings and analyzed the data using the qualitative content analysis method (Mayring, 2015), which allowed me to study the learners' related behavior and actions when using the VLS.

2.4 Data collection

The data collection occurred in five consecutive lessons over three weeks in June of 2021. The data was collected from indirect observations based on video and audio recordings as the research instrument (Janík et al., 2013).

While the video recordings enabled me to focus on the participants' verbal and non-verbal expressions, the audio recordings supplemented the video's inaudible tracks. The subjects of investigation was the exploration of the meanings of new German vocabulary.

Figure 1 depicts the location of cameras in the classroom. Pairs participating in the study were situated at the back of the classroom, whereas regular German lessons for other learners took place in the front. The individual pairs were separated from each other with room dividers so they would not interfere with each other during their interactions. One camera was focused on each pair, and a recorder was placed on each desk, recording the sound, which was inaudible on the video recording.

Before the data collection, a category system[5] based on Schmitt's taxonomy (1997) was developed. Schmitt (1997) established the taxonomy of VLS based on Oxford (1990), extending it with determination strategies that support the students in uncovering a new or unknown word's meaning without a qualified person's help. The category system was created to recognize the learners' first VLS use and to determine the teaching aids to be provided while working on the task.

The teaching aids for the learner strategy elicitations were (1) the text "The Timid Rabbit" (Shaw, 2015), (2) a list of content vocabulary, and (3) pictures available in the text. Content vocabulary consists of nouns, adjectives, verbs, adverbs, and pronouns and is essential for understanding a written text (Roche, 2005). Concerning the text (Shaw, 2015), a vocabulary analysis was first conducted to confirm its A1 level (Glaboniat, 2005). The optimal learning level should not exceed one level higher than the level learners have currently attained, i.e., A2 level (Hufeisen & Riemer, 2010). In view of this recommendation, any B1+ vocabulary was replaced with another word at a lower level[6]. Another vocabulary learning suggestion is to encounter a maximum of twelve new words in one lesson (Gairns & Redman, 1986). Words contained in the textbook (Friedericke et al., 2007) with which the learners worked in regular German lessons were considered known. In contrast, the new words were those that learners had not encountered in the textbook. The teaching aids were available for each pair.

---

5 The category system is attached in the appendix A.

6 E.g., the word "sich wälzen" was replaced by "sich rollen."

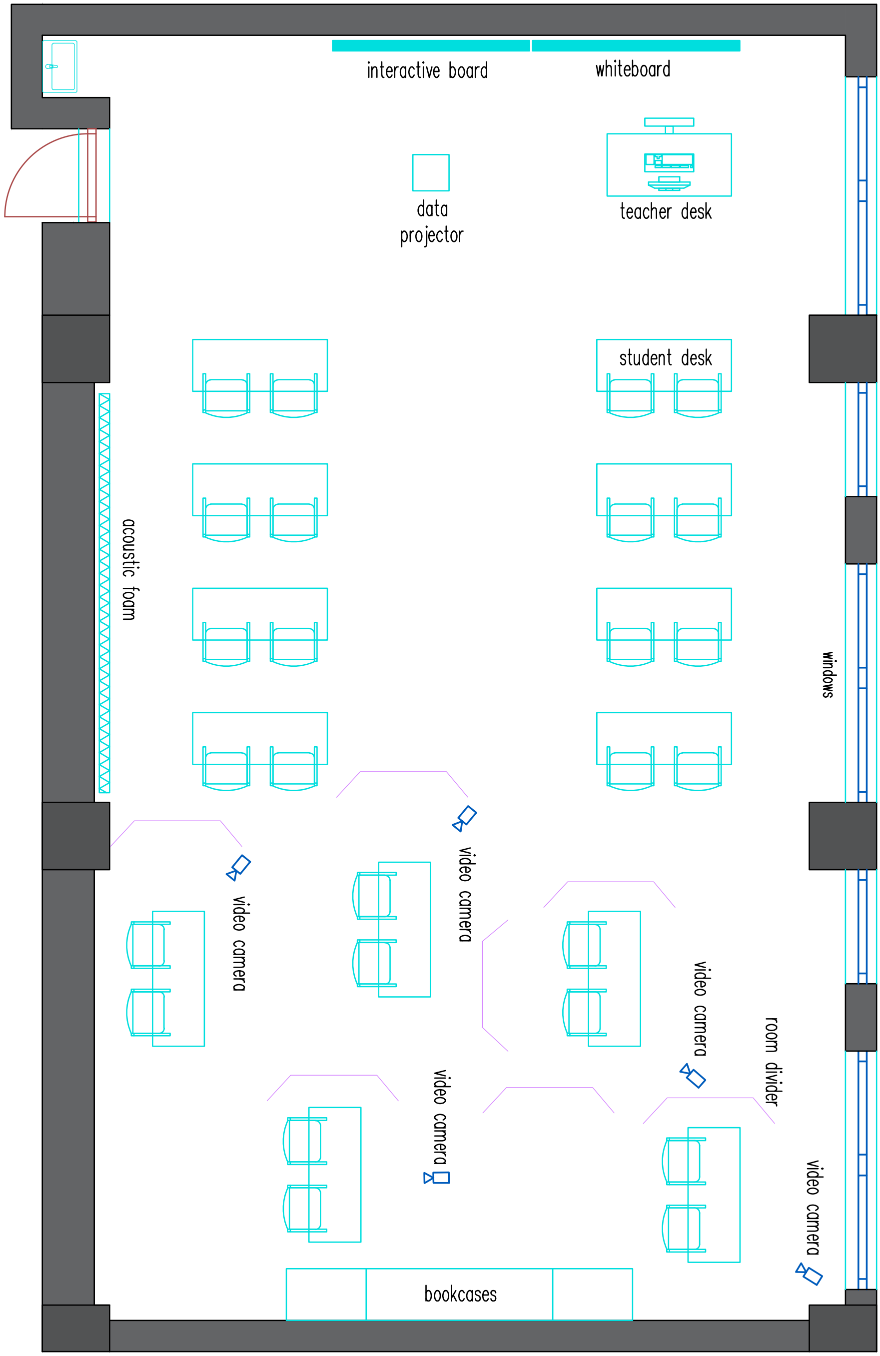


Figure 1
Floor plan and camera placement

2.5 Data analysis

Qualitative content analysis was selected, and the units of analysis were segments in which the learners dealt with word meanings. The analysis took place deductively, i.e., according to the category system, and inductively, i.e., other strategies were derived from the data (Schreier, 2014). The analysis was carried out following the steps reported by Kuckartz (2018).

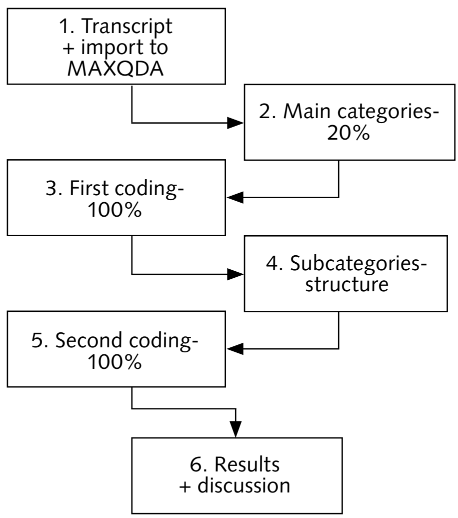


Figure 2
Steps of analysis

The first step was transcribing verbal and non-verbal data and transferring them to the MaxQDA software (VERBI Software, 2019), in which the text transcript was synchronized with the video recordings. The verbal data was transcribed according to Kaderka and Svobodová (2006), and the nonverbal utterances according to Silverman (2011). The second step consisted of determining the main categories, which corresponded with the three categories from Schmitt's taxonomy (1997) – determination, social, and metacognitive strategies. In this step, 20% of data were coded with these categories. The third step of the analysis was to establish the segments, where each segment began with the learner or both learners starting to deal with a word's meaning and finished with them moving on to another word. Subsequently, the above-mentioned main categories were assigned to the individual segments and the entire data corpus was coded. The next step was to create a subcategory structure with all potential subcategories. Initially, the subcategories were identified based on the category system and then generated from the data. Step five involved coding the whole dataset with the main categories and subcategories. The last step of the analysis is addressed in the corresponding chapters of this article.

3 Results

This chapter presents the results connected to the formulated research questions from the methodology section. The results are structured according to each pair working in pair work and the number of lessons in which the pair was recorded for better clarity.

3.1 Which VLS are employed during pair work?

The specific VLS used by the learners in pair work within the five recorded lessons are structured in Table 3 below according to the categories from the developed category system. The VLS strategies are divided into three categories according to whether the strategy was used only by one of the learners in a pair to estimate a new word's meaning: determination strategies (DET), or if the usage of a strategy involved another person, whether it was the learner with whom the person worked in a pair, or the teacher, or someone from another pair participating on the research: social strategies (SOC). According to the literature (Schmitt, 1997; Oxford, 1990), metacognitive strategies (MET) were the last category group. These strategies usually did not lead to inferring the meaning of a new word without connection to other strategies but were inductively produced from the data and, therefore, considered helpful in establishing a new word's meaning.

The numbers show the usage of a strategy during a particular recorded lesson, where empty fields mean that the pair did not use the strategy. As seen in Table 3, some pairs used a wide range of VLS (e.g., pairs 2 and 3). On the other hand, some pairs used a limited number of VLS repeatedly (e.g., pair 1, pair 5), and there is also one pair (pair 4) who, in the last recorded lesson, did not use any VLS to solve the given task.[7]

Table 3 shows that the most frequently used determination strategies across the pairs were *guessing from textual context* and *from available pictures*. From the social strategies, they *asked classmates for meaning* and *made sure about meaning,* and from the metacognitive strategies, it was *linking with already known material*. Details regarding the usage of specific VLS are discussed in the following section.

3.2 How are the identified strategies used during pair work?

Analyzing parts of speech within a text did not often reveal a word's meaning. Instead, it helped the learners understand the word's significance, deciding whether to explore its meaning further or skip it based on its perceived

[7] The distribution of VLS is discussed in section *3.2*.

Table 3

VLS usage count by pair and recorded lesson

Vocabulary learning strategy type	**Pair 1**					**Pair 2**					**Pair 3**					**Pair 4**					**Pair 5**				
	Lesson No.					Lesson No.					Lesson No.					Lesson No.					Lesson No.				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
DET1: Analyze part of speech	1					1	1	3		1	5	1	3	2	3		1				3	1		1	1
DET2: Analyze affixes and roots											2			1										2	
DET3: Guess from textual context	11	16	19	8	13	18	22	23	20	16	9	10	18	18	15	6	12				18	11	25	10	23
DET4: Analyze any available pictures	5	7	5	9	8	11	9	7	18	4	8	6	8	8	7	2	9	2	5		16	4	6	11	7
DET5: Bilingual dictionary														19											
DET6: Word lists						3	1	1			16	8	8	8	2	2	1								
DET7: Spelling												1													
DET8: Splitting word in parts: composites						1				1	1	1													
DET9: Sound associations from L1							4		1	2	1	2		1	1						2			1	
DET10: Sound associations from L2								1		3		1	1								1				
DET11: Sound associations from L3+	2					1						3	3	1	2	3		1				2			
SOC1: Ask classmates for meaning	2	3	1		1	14	6	16	8	8	8	5	8	2		3	10				1	3		3	1
SOC2: Ask classmates for meaning (other pair)						1				1	1														
SOC3: Association						1		1																	
SOC4: Copying (from other pairs)																2									
SOC5: Making sure about a meaning	2		5	1		2	6	6		6	4	11	4	5	6	4	8	3			5	8	8		1
SOC6: Ask teacher for an L1 translation						1	1	4	2	6	8	10		1	2	7	5		3						
MET1: Skip or pass new word							1					1									2				
MET2: Linking with already known	8	11	22	8	9	17	25	12	12	16	24	29	16	13	27	14	19	2	2		27	25	26	9	26
MET4: Self-correction: pictures/ textual context															2						1		2		

importance. Throughout the five recorded lessons, all pairs utilized this strategy, with pair 3 employing it most frequently: *V: Gibt could be conjugated from geben*. This approach aligns with a study by Sukanya and Nutprapha (2017) highlighting the importance of understanding words based on their parts of speech in educational news articles. The study recommends analyzing the parts of speech in any text to create practical teaching resources, mainly focusing on high-frequency parts of speech, such as nouns, enabling learners to expand their vocabulary with commonly used words.

Analyzing affixes and roots represents a strategy notably beneficial for learning the German language. This approach allows learners to separate the components of a word, the root from its prefixes, to estimate its meaning. In the following excerpt, pair 5 engages in determining the meaning of the word "weglaufen," which translates to "run away" in English: *T: Wait, der Fuchs läuft, that means that the fox ran (…). She ran away, isn't it? L: Yes, somehow away.* A study by Iseni and Rexhepi (2023) on Germanic prefixes emphasizes their vital role in word formation, altering the base word's meaning. The authors underscore how this knowledge empowers learners to navigate the complexities of Germanic languages, significantly enhancing their comprehension abilities. Additionally, the specific success of pair 5 in separating the root "läuft" from the prefix "weg" and determining the word's meaning strongly supports the effectiveness of this strategy in practice.

Guessing a word's meaning from textual context emerges as one of the most commonly employed strategies. Pair 2 notably excelled in this strategy, frequently integrating the guessed word's meaning into the sentence to assess its contextual coherence. For instance, *V: Well, kleinen, which means small*, in this excerpt the learners incorporated the word *kleinen* into a sentence o*nce upon a time, there was a small, timid rabbit* to infer its meaning from the context. Bai (2018) classifies this strategy as one of the guessing strategies for estimating word meanings, emphasizing its effectiveness. Supporting this, Rahmani's study (2023), focusing on using VLS of Afghan EFL learners, found that most participants (85.82%) relied on context-based guesswork to comprehend new words. Their approach involved leveraging logical development, common sense, and knowledge to infer word meanings during text reading.

Analyzing any available pictures was a strategy based on using the enclosed visuals in the text, which offered learners additional support in comprehending the material. Some pairs relied heavily on the literal meaning of pictures to estimate the meaning of certain words. For instance, in the case of "Dunkelheit," several pairs directly associated the word with a picture where a rabbit was hidden beneath a blanket. This overly literal approach is evident in the following excerpt from pair 5: *L: I would say this is something like a blanket*, or pair 2: *V: That's paying attention under the blanket?* Vivaldi and Allen (2021) examined children's understanding of pictures, discovering that the interpretation of a picture, whether literal or nonliteral, hinged upon various contextual cues. As depicted in the provided excerpts, pairs 2 and 5 failed to consider the contextual aspects while estimating the word "Dunkelheit". Their literal interpretation of the picture content hindered them from inferring its intended meaning.

A *bilingual dictionary* was introduced during the fourth recorded lesson, in which the learners encountered over twelve new words (Gairns & Redman, 1986). Pair 3 exclusively decided to use this reference material, distinguishing themselves from the other pairs with different strategies. The approach from pair 3 involved confirming previously guessed word meanings by referring to the dictionary for verification, as seen in the excerpt: *V: Beule, I suppose that (nn) we guessed correctly (+ is looking in the dictionary) (...) Be-, Beu-, Beu- (16) Beu, bulge, nice.* Bai (2018) emphasizes that learners utilize dictionaries to understand word meanings and confirm their knowledge for accurate usage. This aligns with Pair 3's practice of confirming their guessed meaning of "Beule" as "bulge," confirming the guessed meaning using the bilingual dictionary.

Spelling emerged as an exclusive strategy employed by Pair 3 during their investigation of the compound word "Angsthase," which translates as "timid rabbit" in English. Their approach involved splitting the word into two components, "Angst" and "Hase," making sure about the word's spelling to avoid mistaking it with another word: *V: wait, so A-N-G-S-T A: [Here] V: And there is Hase A: [She said] m- V: So it's Hase and Angst. And one of the words means dark, and the other one is hair.* Plonsky (2011) investigated the practices of successful language learners, discovering that they consciously focus on spelling and form when learning new words. The strategy of *splitting words into parts* was derived from the data, initially considered part of *analyzing affixes and roots*. However, it was later recognized that compounds cannot be strictly categorized as having affixes, thus creating a new strategy category. Hubáčková (2015) conducted a study on German compounds in which she stated that it is almost impossible to guess the meaning of a compound based on its components only. As seen in the excerpt, the pair refers to "Angst" as "dark," which suggests a previous encounter with this word, in which the pair estimated the meaning as stated above.

The last three identified determination strategies were linked to *sound associations*, exclusively used by three pairs. These strategies involved seeking resemblances in sound between the new word and words from the learners' native language (Czech), first foreign language (English), or other foreign languages (German, Russian, French, etc.). However, in the excerpts provided, none of the learners successfully estimated the word's meaning, resulting in interference rather than aiding comprehension. *Sound associations from L1* were most frequently used by pair 2, with an example such as "Fuchs" being compared to the Czech word "fuška," interpreted as "hard work." *V: Der Fuchs, like fuška, that something is hard.* Similarly, *sound associations from L2* did not assist pair 3 in estimating the correct word meaning, as seen in the *excerpt: A: Frei, so frei (+ reads from the word list), those are French fries.* This group's third and last strategy was the *sound association from L3*, used most frequently by pair 3. In the following excerpt, they grapple with the meaning of the word "mutig" incorrectly as "Mutter" because of its sound similarity: *V: Mut, man (…) that's something like A: it reminds me of Mutter, that's mom…* The correct meaning of the word was "brave" in English. De Bruin et al. (2023) confirm that cross-language intrusions between L1, L2, and L3 can disrupt the language learning process, aligning with the observed interferences caused by sound association strategies.

Among the varied strategies employed by learners, *asking classmates for meaning* was one of the most frequently used strategies from the group of social strategies. This strategy is commonly adopted when encountering unfamiliar words, requiring learners to seek clarification from peers or teachers. Instances exemplify the use of this strategy, such as when learner *J* inquires about the word "einfach": *What does einfach mean (+ reads from the word list)*, or when learner *E* asks a learner from another pair for the meaning of "fürchtest": *E: Do you know, V., what fürchtest means?* In specific scenarios, this approach was found inadequate, prompting students from pairs 2, 3 and 4 to opt for teacher assistance, as demonstrated in the following excerpt, when learner *V* raises a question about the words "Dunkelheit" and "gespannt": *V: Miss teacher, we have a question (+ is raising hand). We don't know what Dunkelheit means and gespannt. I thought that one might be fever or cold, but (…).* Drawing on Vygotky's (2012) theory, the positive impact of the social environment, peers, and teachers on the learning process is emphasized. Learners actively engage with peers to explore and elicit word meanings, which enhances their ability to infer meanings that might elude them when working independently. Evidently, the strategy's effectiveness in *asking classmates for meaning* depends on the learner's existing knowledge and/or their capacity to infer meaning from *textual context* or *pictures*. On the other hand, *asking the teacher for meaning* consistently leads to an estimation of the word's meaning, whether through providing direct translation in L1 or indirect cues from text and visuals.

Expanding on the previous strategy of *asking for meaning*, another category of strategies, referred to as *making sure about meaning,* was identified from the data. This strategy involved a learner proposing a potential meaning of a word and seeking approval or confirmation from their peer, integrating the learner's existing knowledge into the discussion. An illustration of this strategy is evident in this excerpt: *V: Gute means good, right?* Here, the learner presents their understanding of the word "gute" and seeks confirmation from their partner. Ipek (2009) highlights the significance of approval or praise to reinforce a student's activity, motivating them in their subsequent work.

In the context of language learning, seeking confirmation about the meaning of a word from a peer not only validates one's understanding but also creates a collaborative environment that encourages active participation and reinforces the learning process.

One intriguing strategy identified from the data was *association*, which emerged in the interaction of pair 2. Instead of directly *asking classmates for the meaning* of a new word, one learner prompted the other to draw connections between the new word and their existing knowledge or experiences. This instance is depicted when learner *K* asked, *What does fürchten remind you of?* And answering their question: *Absolutely nothing.* Following this, learner *V* attempted to encourage associative thinking by suggesting: *But maybe (…)*, hinting at a potential association. Drawing from the insights of Manzo and Manzo (1990), the *association* strategy aligns with the subjective approach to vocabulary (SAV). This approach encourages students to draw upon their experiences or associations to complement dictionary definitions of new terms. It focuses on building connections between existing knowledge and new vocabulary, facilitating a more profound and personal understanding of the words encountered.

In the first recorded lesson, a notable strategy emerged utilized by pair 4 as they encountered challenges in advancing through the task. This particular strategy involved what could be identified as *copying from other pairs*, a strategy they resorted to when facing difficulties. This approach became apparent in the following dialogue: *E: Why don't we listen to others? J: That could work*. This exchange highlights their decision to seek information from other pairs, particularly in the case of two words, indicating their reliance on the knowledge of others to infer meanings for the given words. According to a study on English education in larger class settings by Erlina et al. (2022), referring to or replicating others' work is described as a coping mechanism in response to time constraints for completing tasks. It acknowledges the pressures of limited time and indicates that the final product may not solely reflect the individual learners' knowledge.

In utilizing metacognitive strategies, pairs 2, 3 and 5 employed *skipping* or *passing a new word*. This strategy is a response to encountering a word that is challenging for a pair to comprehend, acknowledging the time and effort necessary to understand the word's meaning. This is exemplified in the excerpt from pair 5: *T: I would skip this. We will come back to it later.* This excerpt showcases the decision of learner *T* to skip a problematic word initially, aiming to return to it later. Their action aligns with findings by Aravind and Rajasekaran (2018), indicating that skipping unknown words in the learning process is a time-saving strategy. The research also indicates that many learners tend not to revisit the skipped words due to a lack of persistence in estimating their meanings. However, the instance breaks the trend by the pair returning to word "beißen" and successfully inferring its meaning as "to bite": *L: beißen will be to bite.* This instance stands out as the pair demonstrated persistence by returning to the skipped word, successfully estimating its meaning. Contrary to the norm observed in the study, this pair's perseverance led to accurate comprehension. Their persistence illustrates a determination to comprehend and reflects a thorough approach to inferring the meanings of all the text's words.

The most used metacognitive strategy was *linking with already known material*, which involved associating the meaning of a word with the learner's existing knowledge base. Pair 5 notably exhibited the highest frequency of employing this strategy, showcasing their extensive prior knowledge. This is exemplified in the following excerpt: *T: Klein, which means small,* and *L: Grandmother, Oma.* These instances demonstrate their immediate recognition and accurate estimations of word meanings, indicating their strong association between known words and their meanings. The lack of hesitation in their statements indicates a confident and direct link to their existing knowledge. The successful and confident estimations of word meanings by pair 5 and their high engagement indicate a positive impact of prior knowledge on learning. The interaction showed that their broad prior knowledge enabled swift and accurate connections between known and new words, leading to confident estimations. This aligns with the findings of Dong et al. (2020), which suggest that prior knowledge positively influences learning engagement. It allows students to expand their working memory, facilitating the acquisition of new knowledge and enhancing overall learning and engagement.

The final metacognitive strategy observed in the data was *self-correction* involving textual context and/or pictures. Pair 5 utilized this strategy to rectify previous estimations that did not align with the textual context. Their correction was notably based on their interpretations of the enclosed pictures and the text. This is exemplified in the exchange of pair 5: *T: We put that down, but probably wrong as hide, here, to hide under the blanket, but he doesn't hide in the water, right? L: Well, in that case, Angst. T: That looks like being scared again. L: Well, so this will be to be afraid, fürchtet.* The learners' correction was influenced not only by the text but also by the visual cues in the illustrations, which depicted a rabbit initially under a blanket in a bed and later in front of a lake, exhibiting signs of being notably scared in both scenarios. The learners' use of textual context and pictures for self-correction emphasizes their conscious effort to correct their earlier estimations that did not align with the context provided. Swain (2005) indicated that self-correction requires learners to recognize their errors consciously, and this observation supports the idea that learners can notice and correct their own mistakes. McCormick and Vercellotti (2013) further affirm that learners can self-correct without specific training, mainly when not preoccupied with formulating meaning.

This section delved into the utilization of VLS, primarily focusing on how learners employed specific strategies, and the attempt to interpret the acquired results with the existing literature on the given patterns. Descriptions of these strategies reveal a consistent trend: many of them were not used in isolation, operating independently, but rather in connection with other strategies. Besides the various strategies used interdependently, only three strategies were used in isolation. One such strategy was *asking teacher for meaning*, a standalone approach not combined with other strategies. However, preceding their request to the teacher, learners consistently attempted to estimate a word's meaning, employing various strategies independently, but often perceived these efforts as unsuccessful. Another strategy used independently was the *bilingual dictionary*. Specifically, pair 3 was the soul group utilizing a dictionary to explore new word meanings. Despite its solitary use without combination with other strategies, the dictionary was typically used to confirm the previously estimated meanings by engaging different strategies. The final strategy used individually was *copying from other pairs*. In this case, pair 4 mutually sought assistance from other pairs (without their knowledge) when faced with challenges in estimating word meanings.

In contrast, the remaining strategies showed a notable tendency to co-occur. Learners frequently relied on elicitation materials, such as pictures, text, and word lists, as their primary resources to estimate word meanings. This approach involved using multiple strategies simultaneously, systematically reflecting their procedures in their learning process. They also frequently revisited already estimated meanings and words still in the estimation process, constantly reevaluating and refining their understanding. The only exception to this pattern was observed in pair 4, as they chose not to review or revise their estimated meanings. Their approach prioritized completing the task as quickly as possible, but this came at the expense of the VLS usage, the accuracy of their estimations, and finally, their engagement in the strategy usage. The following section further investigates the process of inferring the word meanings.

3.3 Do the employed VLS lead to inferring the meanings of words?

The investigation into the relationship between VLS usage and the successful determination of word meanings directs our attention to the specific number of VLS employed by individual pairs. Thus, five figures will be presented in the upcoming section to address the third research question. In these figures, the x-axis denotes the recorded lesson numbers (1, 2, 3, 4, 5), while the y-axis illustrates the number of VLS used. The dark grey gridline represents strategies that successfully facilitated word meaning inference, while the light grey gridline indicates strategies with which the learners failed to do so, resulting in unsuccessful inference.

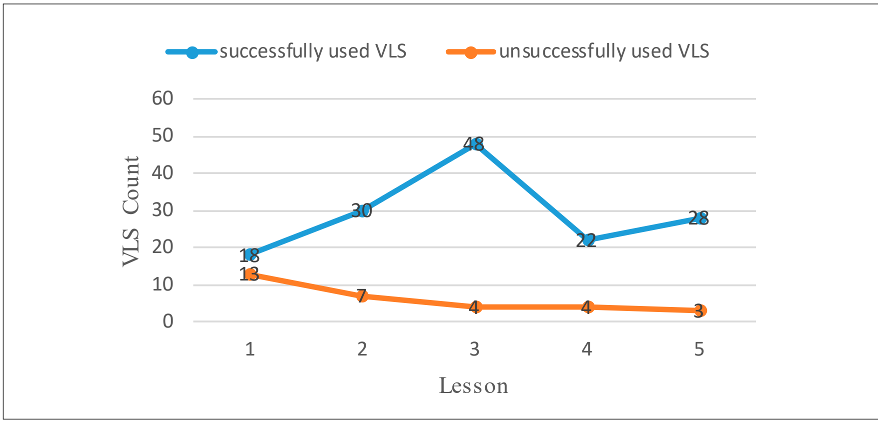


Figure 3
Pair 1: VLS strategy count per lesson for word meaning inference

The data in Figure 3 illustrates the VLS utilized by pair 1 across five recorded lessons. Pair 1 successfully inferred meanings in between 58.06% and 92.31% of instances. Pair 1 used five different strategies: *Guessing from textual context* emerged as the most prevalent and successful strategy, utilized in 37.33% of instances to infer word meanings. Following closely was *linking with already known material*, employed in 36.67% of cases. *Analyzing available pictures* was the third most frequently used strategy, accounting for 18% of instances. However, using the last two strategies, *making sure about meaning* and *asking classmates for meaning*, was almost negligible, at 4.67% and 0.67%, respectively.

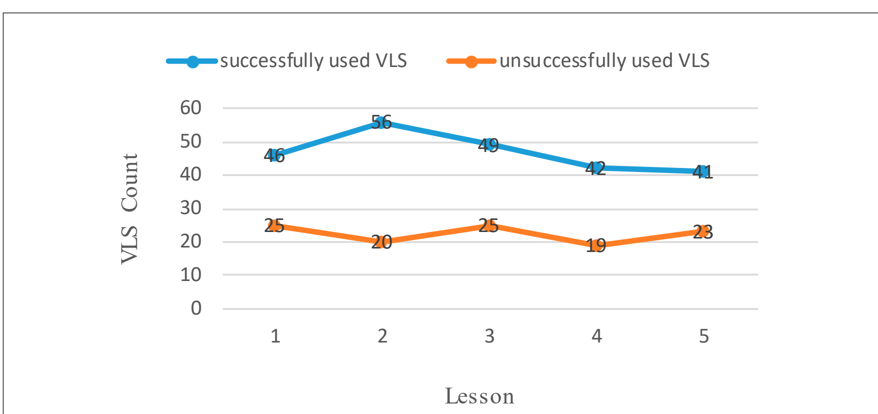


Figure 4
Pair 2: VLS strategy count per lesson for word meaning inference

The data presented in Figure 4 showcases how pair 2 utilized VLS. The success rates across the lessons were notably stable, consistently maintaining a relatively steady level of success, between 64.06% and 73.68%. The pair consistently favored *linking with already known material.* This strategy was prominently employed in 32.48% of instances across all lessons, demonstrating its recurrent significance for inferring word meanings. Throughout the five lessons, pair 2 used a total number of ten different strategies. The second most often used strategy, *guessing from textual context,* was used in 30.34% of instances. The third most often used strategy was *asking classmates for meaning* in 14.96% of instances. *Analyzing available pictures* followed at 14.96% and *asking teacher for meaning* at 5.98%. Additionally, *making sure about meaning* was used in 5.56% of instances. Other strategies were used in less than 1% of cases.

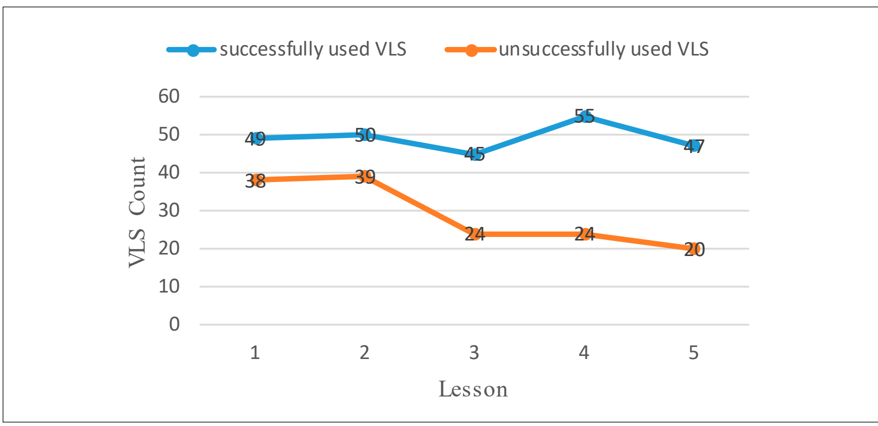


Figure 5
Pair 3: VLS strategy count per lesson for word meaning inference

Figure 5 displays the usage of VLS by pair 3. This pair showcased a stable success rate, fluctuating from 56.32% to 70.15%. Pair 3 employed a total of 15 strategies successfully for inferring word meanings. Notably, the most frequently used strategy was *linking with already known material*, utilized in 36.99% of cases, followed by *guessing from textual context* at 15.85%. *Word lists* as reference material were the third most commonly used strategy at 8.94%. Other strategies, such as using a *bilingual dictionary*, *asking teacher for meaning*, *analyzing any available pictures,* or *making sure about meaning*, were used with percentages lower than 8%.

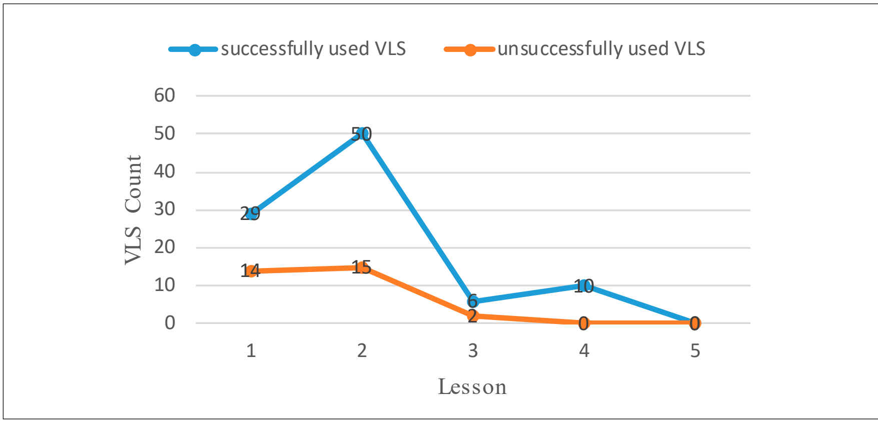


Figure 6
Pair 4: VLS strategy count per lesson for word meaning inference

Data in Figure 6 depicts the VLS usage from pair 4. This pair achieved high success rates, ranging from 67.44% in the first lesson to 100% in the fourth lesson. However, during the fourth lesson, they only used VLS ten times. In the last lesson, the pair chose to skip the process of guessing word meanings entirely. Throughout the five lessons, nine strategies were successfully used to infer meanings. *Linking with already known material* was the most frequently used strategy at 36.84%, followed by *asking teacher for meaning* at 15.79%. *Guessing from textual context* and *analyzing available pictures* were used with identical percentages, at 13.68%. *Making sure about meaning* was also notable at 11.58%. Other strategies were used in less than 5% of cases.

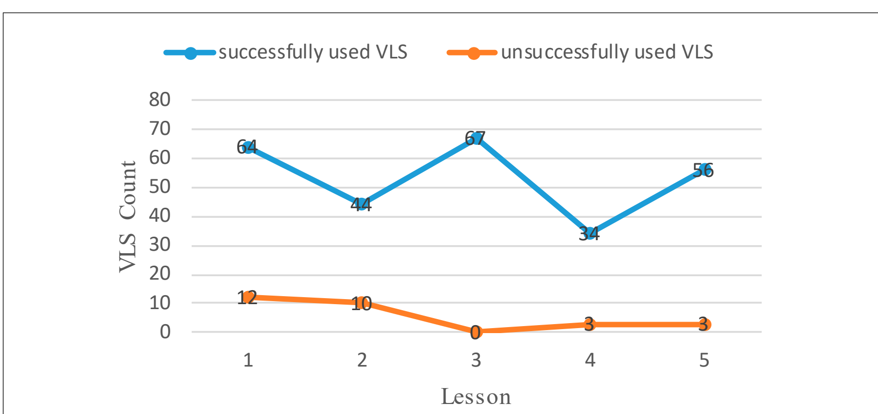


Figure 7
Pair 5: VLS strategy counts per lesson for word meaning inference

Figure 7 presents the VLS use of pair 5. This pair demonstrated the highest success rates among all pairs, ranging from 81.48% in the second lesson to 100% in the third lesson, where all 67 times the strategies used led to inferring word meanings. They employed 13 strategies, *linking with already known material* being the most frequently used at 41.89%, followed by *guessing from textual context* at 30.57% and *analyzing any available pictures* at 15.09%. *Making sure about meaning* was used in 6.79% of cases. The rest of the strategies were used in fewer than 2% of cases.

The analysis revealed distinctive patterns in the usage of VLS and their success in determining word meanings. The strategy most frequently used by many pairs was *linking with already known material*, demonstrating its recurrent significance for inferring word meanings. Some stability in success rates was noted, with pair 2 showcasing the most consistent success rates between lessons. Pair 5 reached all the pairs' highest possible success rates. These observations highlight the impact of strategies like *linking with already known material* and *guessing from textual context*, emphasizing their repeated use and success in understanding word meanings among the diverse pairs. The findings imply that learners with a broader foundation of prior knowledge tended to achieve higher success rates in inferring new word meanings.

3.4 How are the learners engaged in the VLS usage?

The investigation into pair engagement focused on the distribution of strategies employed within the pairs. This analysis distinguished between individual and combined strategy usage. Individual strategy employment refers to instances where only one learner from a pair used a strategy without seeking input from the other. In contrast, combined usage occurred when both learners used strategies to uncover word meanings, negotiate, or agree/disagree on estimated meanings. Moreover, within the combined strategy utilization, the initiator of the strategy use was identified to clarify engagement distribution within each pair. Figure 8 illustrates the degree of engagement by displaying the number of shared (K+F, K+V, A+V, E+J, L+T) and individual uses of VLS. The second part of the figure presents the initiation processes within shared VLS usage, depicted in percentages.

		Pair														
		1			2			3			4			5		
	Engagement	K+F	K	F	K+V	K	V	A+V	A	V	E+J	E	J	L+T	L	T
Lesson	**1**	28	3		61	4	6	49	4	34	32	4	7	57	4	15
	2	35		2	58	4	14	70	5	15	52	6	7	41	11	2
	3	52			69	0	5	64	3	2	4	4	0	61	0	6
	4	26			47	5	9	68	0	11	0	0	10	32	1	2
	5	27			56	2	6	48	0	19	0	0	0	54	0	6
	Initiaion		K	F		K	V		A	V		E	J		L	T
	1		68%	32%		28%	72%		18%	82%		53%	47%		47%	53%
	2		60%	40%		33%	67%		19%	81%		15%	85%		49%	51%
	3		50%	50%		33%	67%		81%	19%		50%	50%		51%	49%
	4		69%	31%		64%	36%		15%	85%		0%	0%		59%	41%
	5		81%	19%		55%	45%		4%	96%		0%	0%		55%	45%

Figure 8
VLS engagement level and initiation in pair-work

Pair 1 primarily favored a collaborative approach, utilizing combined VLS to estimate new word meanings. *Krystof* often initiated strategies, prompting his partner for meanings, while *Frantisek* primarily responded. Their work pattern strongly preferred joint engagement, with *Krystof* initiating strategies more actively. Pair 2 exhibited varying individual VLS usage. *Katerina* and *Viktorie* demonstrated different levels of individual engagement, with *Viktorie* leading in strategy initiation. This imbalance suggests a need for a more balanced contribution from both learners for enhanced outcomes. Pair 3 showed high individual VLS usage, with *Vlasta* significantly dominating in initiating strategy use. This dynamic suggests the potential impact on their collaborative work. Initially collaborating, Pair 4's cooperation declined in later sessions, with *Jaromir* solely employing VLS and disregarding his partner's contributions. There was a shift from collaborative work to independent strategy usage. Pair 5 consistently engaged in shared VLS usage. *Tamara* took the lead in strategy initiation, but *Lenka* actively challenged or disputed her partner's estimations, contributing to their collaborative approach.

The chapter investigates the engagement and strategy usage within pairs. It assesses the level of engagement by analyzing the distribution of strategies employed, distinguishing between individual and shared usage. Pairs 1 and 5 predominantly showed a collaborative approach in VLS, focusing on joint engagement and negotiation of word meanings. In contrast, Pairs 2, 3, and 4 displayed varying levels of individual strategy engagement, implying potential differences in the collaborative dynamics. Pair 3 demonstrated significant individual VLS usage, indicating a dominant force in strategy activation. Pair 2 exhibited more varied individual engagement, suggesting a potential need for balanced contribution to enhance their joint outcomes. These observations emphasize distinct dynamics within each pair, influencing their approaches to collaborative learning.

4 Discussion

Utilizing synchronized video and audio recordings allowed for a comprehensive exploration of how new word meanings were determined in German. This multimodal approach proved to be a foundation for in-depth revisiting and reassessing strategies and cooperation dynamics. As Chan et al. (2020, p. 20) referenced, employing MLA techniques, particularly the combination of video and audio data, enhances reliability and consistency in coding. Incorporating various modalities and extracting diverse features can offer deeper insights into higher-level constructs such as engagement, pair-work dynamics, and self-regulation.

The study explored and categorized VLS utilized during pair work, uncovering diverse patterns in their utilization, including their interdependent manner. It was observed that most strategies were seldom used in isolation, with learners often combining multiple strategies to estimate new word meanings. This observation aligns with findings from Nie and Zhou's study (2017), in which proficient English learners employed a multitude of VLS in combination, rather than isolated, to achieve successful learning outcomes. The study highlighted the effectiveness of employing various strategies collectively, reinforcing the idea that a combined approach enhances learning efficacy.

Data analysis revealed a distinction between successful and unsuccessful strategy applications. Successful strategies led to correctly determining word meanings, whereas unsuccessful strategies resulted in incorrect or undetermined word meanings. The study's outcomes indicate the recurring success of strategies across diverse pairs in inferring word meanings, such as *linking with already known material* and *guessing from textual context*. This observation resonates with O'Malley and Chamot's (1990) concept, suggesting that a strategy, when repeatedly successful, may evolve into an automatic and procedural approach. This transformation likely occurs due to the consolidation of successful and unsuccessful conditions associated with the strategy. Hence, throughout consistent successful practice, learners instinctively employ these strategies when encountering similar learning conditions.

The observations highlight the varying levels of engagement and strategy usage among pairs, emphasizing their different approaches to collaborative learning. Pairs 1 and 5 predominantly displayed joint engagement, while pairs 2, 3 and 4 exhibited diverse levels of individual strategy engagement, hinting at potential discrepancies in pair-work dynamics. For instance, pair 4 took a negative approach, potentially hurting their ability to proceed with the work during the subsequent lessons, giving up on the shared VLS use and meaning negotiations. Chan et al. (2020) emphasize the impact of individual behaviors on pair-work engagement and dynamics. The significant individual usage of VLS across the pairs indicates the need for balanced contributions to optimize shared outcomes. Pair-specific differences in strategy usage and engagement levels may impact their effectiveness in inferring word meanings and overall success in pair work.

4.1 Limitation

The presence of cameras, as an invasive data collection tool, can influence learners' behavior, and the data can be significantly distorted (Laurier & Philo, 2012). Nevertheless, the research was implemented over three consecutive weeks, during which the learners gradually stopped noticing the cameras and started to behave more naturally.

The scope of the study is focused on what is observable in the classroom, omitting the cognitive aspects and out-of-class events. However, Oxford (2017) suggests that the connection between learning strategies and self-regulation involves both sociocultural and psychological dimensions. This implies that the process of strategy use extends beyond what is observable in the classroom. The study's focus on the observable social process within the classroom may limit the exploration of cognitive events, potentially neglecting insights into the broader context. On the other hand, Shum and Ferguson (2012) propose that a deeper insight into the learning process is acquired by observing essential aspects of learning, such as interaction, cooperation, or group processes.

Another limitation of the study is that I work in the group as their teacher and simultaneously as the researcher, which can affect the objectivity and distort the data. However, the essence of the study is to investigate the learners' VLS usage and their connection to inferring word meanings and engagement levels. These findings then provide the learners an insight into their conscious and unconscious learning habits (Juhaňák & Zounek, 2016). The study findings are beneficial in understanding the learners' practices when encountering new vocabulary, especially for me as the group's teacher. This knowledge is helpful for future lesson planning aimed at vocabulary instruction. One possible way to maintain impartiality would be to involve a second researcher to analyze a specific data set and test the inter-coder reliability (Kuckartz, 2018).

Conclusion

The study comprehensively explored determining new word meanings in German in primary school, and utilizing synchronized video and audio recording provided insights into vocabulary learning processes during pair work. The multimodal approach facilitated a deeper understanding of the strategies and dynamics of cooperation within pairs. The findings are supported by previous studies (Nie & Zhou, 2017), emphasizing the effectiveness of employing multiple strategies collectively to achieve successful learning outcomes. Successful strategies, such as *linking with already known material* and *guessing from textual context*, repeatedly led to accurate word meaning inferences across diverse pairs, aligning with the concept that recurrently successful strategies may become automatic over time (O'Malley & Chamot, 1990). The observations revealed varied engagement and strategy usage among pairs, highlighting potential disparities in pair-work dynamics and underscoring the need for balanced contributions to optimize shared outcomes.

The study primarily focused on understanding practices at a micro level, specifically examining a group of learners, aiming to improve teaching practice without generalizations (Juhaňák & Zounek, 2016). The findings provide valuable insights into individual and collective learning processes among learners in a classroom setting. Understanding how learners interact with each other and approach unfamiliar words can significantly contribute to comprehending the overall dynamics in pair work and gaining deeper insight into the needs of learners. This comprehensive understanding could then facilitate more effective planning, task allocation, and assessment of vocabulary-related activities. It creates an environment where tailored strategies can be implemented, optimizing the educational experience for all students. Furthermore, the learners' awareness about their learning practices connects the VLS usage with the very nature of self-regulated learning (Redmer, 2022). Nevertheless, further research is needed to gather data on extensive reflections from the learners on the practices they engage in.

References

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education, 3*(3), 21–34. http://doi.org/10.20472/TE.2015.3.3.002

Alvi, E., & Gillies, R. (2021). Self-Regulated Learning (SRL) perspectives and strategies of Australian primary school students: a qualitative exploration at different year level. *Educational Review, 75*(4), 680–702. https://doi.org/10.1080/00131911.2021.1948390

Aravind, B.R., & Rajasekaran, V. (2018). A study on vocabulary learning strategies of research scholars. *The International Journal of Research in Teacher Education*, *9*(2), 16–25. http://ijrte.inased.org/makale/493

Bai, Z. (2018). An analysis of English vocabulary learning strategies. *Journal of Language Teaching and Research*, *9*(4), 849–855. https://doi.org/10.17507/jltr.0904.24

Chan, M.C.E., Ochoa, X., Clarke, D. (2020). Multimodal learning analytics in a laboratory classroom. In: Virvou, M., Alepis, E., Tsihrintzis, G., Jain, L. (Eds.) Machine learning paradigms. *Intelligent Systems Reference Library*, *158*. Springer. https://doi.org/10.1007/978-3-030-13743-4_8

Common European Framework of Reference for Languages. (2012). https://rm.coe.int/16802fc1bf

Cohen, A. (1990). *Language Learning: Insight for Learners, Teachers and Researchers*. Heinle & Heinle.

Cohen, A. (2011). *Strategies in Learning and Using a Second Language*. Pearson.

De Bruin, A., Hoversten, L. J., & Martin, C. D. (2023). Interference between non-native languages during trilingual language production. *Journal of Memory and Language*, *128*, 1–17. https://doi.org/10.1016/j.jml.2022.104386

Dong, A., Jong, M., & King, R.B. (2020). How does prior knowledge influence learning engagement? The mediating roles of cognitive load and help-seeking. *Frontiers in Psychology*, *11*, 1–10. http://doi.org/10.3389/fpsyg.2020.591203

Erlina, D., Marzulina, L., Harto, K., Holandyah, M., Wulandari B.A., Fauzan, M., Fridiyanto, F., & Mukminin, A. (2022). English education and large classes: unpacking the challenges and coping strategies. *Theory and Practice in Language Studies*, *12*(3), 489–497. https://doi.org/10.17507/tpls.1203.08

Friederike, J., Rohrmann, L., & Zbranková, M. (2007). *Prima A1/díl 1 učebnice*. Fraus.

Gairns, R., & Redman, S. (1986). *Success in English Teaching*. Oxford University Press.

Glaboniat, M. (2005). *Profile deutsch: gememeinsamer ëurpäischen Referenzrahmen: Lernzielbestimmungen, Kannbeschreibungen, kommunikative Mittel: Niveau A1-A2, B1-B2, C1-C2*. Langenscheidt.

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, *46*(4), 643–679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

Hsu, A., & Malkin, F. (2011). Shifting the focus from teaching to learning: rethinking the role of the teacher educator. *Contemporary Issues in Education Research, 4*(12), 43–50. https://doi.org/10.19030/cier.v4i12.6661

Hubáčková, S. (2015). german noun compounds at seminar lessons. *Procedia – Social and Behavioral Sciences*, *186*, 502–506. https://doi.org/10.1016/j.sbspro.2015.04.033

Hufeisen, B., & Riemer, C. (2010). Spracherwerb und Sprachenlernen. In H.-J. Krumm, C. Fandrych, B., Hufeisen. & C. Riemer (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft* (pp. 738–753). de Gruyter.

Ipek, H. (2009). Comparing and contrasting first and second language acquisition: implications for language teachers. *English Language Teaching. 2*(2), 155–163. http://doi.org/10.5539/elt.v2n2p155

Iseni, A., & Rexhepi, A. (2023). Prefixes of Germanic origin. *Anglisticum Journal*, *12*(1), 40–48. http://doi.org/10.58885/ijllis.v12i1.40.ai

Janík, T., Minaříková, E., & Najvar, P. (2013) Der Einsatz von Videotechnik in der Lehrerbildung: Eine Übersicht leitender Ansätze. In U. Riegel & K. Macha (Eds.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (pp. 63–78). Waxmann.

Jimenez-Catalán, R. M. (2003). Sex Differences in L2 Vocabulary Learning Strategies. *International Journal of Applied Linguistics*, *13*(1), 54–77. http://doi.org/10.1111/1473-4192.00037

Juhaňák, L., & Zounek, J. (2016). Analytika učení: nový přístup ke zkoumání učení (nejen) ve virtuálním prostředí. *Pedagogická orientace*, *26*(3), 560–583. https://doi.org/10.5817/PedOr2016-3-560

Kaderka, P. & Svobodová, Z. (2006). Jak přepisovat audiovizuální záznam rozhovoru? Manuál pro přepisovatele televizních diskuzních pořadů. In *Jazykovědné aktuality 3–4*, 18–51.

Kaplan-Rakowski, R. (2019). The effect of stereoscopic three-dimensional images on vocabulary learning. *Contemporary Educational Technology*, *10*(4) 324–337. https://doi.org/10.30935/cet.634172

Karabenick, S. A. & Berger, J. (2013). Help seeking as a self-regulated learning strategy. In H. Bembenutty, T. J. Cleary & A. Kitsantas (Eds.), *Applications of Self-regulated Learning Across Diverse Disciplines. A tribute to Barry J. Zimmerman* (pp. 237–261). IAP.

Kuckartz, U. (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung.* Beltz Juventa.

Laurier, E., & Philo, Ch. (2009). Natural problems of naturalistic video data. In H. Knoblauch, B. Schnettler, J. Raab, Soeffner, H.-J. (Eds.), *Video Analysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology* (pp. 181–190). Lang.

Long, P., & Siemens, G. (2014). Penetrating the fog: analytics in learning and education. *TF63 – Dossier – Learning Analytics 22*(3), 132–137. https://doi.org/10.17471/2499-4324/195

Maheswari, U. V. & Sultana, A. S., (2019). VLS model for content words: An experimental study on engineering students in India. *GEMA Online Journal of Language Studies*, *19*(1), 62–76. https://doi.org/10.17576/gema-2019-1901-04

Manzo, A.V., & Manzo, U.C. (1990). *Content Area Reading: A Heuristic Approach.* Merrill.

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken.* Beltz.

McCormick, D. E., & Vercellotti, M. L. (2013). Examining the impact of self-correction notes on grammatical accuracy in speaking. *TESOL Quarterly, 47*(2), 410–420. https://doi.org/10.1002/tesq.92

Ministerstvo školství, mládeže a tělovýchovy. (2017). Rámcový vzdělávací program pro základní vzdělávání. http://www.nuv.cz/uploads/RVP_ZV_2017.pdf

Nie, Y., & Zhou, L. (2017). A study of vocabulary learning strategies used by excellent English learners. *Research on Modern Higher Education*, *4*, 101–106.

Nunan, D. (1990). *Language Teaching and Methodology.* Prentice Hall.

O'Malley, J.M., & Chamot, A.U. (1990). *Learning strategies in second language acquisition.* Cambridge University Press. https://doi.org/10.1017/CBO9781139524490

Oxford, R.L. (1990). *Language Learning Strategies: What Every Teacher Should Know.* Heinle & Heinle. https://doi.org/10.5070/L411004984

Oxford, R.L. (2013). *Teaching and Researching Language Learning Strategies.* Routledge.

Oxford, R.L. (2017). *Teaching and Researching Language Learning Strategies: Self Regulation in Context.* Routledge.

Ochoa, X. (2017). Multimodal Learning Analytics. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook of Learning Analytics* (pp. 129–141). Society for Learning Analytics Research.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: a meta analysis. *Language Learning. 61*(4), 993–1038. https://doi.org/10.1111/j.1467-9922.2011.00663.x

Rahmani, S. (2023). Vocabulary learning beliefs and strategies of Afghan EFL undergraduate learners. *Cogent Education*, *10*(1), 1–22. https://doi.org/10.1080/2331186X.2023.2194227

Redmer, G. (2022). Self-regulation in an advanced language learner: A case study of language learning strategies. *Studies in Self-Access Learning Journal, 13*(1), 60–76. https://doi.org/10.37237/130104

Roche, J. (2005). *Fremdsprachenerwerb – Fremdsprachendidaktik*. A. Francke.

Rubin, J., & Thompson, I. (1994). *How to be a More Successful Language Learner: Toward Learner Autonomy*. Heinle & Heinle.

Schmitt, N. (1997). Vocabulary Learning Strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary Description, Acquisition and Pedagogy* (pp. 199–227). Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge University Press.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.

Schreier, M. (2014). Qualitative Content Analysis. In U. Flick (Ed.), *The SAGE Handbook of Qualitative Data Analysis* (pp. 170–183). SAGE.

Schunk, D. H., & Greene, J. A. (2017). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of Self-regulation of Learning and Performance* (pp. 10–37). Routledge.

Shaw, E. (2015). *Der kleine Angsthase*. Beltz.

Shum, S. B., & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3–26. https://doi.org/10.1145/2330601.2330616

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, *57*(10), 1380–1400. https://doi.org/10.1177/0002764213498851

Silverman, D. (2011): *Interpreting Qualitative Data: A Guide to the Principles of Qualitative Research*. SAGE.

Soureshjani, K. H. (2011). Gender-oriented use of vocabulary strategies: A comparative study. *Theory and Practice in Language Studies*, *1*(7), 898–902. https://doi.org/10.4304/tpls.1.7.898-902

Stoffer, I. (1995). *University Foreign Language Students' Choice of Vocabulary Learning Strategies as Related to Individual Difference Variables*. The University of Alabama.

Sukanya, K., & Nutprapha, K. D. (2017). A study of parts of speech used in online student weekly magazine. *International Journal of Research – Granthaalayah*, *5*(4), 43–50. https://doi.org/10.5281/zenodo.569978

Swain, M. (2005). The Output Hypothesis: Theory and Research. In E. Hinkel (Ed.), *Handbook on Research in Second Language Teaching and Learning* (pp. 471–484). Routledge.

VERBI Software. *MAXQDA 2020* [computer software]. 2020. VERBI Software. Available at:<maxqda.com.>.

Vivaldi, R. A., & Allen, M. L. (2021). Beyond literal depiction: Children's flexible understanding of pictures. *Journal of Experimental Child Psychology*, *210*, https://doi.org/10.1016/j.jecp.2021.105208

Vygotsky, L. (2012). *Thought and Language: Revised and Expanded Edition*. The MIT Press.

Yaacob, A., Shapii, A., Saad Alobaisy, A., Al-Rahmi, W. M., Al-Dheleai, Y. M., Yahaya, N., & Alamri, M. M. (2019). Vocabulary learning strategies through secondary students at Saudi school in Malaysia. *SAGE open*, *9*(1), 11–12. https://doi.org/10.1177/2158244019835935

Appendix A: Category System

Main Category	Category Type	Subcategory	Description	Examples
Determination	A	Analyze part of speech (DET1)	Learner analyzes or identifies a new word's word class.	*V: Gibt could be conjugated from geben.*
		Analyze affixes and roots (DET2)	Learner examines a new word's root, suffixes, or affixes.	*T: Wait, der Fuchs läuft, that means that the fox ran (…). She ran away, isn't it? L: Yes, somehow away (.).*
		Guess from textual context (DET3)	Learner estimates the meaning of a new word using a textual context by inserting words into sentences an/or deriving a word's meaning from its surrounding words.	*V: Well, kleinen, which means small, once upon a time, there was a small, timid rabbit*
		Analyze any available pictures (DET4)	Learner assumes a new word's meaning from the pictures incorporated in the text.	*L: I would say this is something like a blanket*
		Bilingual dictionary (DET5)	Learner uses a bilingual dictionary to estimate a word's meaning.	*V: Beule, I suppose, that (nn) we guessed correctly (+ is looking in the dictionary) (..) Be-, Beu-, Beu- (16) Beu, bulge, nice.*
		Word lists (DET6)	Learner uses a word list with content words from the text to guess a word's meaning.	*J: What does einfach mean (+ reads from the word list)*
	B	Spelling (DET7)	Learner spells the new word and subsequently attempts to estimate its meaning.	*V: wait, so A-N-G-S-T A: [Here]*
		Splitting words in parts: compounds (DET8)	Learner splits a new word into separate parts and attempts to estimate their meaning.	*V: And there is Hase A: [She said] m- V: So it's Hase and Angst.*
		Sound associations from L1 (DET9)	Learner estimates the meaning of a new word according to its sound similarities to the mother tongue (Czech).	*V: Der Fuchs, like fuška, that something is hard.*
		Sound associations from L2 (DET10)	Learner estimates the meaning of a new word according to its sound similarities to the first foreign language (English).	*A: Frei, so frei (+ reads from the word list), those are fries*
	C	Sound associations from L3+ (DET11)	Learner estimates the meaning of a new word according to its sound similarities to the second and other foreign languages (German, Russian…).	*V: Mut, man (.) that's something like A: it reminds me of Mutter, that's mom…*

	A	Ask classmates for meaning (SOC1)	Learner asks their partner from their pair for a meaning of a new word.	*J: What does einfach mean (+ reads from the word list)*
	B	Ask classmates for meaning-other pair (SOC 2)	Learner asks another learner from a different pair for a meaning of a new word.	*E: Do you know, V., what does fürchtest mean?*
		Ask classmates for association (SOC3)	Learners asks their partner what does the new word reminds them of.	*K: What does fürchten remind you of?*
Social	C	Copying from other pairs (SOC4)	Learners from a pair decide to listen to other pairs and copy their estimated meanings.	*E: Won't we listen to others?* *J: That could work.*
		Making sure about meaning (SOC5)	Learner asks their partner about the accuracy of their estimated meaning.	*V: Gute means good, right?*
	A	Ask teacher for meaning (SOC6)	Learner asks the teacher for a meaning of a new word.	*V: Miss teacher, we have a question (+ is raising hand). We don't know what Dunkelheit means and gespannt. I thought that one might be fever or cold, but (…)*
		Skip or pass new word (MET1)	Learner skips the new word.	*T: I would skip this. We will come back to it later.*
Metacognitive	B	Linking with already known (MET2)	Learner activates their previous knowledge aquired either thrugh in-school or out-of-school exposure to language and associates it with a specific word.	*T: Klein, which means small, and L: Grandmother, Oma.*
	C	Self-correction-pictures/textual context (MET3)	Learner corrects their original new word's meaning estimation based on pictures.	*T: We put that down, but probably wrong as hide, here, to hide under the blanket, but he doesn't hide in the water, right?*

Category Type:

A Categories based on Schmitt's taxonomy (1997)

B Categories based on the subcategories structure (step 4 of analysis)

C Catogories established from data (data-driven)

STUDIA PAEDAGOGICA

Studia paedagogica is a peer reviewed journal published by the Faculty of Arts at Masaryk University. The editorial board is located at the Department of Educational Sciences. The publication frequency is three issues per year.

Studia paedagogica publishes papers on education, upbringing and learning from all spheres of social life. The papers are theoretical, but mainly empirical. The journal is interdisciplinary – it covers current topics in educational research while at the same time providing scope for studies grounded in other social sciences.

The journal is indexed in international databases (CEJSH, DOAJ, EBSCO, ERA, ERIH PLUS, PKP, ProQuest, SCOPUS).

Publisher
Masaryk University, Czech Republic
This issue is published in December 2023.

Print
Reprocentrum, a. s., Blansko

Circulation
250 copies

Journal website
www.studiapaedagogica.cz

Contact
Department of Educational Sciences
Faculty of Arts, Masaryk University
Arna Nováka 1, 602 00 Brno
Czech Republic
e-mail: studiapaedagogica@phil.muni.cz

The journal is a continuation of SPFFBU, U series.
ISSN 1803-7437 (print)
ISSN 2336-4521 (online)

STUDIA PAEDAGOGICA
VOLUME 28 / NUMBER 3 / YEAR 2023

Contents

Department of Educational Sciences
Faculty of Arts, Masaryk University
www.studiapaedagogica.cz
ISSN 1803-7437

MUNI PRESS

MUNI ARTS

9 771803 743005 10