Does Mouse Click Frequency Predict Students' Flow Experience?

Pedro Kenzo Muramatsu
University of São Paulo
kenzo_muramatsu@usp.br

Wilk Oliveira
Tampere University
wilk.oliveira@tuni.fi

Kiemute Oyibo
York University
koyibo@yorku.ca

Juho Hamari
Tampere University
juho.hamari@tuni.fi

Abstract

Designing educational systems able to lead students into flow experience is a contemporary challenge, especially given the positive relationship between flow experience and learning. However, an important challenge within the field of learning analytics is evaluating the students' flow experience during the use of educational systems. In general, such evaluation is conducted using invasive methods (e.g., electroencephalogram, and eye trackers) and cannot be massively applied. To face this challenge, following the trend of utilizing behavioral data produced by users to identify their experience when using different types of systems, in our study, we evaluated the applicability of employing one single type of behavior data (i.e., mouse click frequency) as an exclusive metric to model and to predict students' flow experience. By conducting two data-driven studies (N1 = 25 | N2 = 101), we identified that the mouse click frequency on its own is not able to predict the flow experience. Our study contributes to the field of learning analytics confirming that it is not possible to predict students' flow experience only with mouse click frequency and paving the way for new studies that use different behavior data to predict students' flow experience.

1. Introduction

Flow, as first introduced by Csikszentmihalyi (1975), can be defined as "an optimal experience during an activity in which an individual is deeply engaged and presents high levels of fulfillment, focus, and enjoyment" (Csikszentmihalyi, 1975; Csikszentmihalyi and Csikszentmihalyi, 1992). Especially over the last 30 years, a plethora of studies have identified a positive influence of flow experience on students' learning outcomes in educational systems (*e.g.*, students who engage in a high flow experience state are more prone to achieve a satisfactory learning experience (Özhan and Kocadere, 2020) or that flow experience has a positive effect in both comprehension and memorization (Erhel and Jamet, 2019). In summary, literature demonstrates that the flow experience is linked to positive affects on students' learning experience (Oliveira et al., 2018; Oliveira, Pastushenko, et al., 2021; Perttula et al., 2017).

Thus, over the last few years, several researchers have invested in developing different types of educational systems (*e.g.*, self-regulation systems (Wan et al., 2020), games (Chou et al., 2021), and gamified systems (Oliveira, Toda, Toledo, et al., 2020)) to lead students to a flow experience. This is not an easy task due to two different reasons, first, to understand which design aspects lead students to a flow experience (Gao et al., 2019), and second, to analyze whether the system was able to provide a flow experience to students (Semerci and Goularas, 2021). In particular, the field of learning analytics has focused on the challenge of analyzing whether a student has achieved a flow experience (van Schaik et al., 2012).

Different studies have highlighted that reaching an effective measuring technique for flow experience in educational systems is a significant challenge (Hamari and Koivisto, 2014; Jackson and Marsh, 1996; Lee et al., 2014). As pointed out by Oliveira, Toda, Palomino, Rodrigues, Shi, et al. (2020) and Z. Zheng et al. (2019), this challenge is especially pertinent because instruments adopted in past studies (*e.g.*, electroencephalography (EEG), interviews, questionnaires, eye trackers) present major drawbacks: *i)* high cost and *ii)* the impossibility of conducting a massive application (Oliveira, Toda, Palomino, Rodrigues, Shi, et al., 2020; Z. Zheng et al., 2019). Thus, researchers in the area of learning analytics have sought to create alternative methods to analyze students' flow experience (De Kock, 2014; Lee et al., 2014; Semerci and Goularas, 2021).

In face of such limitations and aiming to create these alternative methods, the use of students' behavior data (*i.e.*, user' data logs) to measure flow experience in educational systems has recently gained traction as a scalable and low-cost strategy (Lee et al., 2014; Oliveira et al., 2019; Oliveira, Toda, Palomino, Rodrigues, Shi,

URI: https://hdl.handle.net/10125/102788
978-0-9981331-6-4

et al., 2020; Semerci and Goularas, 2021). Despite recent promising results (De Kock, 2014; Lee et al., 2014; Oliveira et al., 2018; Oliveira et al., 2019; Semerci and Goularas, 2021), this area of study is yet to be comprehensively explored, highlighting the urgent need for empirical data-driven studies investigating the relationship between behavior data and flow experience in educational systems. In particular, it is necessary to investigate how different behavior data can be used to model and predict students' flow experiences.

As explained by Pentel (2015), mouse dynamics can be used in conjunction as features in training data sets (Chen et al., 2017; Guo and Agichtein, 2010; Pentel, 2015; N. Zheng et al., 2011). Among those metrics, the mouse click frequency has particularly shown favorable results in identifying certain experiences (Oliveira, Isotani, et al., 2021; Semerci and Goularas, 2021). At the same time, previous studies on the use of behavior data generated by the mouse presented significant results in the prediction of the users' flow-related experience (*e.g.*, using mouse logs to predict the user's confusion (Pentel, 2015) and satisfaction (Chen et al., 2017)), which demonstrates its potential as a valid metric for different users' experiences (*e.g.*, flow experience).

Facing the challenge of flow experience measurement, in this paper we present the results of two data-driven studies (N1 = 25 | N2 = 101) investigating if the mouse click frequency (during the system usage) can be used independently to predict students' flow experience in educational systems. Thus, we aim to answer the following research question: **Does mouse click frequency predict students' flow experience in educational systems?** In the two studies conducted to answer this question, students used different educational systems (while the mouse click frequency was collected) and then answered a scale to identify their flow experience during the system usage.

Our main results indicate that there is a slight relationship (β) between the mouse click frequency and the students' flow experience. However, such a relationship fails to provide significant predictive power (Adjusted R^2), implying that the mouse click frequency on its own cannot be used as a metric for flow experience prediction. Our study contributes to the field of learning analytics, as it presents empirical evidence that employing the frequency of the mouse as a solo metric is inadequate to predict students' flow experience.

2. Background

In this section, we introduce the main topic addressed in this paper (*i.e.*, flow experience measurement in educational systems), as well as the main related works.

2.1. Flow experience and learning

The flow experience (*i.e.*, "an optimal experience that people have as a motivating factor in their daily activities" (Faiola et al., 2013)) is one of the most important experiences investigated in the field of educational technologies (Oliveira et al., 2018; Oliveira, Pastushenko, et al., 2021; Perttula et al., 2017; Semerci and Goularas, 2021). This is because the flow experience is considered directly related to the learning experience, and when a student is experiencing the flow, they tend to have a high learning experience (Özhan and Kocadere, 2020; Yen and Lin, 2020).

The flow experience is composed of nine different psychological dimensions (also independent experiences) (Csikszentmihalyi, 1997b; Csikszentmihalyi and Csikszentmihalyi, 1992; Jackson et al., 2011): *i)* challenge-skill balance; *ii)* action-awareness merging; *iii)* clear goals; *iv)* unambiguous feedback; *v)* total concentration on the task at hand; *vi)* sense of control; *vii)* loss of self-consciousness; *viii)* transformation of time; and *ix) autotelic* experience. Researchers argue that an individual must reach the nine dimensions to reach the flow experience (see (Csikszentmihalyi, 1997b; Csikszentmihalyi and Csikszentmihalyi, 1992; Jackson et al., 2011) for a thorough review regarding the flow experience dimensions).

The complexity of modeling students' flow experience in an educational system is a pertinent problem (Lee et al., 2014; Oliveira, Isotani, et al., 2021; Oliveira et al., 2019; Semerci and Goularas, 2021). Despite advancements, this technology still has issues, such as application complexity, prompting other ways, such as the use of EEG or eye trackers, to be proposed more recently. All of these approaches, however, have one of three issues: they are either expensive, intrusive, or cannot be used on a large scale.

As a result, more promising solutions include the development of approaches for analyzing flow experience using behavior data generated by users in educational systems (Lee et al., 2014; Oliveira et al., 2019; Oliveira, Toda, Palomino, Rodrigues, Shi, et al., 2020; Semerci and Goularas, 2021). Researchers have used techniques ranging from step regression (Lee et al., 2014) to deep neural networks (Semerci and Goularas, 2021), and structural equation modeling (Oliveira, Isotani, et al., 2021) in an attempt to find behavioral data that are able to predict the flow experience.

2.2. Related Works

Aiming to identify the main related works and provide a deep field understanding, we started analyzing the results of three systematic literature reviews conducted by Oliveira et al. (2018), Oliveira, Pastushenko, et al. (2021), and Perttula et al. (2017), which were conducted to describe the state of the art in Flow Theory and educational technologies (including the most used methods for identifying the students' flow experience in educational systems). From the studies identified in the literature reviews, we also found other external studies.

Lee et al. (2014) were one of the pioneer studies dedicated to investigating the use of behavior data to measure flow experiences in an e-learning environment. Their work presented an automated detector able to distinguish whether a student is experiencing flow, boredom, and frustration. The approach proposed by Lee et al. (2014) was based on a step regression, which utilized the data sourced from college students when interacting with a step-based tutoring system. Their findings support the use of affect detectors on flow conditions (San Pedro et al., 2013). The average mouse click duration, for instance, was found to be positively correlated with boredom but negatively correlated to flow. This also suggests that standard input devices used commonly by students, such as the mouse, are a potential source of abundant information for developing low-cost detectors. The incorporation of automatic flow detection tools is an improvement to ITSs as this feature can be used to predict students' behavior and tailor effective ways to improve their educational performance as well as learning experience (Egbert, 2004).

In the same realm, Oliveira et al. (2019) introduced a theory-driven based conceptual model, which associates each of nine the flow experience dimensions (Csikszentmihalyi, 1997b; Csikszentmihalyi and Csikszentmihalyi, 1992; Jackson et al., 2011) with student interaction data logs. In a further study, Oliveira, Toda, Palomino, Rodrigues, Shi, et al. (2020) proposed the use of the think-aloud protocol (*i.e.*, a method that provides rich verbal data on reasoning during a problem-solving task (Fonteyn et al., 1993)) as a tool to link users' data logs with their flow experiment in an educational system. In yet another study, they set forth the usage of structural equation modeling to model flow in students performing tasks in a gamified system (Oliveira, Isotani, et al., 2021).

Semerci and Goularas (2021) presented a solution that measures the flow state of students by employing heat-maps and deep neural networks (*i.e.*, convolutional autoencoders). This flow theory-based method provides information on the students' flow state in an e-learning platform, which is calculated by also taking into account their grades. Besides collecting interaction data from e-learning platform (*i.e.*, mouse click coordinates, the locations visited with the mouse, access times, etc.), Semerci and Goularas (2021)' method also sourced data from post-lecture quiz results and student surveys. Their findings indicate that while other methods did badly in terms of *p-values* and *r-values*, deep learning methods were able to extract student behavioral patterns based on mouse interactions with significantly better performance.

In addition to the aforementioned approaches, Z. Zheng et al. (2019) proposed a hierarchical recognition model to track the flow experience of computer programmers by tracing their computer interactions (*e.g.*, keyboard, mouse, IDE functions, and switching applications windows) when engaged in software development. Z. Zheng et al. (2019) implemented a non-invasive flow state tracking system whose performance was assessed using a real-world data set from a medium-sized IT company in China. Their results reached the highest recognition accuracy of 92.6%, an achievement that is suited for performing real-time recognition.

In summary, previous studies have investigated the use of behavioral data to analyze students' flow experience. However, as far as we know, no data-driven study has yet aimed to model and predict flow experience in educational systems by employing a sole type of behavior data (*i.e.*, mouse click frequency).

3. Study Design

In this paper, we present the results of two data-driven studies investigating if is possible to use mouse click frequency to predict students' flow experience. In both cases, *i)* the number of students' mouse clicks when using different systems was collected, then *ii)* the students answered a questionnaire to analyze their flow experience, and finally, *iii)* statistical analyzes were performed to identify if the mouse click frequency would be able to predict the students' flow experience (*i.e.*, learning analytics).

3.1. Materials and method

To collect the students' mouse click frequency, we used two different systems. For both experiments, there were no selection criteria for participants (*e.g* gender, age group, among others), which was a decision made to achieve a heterogeneous population. In addition to being research-oriented, they provide straightforward access to various student interaction data logs, which allows for conducting the analyzes proposed in this study.

For the first study, we used Eagle-edu[1], a gamified educational system design according to the 21 gamification elements introduced by Toda et al. (2019). As a gamified learning platform, Eagle-edu was chosen as it is a commercial system widely used in the territory where the study was conducted (*i.e.* Brazil), which according to Wohlin et al. (2012), is beneficial to the ecological settings of the study. Eagle-edu is comprised of three main pages: *Home*, where users can keep track of their progress in the system; *Profile*, where users can view and edit their personal information; *Learn*, where the user can complete multiple kinds of tasks (*e.g*, *Quiz*, *Complement*, and *Pairs*). Figure 1 presents an interface of this system. Participants enrolled in an English course previously created by a team from a local language school, which involved performing basic tasks (*e.g* pair-matching of Portuguese and English words) and quizzes.

Figure 1. Eagle-edu's course page interface

For the second study, we utilized a gamified educational system named Quick Detector on Gender Flow (QDGF). In QDGF, students would answer a logic reasoning quiz containing 20 questions and, by doing so, participants could earn points, trophies, or even appear in a ranking. Figure 2 exhibits an example of the system.

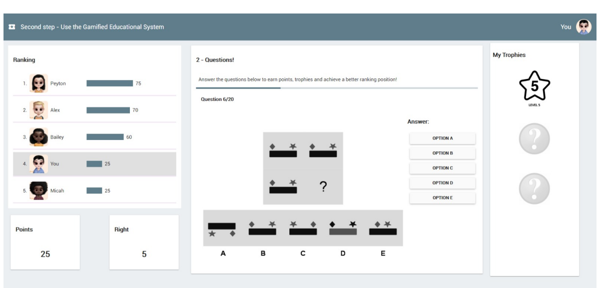

Figure 2. Quick Detector on Gender Flow's quiz page interface

When it comes to relating each of the nine dimensions of flow experience (Csikszentmihalyi, 1997a) to the data log collected (frequency of mouse clicks), both studies used the theoretical model introduced by Oliveira et al. (2019), as it is currently (as far know) the only model available for this purpose. To identify the student's flow experience, we used the flow state scale (FSS) developed by Jackson and Eklund (2002). We used this scale because it had been previously validated by Hamari and Koivisto (2014) to be used in the field of gamification, and according to Oliveira et al. (2018), is one of the most popular scales in studies in the area of educational technologies. To analyze the data, we used SmartPLS 3[2], a software for variance-based SEM using the partial least squares path modeling (PLS-PM) method (Wong, 2013).

Both studies were organized in **five** main steps: *i)* invitation *ii)* participation agreement *iii) and iv)* system login and usage *v)* flow experience self-report *iv)* data analysis. Figure 3 illustrates the steps of the process. In steps *iv) and v)*, the mouse click frequency is collected and then submitted to data analysis.

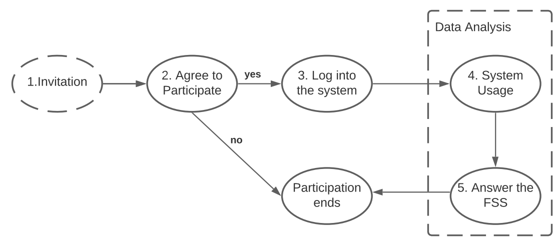

Figure 3. Method (step-by-step)

In the *first step*, the participants were invited to use the system. In the *second step*, participants had the opportunity to read the "Informed Consent Form" and decide about their participation in the study. In the *third step*, those who agreed to participate in the study logged into the system. In the *fourth step*, participants used the system (*i.e.*, a beginner-level English language course for Eagle-edu and a logic reasoning quiz for QDGF). At this stage, the number of students' mouse clicks was collected during the system usage. Immediately after completing the tasks designated by the system, in the *fifth step*, participants proceeded to answer the FSS (Hamari and Koivisto, 2014; Jackson and Eklund, 2002).

3.2. Participants and data analysis

In the first study (conducted in the system Eagle-edu), initially, a total of 25 students participated

[1] https://eagle-edu.com.br/

[2] https://www.smartpls.com/

voluntarily. Participants were invited through social media so that a heterogeneous sample could be achieved. Self-reportedly, 19 of them identified as males and six as females. Education levels included: high school, undergraduate, or graduate. The average age was 25 years, with a standard deviation of $\sigma = 3.13$. One participant, however, was excluded as they were not a native speaker of the language of the scale used in this study.

In the second study, we had 101 participants ranging from six levels of education (49 graduate students, 30 from high school, six at an MBA level, 10 Masters, one Ph.D. recipient, and four participants with other responses) whose age averaged approximately 18 years old (standard deviation of $\sigma = 10.55$) and originated from 15 countries (most significantly the USA with a total of 39; 15 from India; 12 from Portugal and 11 from Poland). Each participant received \$ 0.25 and was invited through a crowd-sourcing marketplace "Amazon Mechanical Turk (MTurK)"[3]. In both studies, we included an "attention-check" question.

To conduct the data analysis we used partial least squares structural equation modeling (PLS-SEM). This class of SEM is a well-established approach to developing theories on exploratory research. It achieves this by focusing on explaining the variance in the dependent variables when examining the model. We chose this method since it enables us to incorporate unobservable variables measured indirectly by indicator variables. Furthermore, PLS-SEM does not only allow the estimation of complex cause-effect relationship models with latent variables, it is robust despite a small sample (J. Hair et al., 2016; J. Hair et al., 2017). Following suggestions from J. Hair et al. (2016), J. Hair et al. (2017), we used β values to map/model the relationships between mouse frequency and flow experience and Adjusted R^2 to report the predictive power.

4. Results

Initially, given that the method employed by SmartPLS (PLS-SEM) is non-parametric (J. Hair et al., 2014), there was no need to analyze the distribution (normality) of the data. Furthermore, we estimated model reliability to ensure that the FSS matched the study data. Results are shown in Table 1 for the first study, and Table 2 for the second study, demonstrating that it was satisfactory for most dimensions of flow experience. Ideal rates are: $\alpha \geq 0.70$, RHO A ≥ 0.70, CR $\geq$0.70, AVE ≥ 0.50.

We then measured the discriminant validity. In the measurement model for study 1, the discriminant validity was measured using the cross-loading metric. The result showed that no construct's items loaded higher on another construct than itself (J. F. Hair et al., 2014). Furthermore, the LSC and CG constructs were removed from the model of Study 1 (Figure 4) because the composite reliability criterion for including them in the structural model was not satisfied. In other words, the AVE metric was less than 0.5 for both constructs. The second study presented the following results: $Flow = 0.498$ and $NMC = -0.203$. Table 3 and Table 4 present the discriminant validity of the studies.

Next, we analyzed if the mouse click frequency predicts students' flow experience by using PLS-SEM. We used the regression coefficient (β) to analyze the relationship between the variables and in the *p-values* related to each β (this is done to check if the relationships are indeed significant). Furthermore,

Table 1. Reliability results for the flow experience dimensions for Study 1

	α	RHO A	CR	AVE
AE	0.963	-1.563	0.839	0.575
CSB	0.796	0.826	0.851	0.59
CTH	0.813	0.77	0.825	0.618
CG	0.809	0.43	0.644	0.389
LSC	0.853	-1.813	0.016	0.122
MAA	0.779	1.02	0.852	0.602
SC	0.699	0.498	0.697	0.459
TT	0.908	0.669	0.900	0.695
UF	0.896	1.004	0.922	0.750

Key:α: Cronbach's; RHO A: Jöreskog's rho; CR: Composite Reliability; AVE: Average; Variance Extracted; AE: Autotelic Experience; CSB: Challenge-Skill Balance; CTH: Concentration on the Task at Hand; CG: Clear Goals; LSC: Loss Self-Consciousness; MAA: Merging of Action and Awareness; SC: Sense of Control; TT:UF: Unambiguous Feedback Transformation of Time;

Table 2. Reliability results for the Flow Experience for Study 2

	α	RHO A	CR	AVE
Flow	0.735	0.383	0.704	0.248

Key:α: Cronbach's; RHO A: Jöreskog's rho; CR: Composite Reliability; AVE: Average; Variance Extracted; AE: Autotelic

[3]https://www.mturk.com/

Table 3. Discriminant Validity for Study 1

	AE	CSB	CTH	CG	LSC	MAA	NMC	SC	TT	UF
AE	0.758									
CSB	0.371	0.768								
CTH	0.438	0.419	0.786							
CG	0.173	0.607	-0.081	0.623						
LSC	-0.224	-0.28	-0.439	-0.084	0.35					
MAA	0.124	0.255	-0.121	0.14	0.193	0.776				
NMC	-0.243	-0.401	0.304	-0.547	-0.24	-0.326	1			
SC	0.37	0.511	0.123	0.692	-0.162	0.265	-0.63	0.677		
TT	0.565	0.18	0.34	0.137	-0.157	0.131	-0.198	0.453	0.833	
UF	0.261	0.581	-0.003	0.821	-0.254	0.067	-0.368	0.526	0.095	0.866

Key: AE: Autotelic Experience; CSB: Challenge-Skill Balance; CTH: Concentration on the Task at Hand; CG: Clear Goals; LSC: Loss of Self-Consciousness; MAA: Merging of Action and Awareness; SC: Sense of Control; TT: Transformation of Time; UF: Unambiguous Feedback.

Table 4. Discriminant Validity for Study 2

	Flow	NMC
Flow	0.498	
NMC	-0.203	1

Key: NMC: Number of Mouse Clicks;

this study also utilized the $R^2 - value$ so that we could acknowledge how effectively mouse clicks predict flow experience (for greater reliability, we reported the adjusted $R^2 - value$). Table 5 presents the results for Study 1 with the aforementioned metrics. Likewise, Table 6 present the results for Study 2. Figure 4 and Figure 5 present the path model for both studies.

In summary, the correlation between a single type of data log (mouse click frequency) and the flow experience fails to provide significant predictive power, as revealed by the values of R^2Adj ($R^2Adj = 0.031$ in Study 1 and ranging from $R^2Adj = -0.004$ in the TT dimension to $R^2Adj = 0.370$ in the SC dimension in Study 2) and β ($\beta = -0.530$ in Study 1 and ranging from $\beta = -0.630$ in the SC dimension to $\beta = 0.304$ in the CTH dimension as observed in Study 2) in both studies. Our experimental results imply that the mouse click frequency by itself cannot be used as a metric for flow experience.

4.1. Discussion

In this paper, we analyze the possibility of predicting students' flow experience based only on the mouse click frequency performed by the student when using gamified educational systems. Throughout two studies, we identified that the mouse click frequency alone is not able to predict the flow experience of students when using such systems.

Our results can be considered surprising, as they are contrary to the results of recent studies which identify that the mouse click frequency can be used to predict experiences related to flow experience (Oliveira, Isotani, et al., 2021; Z. Zheng et al., 2019), such as stress and anxiety, as well as the theoretical model proposed by Oliveira, Toda, Palomino, Rodrigues, Shi, et al. (2020), who states that the mouse click frequency would be related to some flow experience dimensions.

The study conducted by Z. Zheng et al. (2019) utilized mouse activities inside IDEs (*e.g.*, mouse clicks) to track developers' interactions during work to recognize the flow state. Their findings point out an overall recognition accuracy of 92.6%, suggesting the *effectiveness* of their proposed approach. Oliveira, Toda, Palomino, Rodrigues, Shi, et al. (2020) also presented optimistic results as their work identified a significant relationship between four types of data logs (including NMC) and seven out of the nine aforementioned flow experience dimensions.

In the study conducted by Oliveira et al. (2019), the number of mouse clicks is seen as a theoretically related behavior data to total concentration on the task at hand (*i.e.*, one of the flow experience dimensions). Oliveira et al. (2019), despite putting forward this relationship and presenting theoretical reasons for it, also proposes that experimental studies be carried out to investigate these relationships. In this way, our study brings insights in the sense that possibly the theoretical relationships proposed by Oliveira et al. (2019) are not confirmed in practice.

Other studies were also conducted investigating the relationships of user behavior data in educational systems with their flow experience, however, studies that

Table 5. Correlational matrix for Study 1

	β	P-values	CI 2,5%	CI 97.5%	Adj. R^2
NMC →AE	-0.243	0.468	-0.553	0.469	0.016
NMC→CSB	-0.401	0.417	-0.777	0.632	0.123
NMC→CTH	0.304	0.399	-0.670	0.604	0.051
NMC→CG	-0.547	0.262	-0.826	0.502	0.268
NMC→LSC	-0.240	0.530	-0.715	0.553	0.015
NMC→MAA	-0.326	0.247	-0.717	0.491	0.066
NMC→SC	-0.630	0.235	-0.843	0.681	0.370
NMC→TT	-0.198	0.533	-0.557	0.559	-0.004
NMC→UF	-0.368	0.419	-0.782	0.523	0.096

Key: β: Regression Coefficient; CI: Confidence Interval; NMC: Number of Mouse Clicks; AE: Autotelic Experience; CSB: Challenge-Skill Balance; CTH: Concentration on the Task at Hand; CG: Clear Goals; LSC: Loss of Self-Consciousness; MAA: Merging of Action and Awareness; SC: Sense of Control; TT: Transformation of Time; UF: Unambiguous Feedback

Table 6. Correlational Matrix for Study 2

	β	P-Values	R^2 Adj	CI 2.50%	CI 97.50%
NMC→ Flow	-0.203	0.334	0.031	-0.474	0.376

Key:β: Regression Coefficient; CI: Confidence Interval

analyzed mouse-related items performed the analyzes comparing different types of associated data (different types of related data mouse), but without analyzing the data separately. One of the studies that directly compared the relationship between user behavior data concerning the mouse and the flow experience of participants was that of Lee et al. (2014). In their study, they identified that the average mouse click duration was positively correlated with boredom but negatively correlated to flow. In a comparative analysis, it is possible to argue that although the mouse click frequency is insufficient to predict students' flow experience, associating this information with other data related to mouse actions can be an equally affordable and viable alternative to predicting flow experience.

The study by Semerci and Goularas (2021) identified that deep learning methods were able to extract student behavioral patterns based on mouse interactions. Although these behaviors cannot be generalized, given the sample size of Semerci and Goularas (2021), it is also an indication that a single mouse-related data is not effective in predicting the flow experience. However, by gathering more than one type of data log, it may become a feasible approach to predicting flow experience, or some of the dimensions that compose flow experience.

These results demonstrate that despite the use of user behavior data in educational systems presenting weak predictive power to model flow experience, they are a promising alternative to predict different user experiences, as the rapid growth in online educational systems generates more and more behavior data logs (Eberle and Hobrecht, 2021). However, our findings make it clear that the mouse click frequency alone cannot be used to predict the flow experience. Nonetheless, the fact that the path coefficients in the models are numerically high, although non-significant, calls for further studies with larger sample sizes that will allow for the stratification of the data.

In summary, our results indicate that predicting the flow experience of students based only on the number of mouse clicks during the use of an educational system is not plausible. However, by comparing our results with different studies, it can be argued that adopting multiple mouse-related behavior data logs is the pathway to accurately modeling individual dimensions of or flow experience itself.

4.2. Limitations

This study has some limitations stemming from its characteristics which we seek to attenuate throughout. Initially, flow is a complex experience to measure, with a total of nine dimensions observed in this study. Also, The study was conducted remotely, thus,

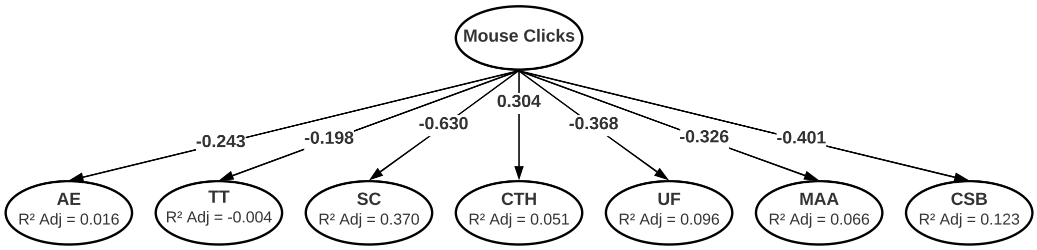

Figure 4. Path Model for Study 1

Figure 5. Path Model for Study 2

non-observable external variables may interfere with the participants' experience. To mitigate this limitation, we used only empirically validated instruments. Another limitation comes from the limited sample size of the first study, which might have impacted the robustness of the calculated values obtained through SmartPLS. In both studies conducted, only one type of education system was used (*i.e.*, gamified education systems). Thus, we cannot guarantee that the results can be generalized to other types of systems. The use of only one predictor variable can reduce the results' robustness, as the surrounding noise is not captured. Thus, complementary metrics, such as cursor position and behavior on the screen, may be used along with other mouse-related behavior data analyses.

4.3. Recommendations for future studies

Although we utilized databases from two different studies, the number of participants is still limited to conduct robust analyses, for example, machine learning/deep learning. Thus, we suggest future studies carry out experiments with a large number of participants, which allows the use of this type of technique for machine-learning-based predictive modeling (possibly increasing the generalization of the results).

Recent studies have increasingly realized that students' experiences can change according to their characteristics. Therefore, further studies can collect more data from students that allows for identifying whether the prediction of the flow experience may vary according these and other factors of students. We did not have an adequate sample that will allow for stratified (multi-group) analysis to uncover the moderating effect of key demographic variables such as age, gender, etc. Thus, future studies (with larger sample sizes) should conduct multi-group analyses to uncover the moderating effect of demographic variables.

In our study, we only used the number of mouse clicks in our analysis, and our findings demonstrate that it has little predictive power for flow experience. However, other types of mouse-related data can also be collected as well (*e.g.*, mouse pointer movements). If, on the one hand, our results show that the number of mouse clicks alone is not enough to predict the flow experience, on the other hand, the possibility of using different types of mouse-related data also represents an opportunity to identify a new low-cost way of predicting the flow experience. The same can be applied to different peripheral devices as Students may prefer to use touchscreens or touch pads to interact with educational systems. Thus, future studies may explore all permutations of different data logs and input devices in order to observe which combination leads to a more accurate ecosystem for flow experience modeling.

5. Concluding Remarks

We conducted two data-driven studies analyzing if mouse click frequency can be used as a single variable to predict students' flow experience. Our results demonstrate that relying on such a strategy to model flow experience is not reliable, as the mouse click frequency alone does not present significant predictive power. In future studies, we aim to replicate this experiment by collecting data from various types of educational systems. We also aim to include new kinds of mouse-related behavior data (*e.g.*, mouse movements) and using novel data analysis methods, such as data mining and deep learning techniques.

Acknowledgment

This work has been supported by the Academy of Finland Flagship Programme [Grant No. 337653 - Forest-Human-Machine Interplay (UNITE)].

References

Chen, Y., Liu, Y., Zhang, M., & Ma, S. (2017). User Satisfaction Prediction with Mouse Movement Information in Heterogeneous Search Environment. *IEEE Transactions on Knowledge and Data Engineering*, *29*(11), 2470–2483.

Chou, Y.-S., Hou, H.-T., Chang, K.-E., & Su, C.-L. (2021). Designing cognitive-based game mechanisms for mobile educational games to promote cognitive thinking: An analysis of flow state and game-based learning behavioral patterns. *Interactive Learning Environments*, 1–18.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. Jossey-Bass Publishers. http://books.google.com/books?id=afdGAAAAMAAJ

Csikszentmihalyi, M. (1997a). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

Csikszentmihalyi, M. (1997b). Flow and education. *NAMTA journal*, *22*(2), 2–35.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press.

De Kock, F. G. (2014). The neuropsychological measure (eeg) of flow under conditions of peak performance.

Eberle, J., & Hobrecht, J. (2021). The lonely struggle with autonomy: A case study of first-year university students' experiences during emergency online teaching. *Computers in Human Behavior*, *121*, 106804. https://doi.org/https://doi.org/10.1016/j.chb.2021.106804

Egbert, J. (2004). A study of flow theory in the foreign language classroom. *Canadian modern language review*, *60*(5), 549–586.

Erhel, S., & Jamet, E. (2019). Improving instructions in educational computer games: Exploring the relations between goal specificity, flow experience and learning outcomes. *Computers in Human Behavior*, *91*(September 2018), 106–114. https://doi.org/10.1016/j.chb.2018.09.020

Faiola, A., Newlon, C., Pfaff, M., & Smyslova, O. (2013). Correlating the effects of flow and telepresence in virtual worlds: Enhancing our understanding of user behavior in game-based learning. *Computers in Human Behavior*, *29*(3), 1113–1121.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative health research*, *3*(4), 430–441.

Gao, B., Wan, Q., Chang, T., & Huang, R. (2019). A framework of learning activity design for flow experience in smart learning environment. *Foundations and trends in smart learning* (pp. 5–14). Springer.

Guo, Q., & Agichtein, E. (2010). Towards predicting web searcher gaze position from mouse movements. *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, 3601–3606. https://doi.org/10.1145/1753846.1754025

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, *26*(2), 106–121. https://doi.org/10.1108/EBR-10-2013-0128

Hair, J., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling*. SAGE.

Hair, J., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (pls-sem)*. Sage publications.

Hair, J., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in partial least squares structural equation modeling*. saGe publications.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional flow scale-2. *Computers in Human Behavior*, *40*, 133–143.

Jackson, S., Eklund, B., & Martin, A. (2011). The flow manual the manual for the flow scales manual. sampler set. *Mind*, 1–85.

Jackson, S., & Eklund, R. (2002). Assessing flow in physical activity: The flow state scale–2 and dispositional flow scale–2. *Journal of Sport and Exercise Psychology*, *24*(2), 133–150.

Jackson, S., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of sport and exercise psychology*, *18*(1), 17–35.

Lee, P.-M., Jheng, S.-Y., & Hsiao, T.-C. (2014). Towards automatically detectinfg whether student is in flow. *International Conference on Intelligent Tutoring Systems*, 11–18.

Oliveira, W., Bittencourt, I. I., Isotani, S., Dermeval, D., Marques, L. B., & Silveira, I. F. (2018). Flow theory to promote learning in educational systems: Is it really relevant? *Brazilian Journal of Computers in Education*, *26*(02), 29.

Oliveira, W., Isotani, S., Pastushenko, O., Hruska, T., & Hamari, J. (2021). Modeling students' flow experience through data logs in gamified educational systems. *Proceedings - IEEE 21st International Conference on Advanced Learning Technologies, ICALT 2021*, 97–101. https://doi.org/10.1109/ICALT52272.2021.00037

Oliveira, W., Pastushenko, O., Rodrigues, L., Toda, A., Palomino, P., Hamari, J., & Isotani, S. (2021). Does gamification affect flow experience? a systematic literature review. *GamiFIN Conference 2021: Proceedings of the 5th International GamiFIN Conference*, 1–10.

Oliveira, W., Toda, A., Palomino, P., Rodrigues, L., Isotani, S., & Shi, L. (2019). Towards automatic flow experience identification in educational systems: A theory-driven approach. *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 581–588.

Oliveira, W., Toda, A., Palomino, P., Rodrigues, L., Shi, L., & Isotani, S. (2020). Towards automatic flow experience identification in educational systems: A qualitative study. *Proceedings of the XXXI Brazilian Symposium on Computers in Education*, 702–711. https://doi.org/10.5753/cbie.sbie.2020.702

Oliveira, W., Toda, A., Toledo, P., Shi, L., Vassileva, J., Bittencourt, I. I., & Isotani, S. (2020). Does tailoring gamified educational systems matter? the impact on students' flow experience. *The Hawaii International Conference on System Sciences*.

Özhan, Ş. Ç., & Kocadere, S. A. (2020). The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment. *Journal of Educational Computing Research*, *57*(8), 2006–2031.

Pentel, A. (2015). Using Mouse Logs to Predict User's Emotional State. *Personalization Approaches in Learning Environments*, 40–45.

Perttula, A., Kiili, K., Lindstedt, A., & Tuomi, P. (2017). Flow experience in game based learning–a systematic literature review. *International Journal of Serious Games*.

San Pedro, M. O. Z., d Baker, R. S., Gowda, S. M., & Heffernan, N. T. (2013). Towards an understanding of affect and knowledge from student interaction with an intelligent tutoring system. *International Conference on Artificial Intelligence in Education*, 41–50.

Semerci, Y. C., & Goularas, D. (2021). Evaluation of students' flow state in an e-learning environment through activity and performance using deep learning techniques. *Journal of Educational Computing Research*, *59*(5), 960–987.

Toda, A. M., Oliveira, W., Klock, A. C., Palomino, P. T., Pimenta, M., Gasparini, I., Shi, L., Bittencourt, I., Isotani, S., & Cristea, A. I. (2019). A taxonomy of game elements for gamification in educational contexts: Proposal and evaluation. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, *2161-377X*, 84–88.

van Schaik, P., Martin, S., & Vallance, M. (2012). Measuring flow experience in an immersive virtual environment for collaborative learning. *Journal of Computer Assisted Learning*, *28*(4), 350–365.

Wan, Q., Liu, M., Gao, B., Chang, T., & Huang, R. (2020). The relationship between self-regulation and flow experience in online learning: A case study of global competition on design for future education. *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 365–367.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.

Wong, K. K.-K. (2013). Partial least squares structural equation modeling (pls-sem) techniques using smartpls. *Marketing Bulletin*, *24*(1), 1–32.

Yen, W.-C., & Lin, H.-H. (2020). Investigating the effect of flow experience on learning performance and entrepreneurial self-efficacy in a business simulation systems context. *Interactive Learning Environments*, 1–16.

Zheng, N., Paloski, A., & Wang, H. (2011). An efficient user verification system via mouse movements. *Proceedings of the 18th ACM Conference on Computer and Communications Security*, 139–150.

Zheng, Z., Wang, L., Cao, Y., Zhuang, Y., & Tao, X. (2019). Towards non-invasive recognition of developers' flow states with computer interaction traces. *2019 26th Asia-Pacific Software Engineering Conference (APSEC)*, 300–307.