The Relationship Between Students' Flow Experience and Their Behavior Data in Gamified Educational Systems

Wilk Oliveira
University of São Paulo
wilk.oliveira@usp.br

Kamilla Tenório
University of São Paulo
kktas@usp.br

Juho Hamari
Tampere University
juho.hamari@tuni.fi

Seiji Isotani
University of São Paulo
sisotani@icmc.usp.br

Abstract

The number of students using online educational systems is increasing, especially after the growth of the use of this type of system due to the social isolation caused by the Covid-19 pandemic. This situation highlighted the challenge of analyzing the users' experience in this type of system, especially when evaluating more complex experiences, such as the flow experience. One of the most promising innovative alternatives is to use the behavior data logs produced by students in educational systems to analyze their experiences. In this paper, we conducted a study (N = 24) to analyze the relationships between the behavior data logs produced by students when using a gamified educational system and their flow experience during the system usage. Our results contribute to the automatic users' experience analysis in educational systems.

1. Introduction

Over the years, technological advances have transformed many activities in diverse fields of our society [1, 2]. Due to the emergence of Covid-19, this digital transformation had to be accelerated, and technology was a powerful tool to help combat this pandemic and support the continuation of work in different fields, including medicine, entertainment, finance, and education [3, 4]. In the educational context, learning systems have been rapidly adapted, aiming to promote and support the use of technologies on online learning courses in all educational levels [5].

However, to successfully apply educational technologies in the online context, these environments need to include approaches that improve students' experiences and, consequently, improve students' learning outcomes [6]. An approach that has been increasingly adopted in the educational environments to enhance students' outcomes such as engagement, motivation, and flow, is gamification [7, 8, 9], *i.e.*, "transforming systems, service and activities to better afford similar motivational benefits as games often do" [10].

More specifically, the impact of flow experience (*i.e.*, an optimal experience resulting in intense engagement [11]) has been increasingly investigated by researchers of the educational technology field due to its promising positive impact on students' learning outcomes [12, 13, 14]. Previous studies have pointed out that flow experience had a beneficial influence on students' learning outcomes, such as enhanced students' memorization and comprehension [13], increased students' academic success through the increasing of motivation [12], and improved students' entrepreneurial self-efficacy [14].

Nonetheless, it is still a challenge to measure flow experience in gamified learning systems effectively [15, 16, 17, 18, 19]. Previous techniques adopted to measure flow experiences, such as electroencephalography (EEG), eye trackers, interviews, and questionnaires, presented limitations as high cost and the impossibility of conduct a massive application [17, 20, 21]. A more recent and promising approach is to measure flow experiences through the use of data logs of students in learning systems [17, 18, 19, 22]. However, this field is still poorly explored, and there is a lack of empirical data-driven study to investigate the relation between the users' data logs and their flow experience in gamified educational systems [17, 18, 19, 20, 23].

Considering it, in this study, we conducted data-driven research exploring the relationship between students' flow experience and their behaviour data logs in a gamified educational system. Our results show that data logs related to students' interaction with the gamification elements of the system have a significant relationship with seven flow experience dimensions, while data logs related to students' interaction with learning activities have a significant relationship with five flow experience dimensions. Therefore, we advance state-of-the-art, contributing to the modeling and automatic measurement of students' flow experiences in gamified learning systems.

URI: https://hdl.handle.net/10125/79338
978-0-9981331-5-7

2. Background and Related Works

In this section, we present the main topics related to this study (*i.e.*, *i)* gamification and flow experience in educational systems and *ii)* flow experience measurement in educational settings). We also present the main related works. Since the first time the term *gamification* was used, in 2008, there is a growing interest in this approach [7] and it has been applied in several domains, such as marketing [24], health [25], and education [26], which is the area more explored [7, 6, 27]. In the educational context, the literature points out that gamification can have a positive impact on student's flow when the gamification design is well planned [28]. When gamification facilitates the flow experience, it leads to better students' learning outcomes in educational environments [12].

Flow experience, according to the Flow Theory [11], can be summarized as an experience of deep engagement that a person can achieve in a given activity [11, 29, 30]. Due to the potential of positive effects of flow experience, there has been an increase of studies been conducted concerning flow experience in different fields such as sports [31], video games [32], besides the educational field [33]. This experience is represented by the interconnection of the following nine dimensions [29, 34, 35]: (1) challenge-skill balance (CSB); (2) action-awareness merging (MMA); (3) clear goals (G); (4) unambiguous feedback (F); (5) total concentration on the task at hand (C); (6) sense of control (CTRL); (7) loss of self-consciousness (LSC); (8) transformation of time (T); and (9) *autotelic* experience (A). It means that to achieve the flow experience, it is necessary that a person simultaneously reach these nine dimensions.

In order to measure flow experience, different methods were developed and adopted over the years [36, 37]. Originally, to measure the flow experience, it was used interviews and focus groups [29], and a system that requested people pressed a button when they reached an experience [36, 37]. However, these methods presented some limitations such as expensiveness and impossibility of conduct a massive application [21]. Therefore, other methods have been proposed afterward.

Another proposed method to measure flow experience is the use of questionnaires [36, 37]. Over the years, this method has become the most used method, and different questionnaires have been proposed and validated for different fields, such as physical activity [38], sports [39], and gamification [16]. Nevertheless, this method also presented some limitations, such as the difficulty of application in the context of distance learning, difficulty in interpreting the results, and the fact that it cannot be used continuously. Based on it, more recent methods have been proposed in the literature to measure flow experiences, such as EEG and eye trackers [23, 40, 41, 42]. However, these methods also presented limitations related to cost, no effective flow detection, and the impossibility of conducting a massive application [21].

Considering the limitations of the previously presented methods, researchers are investigating the possibility of measuring flow experience through data logs that are produced by the users while using the systems, and it has been presenting promising results in the educational technology field [17, 18, 22, 19].

Among the studies that investigated the use of data logs to measure flow experiences in an educational technology context is Lee *et al.* [17]. Lee *et al.* [17] were one of the firsts to investigate the relationship between students' data logs and their flow experience in learning systems. Lee *et al.* [17] presented an automated detector, using a step regression approach, to identify the students' flow experience during the learning process in a step-based tutoring system.

Another related work was conducted by Oliveira *et al.* [18], which proposed a theory-driven conceptual model, aiming to associate student's interaction data logs with each of the nine flow experience dimensions. Oliveira *et al.* [22], in another study, conducted a qualitative study using the think-aloud protocol to associate users' data logs with their flow experience in an educational system. In two other studies from the same project, Oliveira *et al.* [43, 44], advanced the first studies, using structural equation modeling to model and predict students' experience through data logs in a gamified task.

Semerci and Goularas [19] proposed a method based on flow theory to provide information about the students' flow state in a learning system. In order to measure flow experience, they used activity heatmaps, deep neural networks, and students' grades [19]. Table 1 present a comparison between the related works (previously presented in this section) and our study (presented in this paper).

Although previous studies have investigated the use of users' data logs to measure flow experiences in educational contexts, no study seeks to identify pattern relation between the users' data logs and their flow experience in a gamified educational system. Therefore, considering the beneficial effects of flow experience on students' learning in gamified educational systems [12], as far we know, our study is the first to explore the relationship between students' flow experience and their data logs in a gamified learning environment through empirical data-driven research and using modern, robust data analysis techniques.

Table 1. Comparison between the related works

	UDL	GE	NFeD	TofS	DaT
[17]	Yes	No	1	DD	SR
[18]	Yes	No	9	TD	NA
[22]	Yes	No	9	QD	TA
[43, 44]	Yes	No	9	DD	SEM
[19]	Yes	No	9	DD	DNN
Our study	Yes	Yes	9	DD	SEM

Key: UDL: analyzed user data logs; GE: analyzed gamification elements; NFeD: Number of flow experience dimensions analyzed in the study; TofS: type of study; DaT: Data analysis technique adopted; DD: Data-driven study; TD: Theory-driven study; QD: Qualitative-driven study; SR: Step regression; NA: Not applicable; TA: Think aloud; DNN; SEM: Structural equation modeling; Deep neural network.

3. Study Design

The objective of this study was to analyze the relationship between students' flow experience when using a gamified educational system and their behaviour data logs during the system usage. To achieve this goal, we conduct a data-driven study, analyzing ecological (*i.e.*, non-simulated data) student data when using a gamified educational system.

3.1. Materials

To conduct the research, we used a gamified educational system called "Eagle-edu[1]". The system is used nationally and allows teachers/instructors to create different types of courses and distribute them to their students. The system consists of three main pages: *Home*, where students can track their progress and evolution in the system; *Learn*, where students can access the educational activities; and *Profile*, where students can access their general information. The main pages have a sub-page called *Store* where students can buy virtual items (with the fictitious virtual coins you earned when doing the activities) and another sub-page called *Friends*, where students can view other colleagues and follow them. The system have three kind of tasks (to compose a mission) in the *Learn* page, **quiz**: simple questions with different answer options; **complement**: simple sentences with white space to complete; and **pairs**: presentation of options for students to choose pairs.

Missions consist of different tasks. The teacher/instructor defines the number of tasks (that compose a mission) and missions (that compose a subject). The mission is only closed when the student finishes all tasks. The system gamification design is composed of 21 gamification elements and organized in five dimensions, as defined by Toda *et al.* [45]:

- **Performance/measurement**: This dimension is composed by five gamification elements (*i.e.*, Acknowledgment, Level, Point, Progression, and Stats). These are elements related to the environmental response, which can be used to provide feedback to the learner [45].
- **Ecological**: This dimension is represented by five gamification elements (*i.e.*, Chance, Economy, Imposed Choice, Rarity, and Time Pressure). This context is related to the environment that the gamification is being implemented [45]. These elements can be represented as properties [45].
- **Social**: This dimension is composed by four gamification elements (*i.e.*, Competition, Cooperation, Reputation, and Social Pressure). This dimension is related to the interactions between the learners presented in the environment [45].
- **Personal**: This dimension is composed by five gamification elements (*i.e.*, Novelty, Objectives, Puzzles, Renovation, and Sensation). This dimension is related to the learner that is using the environment [45].
- **Fictional**: This dimension is composed by two gamification elements (*i.e.*, Narrative and Storytelling). It is the mixed dimension that is related to the user (through Narrative) and the environment (through Storytelling), tying their experience with the context [45].

The system was chosen because it is widely used in the country where the study was conducted and can be used free of charge for research purposes, with all its resources available, improving the ecological settings of the study [46]. The system also allows access to various student interaction data logs, which allows conducting the analyzes proposed in this study. Another important point is that the system was entirely implemented (including the gamification design) for educational purposes. The study was carried out using an English course previously created by a team from a language school.

To select the students' data logs to be collected in the system, we use the theoretical model proposed by Oliveira *et al.* [18]. The model proposes a theoretical

[1] `http://eagle-edu.com.br/`

relation between different types of students' interaction data logs and the nine original flow experience dimensions proposed by Csikszentmihalyi [30]. The model was chosen because, as far as we know, it is the only model for this purpose. At the same time, it relates data logs individually to each of the original flow experience dimensions [30], which tends to deepen the analysis. For this study, we collected six data logs from the theoretical model proposed by Oliveira *et al*. [18]: *i)* number of mouse clicks; *ii)* number of completed missions; *iii)* number of completed tasks; *iv)* number of wrong tasks; *v)* active time in the system; and *vi)* average response time. Additionally, we collected two data logs regarding system gamification *i)* number of points; and *ii)* economy (number of virtual coins).

To analyze the students' flow experience, we used the flow state scale (FSS) proposed by Jackson and Eklund [38]. The scale consists of 36 questions, representing the nine original dimensions of the flow experience [30]. We used a 5-point Likert scale [47], following the recommendations present in the scale application manual [35]. The scale was chosen because, according to Oliveira *et al*. [20], it is the most used in the area of educational technologies. Another important point is that as far we know, this scale is the only one that has been validated for the gamification domain [16] and it was also validated in the native language of the study participants [48]. In addition, following the example of recent studies [49, 50, 51], we inserted an "attention-check" statement (*i.e.*, "This is an attention-check question, if you have read this question, check option 3") to verify whether participants were paying attention to the questions on the scale. Responses from participants who mistake the attention check statement are excluded from the analyzes.

3.2. Procedure

We organized the study in four general steps. In the *first step*, we invited high school, undergraduate, and graduate students through the dissemination of the research through email lists (academic and non-academic) and social networks (Facebook and WhatsApp). Participants who opened the invitation link were able to accept participate or not. In the *second step*, participants who agreed to participate in the study were asked to create an account in the system and then log in to the system. Those who did not accept to participate had their participation in the study immediately ended.

After logging into the system, in the *third step*, participants could choose an avatar in the system (this choice is not mandatory in the system). Posteriorly, the participants were able to make the activities (missions) available in the system. In the interval between the missions, the students could access the store, where they could buy virtual items with the coins earned when doing the activities and access the friends' page. It is important to note that the store and the friends' page were not mandatory activities in the system.

In the *fourth step*, immediately after ending the system usage, the participants were redirected to respond to the FSS. Considering that it is about measuring an experience when using an educational system, we asked participants to start responding to scale immediately after using the system. In our study, we analyzed only responses inserted immediately after the last activity in the system, thus, avoiding responses where the participant may not be specifically referring to their experience in the system. Figure 1 present a summary for the study procedure.

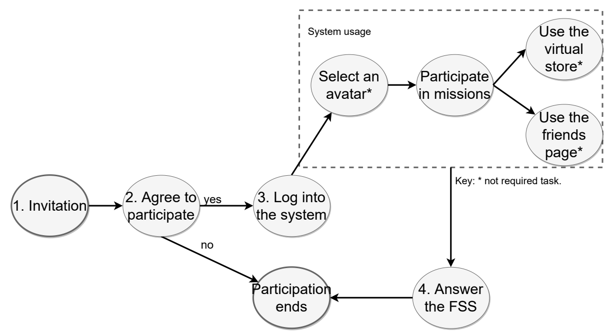

Figure 1. Step-by-step of the study

3.3. Participants

Initially, our participants were 25 students, 19 self-reported as males and six as females. They were high school, undergraduate, or graduate students with an average age of 25 years. One participant was excluded for not being native in the language of the scale in which the study was conducted. None of the participants failed the other exclusion criteria, so we kept them in the research. Due to the number of participants and the nature of the research, no outliers were analyzed.

4. Results

Initially, to define the most suitable strategies for data analysis, following the recommendations of Wohlin [46], we analyzed the distribution (normality) of the data. For this, as our sample is less than 30 participants, we used the Shapiro-Wilk test [52], as recommended by Wohlin [46]. The results indicate that the data do not come from a normal distribution. Therefore, to ensure that the scale (*i.e.*, the FSS proposed by

Jackson and Eklund [38]) used in the study matched the study data, we analyzed the internal reliability for each flow experience dimension (using Cronbach's alpha test [53]). The results show that the internal validity was acceptable in all sub-scales ($\alpha \geq 0.70$).

After that, we analyzed the correlation between each student's data logs and each of the flow experience dimensions. As the data does not follow a normal distribution, following Wohlin's [46] recommendations, we used Kendall's tau-b test [54] to analyze the correlations. The results indicate that most of the correlations were low or moderate, while only three correlations were significant. Results indicate that the number of virtual coins has a moderate correlation with with the dimensions of concentration and *autotelic* experience. Moreover, average response time has a moderate correlation with with the *autotelic* experience dimension. Table 2 present the correlations. We used the software IBM SPSS 27 (2021)[2] to conduct the described analysis.

To analyze the relationships between students' data logs and each of the flow experience dimensions, we used partial least squares structural equation modeling (PLS-SEM), a method of structural equation modeling that allows estimating complex cause-effect relationship models with latent variables [55]. More specifically, we use PLS-SEM because it is a technique capable of analyzing models of relationships between variables and estimating cause-and-effect relationships robustly, even when using a small sample, according to Hair *et al.* [56] . We used the software SmartPLS[3], that provides a graphical interface to calculate PLS [57]. Initially, we estimated the model reliability. We present the reliability results in Table 3. Overall, the reliability was acceptable ($\alpha \geq 0.70$, RHO A ≥ 0.70, CR ≥ 0.70, AVE ≥ 0.50) for all flow experience dimensions, except the "action-awareness merging" and "sense of control" dimensions.

Finally, we estimated the relationships between each student's data logs and each flow experience dimension. In this study, we are especially interested in the regression coefficient (β) to analyze the direction of the relationship between the variables and in the *p-values* related to each β (to analyze whether the relationships are significant). Moreover, we are also interested in the R^2-*value* to know how much each data log predicts each flow experience dimension (for greater reliability, we reported the adjusted R^2). Table 4 present the relationship between students' flow experience and their data logs and organize the other values.

[2] Available in: `https://www.ibm.com/products/spss-statistics`

[3] Available in: `https://www.smartpls.com/`

In general, the results show low or time-consuming relationships between variables, with a total of 19 significant relationships being identified. Points present a significant relationship with action-awareness merging (β = -0.494 | R^2 = 0.210), concentration (β = -0.368 | R^2 = 0.096), and *autotelic* experience (β = -0.279 | R^2 = 0.036). Economy presented a significant relationship with clear goals (β = 0.398 | R^2 = 0.067), unambiguous feedback (β = -0.296 | R^2 = 0.046), total concentration on the task at hand (β = 0.366 | R^2 = 0.095), transformation of time (β = 0.347 | R^2 = 0.080), and *autotelic* experience (β = 0.323 | R^2 = 0.064). Number of mouse clicks showed no significant relationship.

Number of completed missions presented a significant relationship with action-awareness merging (β = -0.541 | R^2 = 0.261), concentration (β = 0.368 | R^2 = 0.096) and *autotelic* experience (β = -0.272 | R^2 = 0.032). Number of completed tasks presented a significant relationship with action-awareness merging (β = -0.523 | R^2 = 0.241) and concentration (β = 0.372 | R^2 = 0.099). Number of wrong tasks presented a significant relationship with concentration (β = -0.283 | R^2 = 0.038) and *autotelic* experience (β = -0.275 | R^2 = 0.034). Active time in the system also presented a significant relationship with action-awareness merging (β = -0.442 | R^2 = 0.159), concentration (β = -0.310 | R^2 = 0.055) and *autotelic* experience (β = 0.281 | R^2 = 0.037). Finally, average response time presented a significant relationship with concentration (β = 0.306 | R^2 = 0.053). In general, the R^2 values were low, indicating low predictive power. However, this can be related to the small sample size, making room for the replication of this study with larger samples.

4.1. Discussion

In this study, we analyzed the relationships between students' data logs in a gamified educational system and their flow experience when using the system. In total, we explored the relationships between eight data logs and the nine original flow experience dimensions. The results demonstrate that, in general, the data logs and the flow experience dimensions are moderately related. Behaviour data logs related to the gamification of the system showed significant relationships with seven flow experience dimensions. Activity data logs in the system are related to five flow experience dimensions.

Overall, previously published studies usually described the flow experience only as the balance between the participants' skill level and the task's challenge level [17, 40, 58], which is only one of the dimensions of the flow experience [11]. In our

Table 2. Kendall's tau-b correlation between flow experience dimensions and students' data logs

	Pts	Ecm	NMC	NCM	NCT	NWT	ATS	ART
CSB	.065	.150	-.038	.055	.072	.108	.151	.212
MMA	-.182	-.113	-.119	-.229	-.205	-.205	-.152	-.122
G	.095	.227	.008	.053	.054	-.134	.066	-.016
F	.164	.288	.138	.120	.140	-.109	.095	-.075
C	.242	**.370***	.248	.252	.281	.278	.247	.135
CTRL	.147	.303	.000	.134	.157	.004	.179	-.027
LSC	-.084	-.172	-.063	-.090	-.116	-.296	-.086	-.126
T	.095	.267	-.030	.105	.111	.180	.138	.191
A	.137	**.331***	.120	.119	.138	.258	.273	**.390***

Key: * $p<.05$; Pts: Points; Ecm: Economy; NMC: Number of mouse clicks; NCM: Number of completed missions; NCT: Number of completed tasks; NWT: Number of wrong tasks; ATS: Active time in the system; ART: Average response time; CSB: challenge-skill balance; MMA: action-awareness merging; G: clear goals; F: unambiguous feedback; C: total concentration on the task at hand; CTRL: sense of control; LSC: loss of self-consciousness; T: transformation of time; and A: *autotelic* experience. Red color indicates negative correlation and green color indicates positive correlation. The intensity of the color indicates the level of correlation.

Table 3. Reliability results for the flow experience dimensions

	α	RHO A	CR	AVE
CSB	0.796	0.862	0.829	0.571
MMA	0.779	-0.314	0.006	0.224
G	0.809	0.916	0.872	0.635
F	0.896	0.253	0.839	0.590
C	0.813	0.890	0.887	0.676
CTRL	0.699	0.715	0.804	0.519
LSC	0.853	1.078	0.885	0.661
T	0.908	0.946	0.934	0.779
A	0.963	1.075	0.971	0.894

Key: α: Cronbach's; RHO A: Jöreskog's rho; CR: Composite Reliability; AVE: Average Variance Extracted; CSB: challenge-skill balance; MMA: action-awareness merging; G: clear goals; F: unambiguous feedback; C: total concentration on the task at hand; CTRL: sense of control; LSC: loss of self-consciousness; T: transformation of time; and A: *autotelic* experience.

study, we do not identify any significant relationship between data logs and the challenge-skill balance dimension. However, we advance the current literature describing new relationships with other flow experience dimensions few (or not) explored in previous studies (see Table 4).

Regarding the gamification elements, a negative relationship was identified between points and action-awareness merging. It could occur since when someone does not achieve an action-awareness merging experience, they may not be able to become aware of the activities and consequently, increasing the number of errors, which leads to earning fewer points. This relationship may be related to another result obtained in the study, indicating a positive relationship between points and concentration, which is related to the fact that a person with a high level of contraction tends to hit more activities and consequently increase the number of points.

Concerning the economy element, despite some significant relationships, these significant relationships are weak or null. This indicates that in fact, there is possibly no direct relationship between the data logs economy element. This also indicates that long-term studies can be conducted specifically, analyzing the relationship between some game elements and students' data logs. As far we know, our study is the first to present the relationship between gamification elements and students' flow experience.

Oliveira's *et al.* [43], was one of the previous studies that analyzed the relationship between data logs and the nine flow experience dimensions, despite demonstrating a relationship between data logs and the participants' flow experience, it does not make it clear exactly which specific data logs relates to each of the nine flow experience dimensions. Therefore, in our study, we were able to advance the literature by providing a clearer description of which data logs relate to each flow experience dimension.

We identified that active time in the system is negatively associated to action-awareness merging and

Table 4. Relationship between students' flow experience and their data logs

			CI						CI		
	β	***P-values***	**2.5%**	**97.5%**	**Adj. R^2**		β	***P-values***	**2.5%**	**97.5%**	**Adj. R^2**
Pts → CSB	0.341	0.427	-0.694	0.696	0.076	**Ecm → CSB**	0.359	0.229	-0.666	0.562	0.090
Pts → MMA	-0.494**	0.007	-0.688	0.452	0.210	**Ecm → MMA**	-0.493	0.073	-0.772	0.385	0.208
Pts → G	0.247	0.208	-0.554	0.370	0.018	**Ecm → G**	0.328**	0.002	-0.587	0.456	0.067
Pts → F	0.220	0.268	-0.614	0.406	0.005	**Ecm → F**	0.296**	0.002	-0.477	0.427	0.046
Pts → C	0.368*	0.001	0.125	0.533	0.096	**Ecm → C**	0.366***	0.000	0.168	0.481	0.095
Pts → CTRL	0.247	0.470	-0.809	0.463	0.018	**Ecm → CTRL**	0.386	0.430	-0.734	0.704	0.111
Pts → LSC	-0.472	0.343	-0.840	0.721	0.187	**Ecm → LSC**	-0.486	0.230	-0.722	0.784	0.201
Pts → T	0.249	0.509	-0.854	0.482	0.019	**Ecm → T**	0.347*	0.015	-0.637	0.451	0.080
Pts → A	0.279*	0.042	-0.172	0.466	0.036	**Ecm → A**	0.323***	0.000	0.114	0.459	0.064
NMC → CSB	-0.401	0.415	-0.766	0.695	0.123	**NCM → CSB**	0.398	0.379	-0.557	0.741	0.121
NMC → MMA	-0.326	0.200	-0.539	0.613	0.066	**NCM → MMA**	-0.541***	0.000	-0.766	0.289	0.261
NMC → G	-0.547	0.253	-0.858	0.456	0.268	**NCM → G**	0.232	0.331	-0.652	0.373	0.011
NMC → F	-0.368	0.410	-0.763	0.518	0.096	**NCM → F**	0.203	0.364	-0.656	0.403	-0.002
NMC → C	0.304	0.369	-0.824	0.447	0.051	**NCM → C**	0.368**	0.005	-0.209	0.548	0.096
NMC → CTRL	-0.630	0.227	-0.819	0.703	0.370	**NCM → CTRL**	0.207	0.578	-0.721	0.480	-0.001
NMC → LSC	-0.240	0.526	-0.757	0.404	0.015	**NCM → LSC**	-0.428	0.376	-0.715	0.777	0.146
NMC → T	-0.198	0.522	-0.422	0.671	-0.004	**NCM → T**	0.282	0.384	-0.745	0.513	0.038
NMC → A	-0.243	0.449	-0.717	0.345	0.016	**NCM → A**	0.272*	0.066	-0.209	0.460	0.032
NCT → CSB	0.347	0.410	-0.659	0.658	0.080	**NWT → CSB**	0.265	0.222	-0.753	0.389	0.028
NCT → MMA	-0.523**	0.009	-0.725	0.496	0.241	**NWT → MMA**	-0.589	0.071	-0.811	0.553	0.318
NCT → G	0.244	0.306	-0.660	0.405	0.017	**NWT → G**	0.210	0.466	-0.611	0.403	0.001
NCT → F	0.221	0.301	-0.548	0.427	0.006	**NWT → F**	0.129	0.562	-0.538	0.316	-0.028
NCT → C	0.372**	0.001	-0.177	0.508	0.099	**NWT → C**	0.283**	0.009	-0.468	0.408	0.038
NCT → CTRL	0.251	0.498	-0.734	0.504	0.020	**NWT → CTRL**	-0.368	0.388	-0.734	0.589	0.096
NCT → LSC	-0.441	0.354	-0.717	0.796	0.158	**NWT → LSC**	-0.509	0.089	-0.705	0.779	0.225
NCT → T	0.271	0.423	-0.774	0.497	0.032	**NWT → T**	0.348	0.062	-0.733	0.500	0.081
NCT → A	0.287	0.053	-0.220	0.473	0.041	**NWT → A**	0.275*	0.011	-0.255	0.404	0.034
ATS → CSB	0.262	0.369	-0.749	0.430	0.026	**ART → CSB**	0.308	0.124	-0.518	0.477	0.054
ATS → MMA	-0.442*	0.042	-0.642	0.655	0.159	**ART → MMA**	-0.438	0.366	-0.717	0.536	0.155
ATS → G	0.250	0.268	-0.617	0.429	0.020	**ART → G**	0.186	0.418	-0.542	0.391	-0.009
ATS → F	0.199	0.375	-0.625	0.380	-0.004	**ART → F**	0.115	0.671	-0.512	0.429	-0.032
ATS → C	0.310**	0.009	-0.445	0.447	0.055	**ART → C**	0.306*	0.026	-0.295	0.484	0.053
ATS → CTRL	0.232	0.556	-0.739	0.472	0.011	**ART → CTRL**	0.200	0.528	-0.812	0.434	-0.003
ATS → LSC	-0.421	0.407	-0.759	0.822	0.140	**ART → LSC**	-0.240	0.469	-0.558	0.620	0.015
ATS → T	0.309	0.292	-0.745	0.483	0.054	**ART → T**	0.245	0.333	-0.573	0.478	0.017
ATS → A	0.281*	0.011	-0.207	0.438	0.037	**ART → A**	0.265	0.121	-0.466	0.457	0.028

Key: * $p<0.1$, ** $p<0.05$, *** $p<0.01$; β: Regression Coefficient; Pts: Points; Ecm: Economy; NMC: Number of mouse clicks; NCM: Number of completed missions; NCT: Number of completed tasks; NWT: Number of wrong tasks; ATS: Active time in the system; ART: Average response time; CSB: challenge-skill balance; MMA: action-awareness merging; G: clear goals; F: unambiguous feedback; C: total concentration on the task at hand; CTRL: sense of control; LSC: loss of self-consciousness; T: transformation of time; and A: *autotelic* experience; CI: Confidence intervals (bias corrects); Adj. R^2: Adjusted R^2.

positively associated with concentration and *autotelic* experience. A similar result was also identified in the qualitative study conducted by Oliveira *et al.* [22], and it may be related to the fact that reaching an action-awareness merging (as well as *autotelic* experience) causes a user to remain focused and consequently keep doing an activity (*e.g.*, using a certain system) for a long period. At the same time, if a student does not have an action-awareness merging experience, they may not be able to spend much time on an activity [11, 29, 34].

Also, regarding the relationship between active time in the system and *autotelic* experience, this result is important because the *autotelic* experience is considered one of the most difficult to be measured and with few empirical results related to it [30, 39, 18]. One theoretical issue to explain this relationship is that, according to Oliveira *et al.* [18], if a student spends little time active in each section of the system, or if s/he starts and finishes the sections in the system frequently, it is possible to perceive a clue that the student is is not in "real *autotelic* experience".

At the same time, we identified a relationship between the number of completed missions and the dimensions of action-awareness merging (negative), concentration and *autotelic* experience (positive). These relationships can be explained by the theory that an individual focused on activity is less affected by external factors and, consequently, demanding efforts to perform a certain activity [11, 30, 39]. These dimensions define one of the clearest indications of being in flow, that is, totally focused in the present on a specific task being performed [18].

Specially, we identified a relationship number of completed tasks, number of wrong tasks and average response time with the students' concentration. Both data logs are related to spending more time on the system doing activities. Thus, these relationships can be explained by the fact that a more concentrated student tends to spend more time doing a certain activity.

The results of our study have two practical implications. The first one is related to the automatic identification of the flow experience in educational systems. Through the relationships identified in our research, it is possible to advance in understanding how a student's data logs model their flow experience in the system and consequently use this modeling to identify whether or not a student is having a flow experience when using the educational system. Based on our results, for example, if a student receives few points compared to others in a gamified educational system, they cannot achieve the action-awareness merging experience, and consequently, they does not achieve the flow experience.

The second practical implication concerns the possibility of using the study's insights to adapt the system according to the students' needs. Therefore, once identified that a student could not complete the system's missions, it is possible to infer that the student is not managing to have an action-awareness merging experience and, consequently, seek strategies to adapt the system or interact with students to improve their experience on the system.

4.2. Threats to validate, limitations and recommendations for future studies

As this study involves human beings using an educational system, some threats to the validity and limitations inherent to this type of study were observed. Initially, the number of participants can be considered small for some types of analysis. To mitigate this threat, we use robust data analysis techniques, which work well to analyze cause-and-effect relationships even with small sample sizes. The study data can only be representative of a specific context. To mitigate this threat, we suggest that future studies may replicate our study in different contexts (*e.g.*, different countries). Some study participants may have previously used the system. To mitigate this threat, we removed responses from users that had used the system before.

The experience analyzed in our study (*i.e.*, flow experience) is considered a complex experience to be observed. To mitigate this threat, we used only previously validated instruments to analyze the experience, in addition to analyzing the instrument's internal validity with the study data. The data logs collected in the study may not be sufficient to represent all the possibilities of a system. To mitigate this threat, we collected the data based on a theoretical model that relates students' data logs and the different flow experience dimensions. In our study, we analyze the relationships between students' flow experience and multiple data logs. This multiplicity can lead to unidentified mediating effects among the data logs.

Our study brings results that deepen the literature. However, as well as the other related works, it was conducted in a limited period. Thus, as also recommended by Oliveira *et al.* [43], longitudinal studies must be conducted in the future. In our study, we analyzed data from 24 participants. Although our results bring interesting insights, our results cannot be generalized. Therefore, our study must be replicated with a larger sample size. In our study, we used correlation techniques and SEM to analyze the relationships between students' flow experience and their data logs. We used SEM because it is a powerful technique for analyzing relationships even with small sample sizes. However, we recommend that future studies with a larger sample may use other techniques, such as data mining and machine learning to perform new analyzes.

5. Conclusions

Analyzing the relationships between students' data logs and their experience in educational systems is a contemporary challenge. In this study, we analyze the relationships between students' behaviour data logs in a gamified educational system and their flow experience when using the system. Our results demonstrate that some students' data logs are directly related to different flow experience dimensions, however, our results also demonstrate that the predictive power of these relationships is generally considered low. We advance the literature contributing to the analysis of students' experience in educational systems. In future studies, we aim to replicate the study in different educational systems and, with a larger sample size, increase the power of generalizing the results.

Acknowledgments

The authors would like to thank the grants provided by São Paulo Research Foundation (FAPESP), Projects: 2018/07688-1 and 2020/02801-4 and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Notes

Previous studies of this project have been published: Oliveira *et al*. [20] conducted a systematic literature review about Flow Theory and Educational Technologies; Oliveira [21] presented the project overview; Oliveira *et al*. [18] proposed a theoretical model relating students' data logs and their flow experience in educational systems; Oliveira *et al*. [22] conducted a qualitative study analysing students' data logs and their flow experience in an educational systems; and Oliveira *et al*. [43, 44] conducted data-driven studies modeling and predicting (respectively) students' flow experience based on their data logs in a gamified educational system.

References

[1] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, pp. 44–56, 1 2019.

[2] Y. Shen, P. Yu, H. Lu, X. Zhang, and H. Zeng, "An AI-based virtual simulation experimental teaching system in space engineering education," *Computer Applications in Engineering Education*, p. cae.22221, 4 2020.

[3] A. Sindiani, N. Obeidat, E. Alshdaifat, L. Elsalem, M. Alwani, H. Rawashdeh, A. Fares, T. Alalawne, and L. Tawalbeh, "Distance education during the COVID-19 outbreak: A cross-sectional study among medical students in North of Jordan," *Annals of Medicine and Surgery*, vol. 59, pp. 186–194, 2020.

[4] Ł. Tomczyk and C. Walker, "The emergency (crisis) e-learning as a challenge for teachers in poland," *Education and Information Technologies*, pp. 1–31, 2021.

[5] E. Eryilmaz, B. Thoms, Z. Ahmed, and K.-H. Lee, "Effects of recommendations on message quality and community formation in online conversations," *Education and Information Technologies*, vol. 26, no. 1, pp. 49–68, 2021.

[6] S. Bai, K. F. Hew, and B. Huang, "Is gamification "bullshit"? evidence from a meta-analysis and synthesis of qualitative data in educational contexts," *Educational Research Review*, p. 100322, 2020.

[7] J. Koivisto and J. Hamari, "The rise of motivational information systems: A review of gamification research," *International Journal of Information Management*, vol. 45, pp. 191–210, 2019.

[8] W. Oliveira and I. I. Bittencourt, *Tailored gamification to educational technologies*. Springer, 2019.

[9] K. Tenório, G. C. Challco, D. Dermeval, B. Lemos, P. Nascimento, R. Santos, and A. P. da Silva, "Helping teachers assist their students in gamified adaptive educational systems: Towards a gamification analytics tool," in *International Conference on Artificial Intelligence in Education*, pp. 312–317, Springer, 2020.

[10] J. Hamari, *Gamification*, pp. 1–3. The Blackwell Encyclopedia of Sociology, 2019.

[11] M. Csikszentmihalyi and I. Csikszentmihalyi, *Beyond boredom and anxiety*, vol. 721. Jossey-Bass San Francisco, 1975.

[12] Ş. Ç. Özhan and S. A. Kocadere, "The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment," *Journal of Educational Computing Research*, vol. 57, no. 8, pp. 2006–2031, 2020.

[13] S. Erhel and E. Jamet, "Improving instructions in educational computer games: Exploring the relations between goal specificity, flow experience and learning outcomes," *Computers in Human Behavior*, vol. 91, pp. 106–114, 2019.

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

[15] S. A. Jackson and H. W. Marsh, "Development and validation of a scale to measure optimal experience: The flow state scale," *Journal of sport and exercise psychology*, vol. 18, no. 1, pp. 17–35, 1996.

[16] J. Hamari and J. Koivisto, "Measuring flow in gamification: Dispositional flow scale-2," *Computers in Human Behavior*, vol. 40, pp. 133–143, 2014.

[17] P.-M. Lee, S.-Y. Jheng, and T.-C. Hsiao, "Towards automatically detecting whether student is in flow," in *International Conference on Intelligent Tutoring Systems*, pp. 11–18, Springer, 2014.

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

[19] Y. C. Semerci and D. Goularas, "Evaluation of students' flow state in an e-learning environment through activity and performance using deep learning techniques," *Journal of Educational Computing Research*, p. 0735633120979836, 2020.

[20] W. Oliveira, I. I. Bittencourt, S. Isotani, D. Dermeval, L. B. Marques, and I. F. Silveira, "Flow theory to promote learning in educational systems: Is it really relevant?," *Brazilian Journal of Computers in Education*, vol. 26, no. 02, p. 29, 2018.

[21] W. Oliveira, "Towards automatic flow experience identification in educational systems: A human-computer interaction approach," in *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, pp. 41–46, 2019.

[22] W. Oliveira, L. Rodrigues, A. Toda, P. Palomino, L. Shi, and S. Isotani, "Towards automatic flow experience identification in educational systems: A qualitative study," in *Brazilian Symposium on Computers in Education*, vol. 31, pp. 702–711, 2020.

[23] F. De Kock, *The Neuropsychological Measure (EEG) of Flow Under Conditions of Peak Performance*. Unisa, 2014.

[24] J. Hamari, "Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service," *Electronic Commerce Research and Applications*, vol. 12, no. 4, pp. 236–245, 2013. cited By 362.

[25] D. Johnson, S. Deterding, K.-A. Kuhn, A. Staneva, S. Stoyanov, and L. Hides, "Gamification for health and wellbeing: A systematic review of the literature," *Internet Interventions*, vol. 6, pp. 89–106, 2016.

[26] D. Dicheva, C. Dichev, G. Agre, and G. Angelova, "Gamification in education: A systematic mapping study," *Educational Technology and Society*, vol. 18, no. 3, pp. 75–88, 2015. cited By 595.

[27] A. C. T. Klock, I. Gasparini, M. S. Pimenta, and J. Hamari, "Tailored gamification: A review of literature," *International Journal of Human-Computer Studies*, p. 102495, 2020.

[28] M. Sailer and L. Homner, "The gamification of learning: A meta-analysis," 2020.

[29] M. Csikszentmihalyi and I. S. Csikszentmihalyi, *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press, 1992.

[30] M. Csikszentmihalyi, *Finding flow: The psychology of engagement with everyday life*. Basic Books, 1997.

[31] S. A. Jackson and M. Csikszentmihalyi, *Flow in sports*. Human Kinetics, 1999.

[32] J. Chen, "Flow in games (and everything else)," *Communications of the ACM*, vol. 50, no. 4, pp. 31–34, 2007.

[33] N. Shin, "Online learner's 'flow'experience: an empirical study," *British Journal of Educational Technology*, vol. 37, no. 5, pp. 705–720, 2006.

[34] M. Csikszentmihalyi, "Flow and education," *NAMTA journal*, vol. 22, no. 2, pp. 2–35, 1997.

[35] S. Jackson, B. Eklund, and A. Martin, "The flow manual the manual for the flow scales manual. sampler set," *Mind*, pp. 1–85, 2011.

[36] A. Perttula, K. Kiili, A. Lindstedt, and P. Tuomi, "Flow experience in game based learning–a systematic literature review," *International Journal of Serious Games*, vol. 4, no. 1, pp. 57–72, 2017.

[37] W. Oliveira, I. I. Bittencourt, S. Isotani, D. Dermeval, L. B. Marques, and I. F. Silveira, "Flow theory to promote learning in educational systems: Is it really relevant?," *Brazilian Journal of Computers in Education*, vol. 26, no. 02, p. 29, 2018.

[38] S. A. Jackson and R. C. Eklund, "Assessing flow in physical activity: The flow state scale–2 and dispositional flow scale–2," *Journal of Sport and Exercise Psychology*, vol. 24, no. 2, pp. 133–150, 2002.

[39] S. A. Jackson, S. K. Ford, J. C. Kimiecik, and H. W. Marsh, "Psychological correlates of flow in sport," *Journal of Sport and exercise Psychology*, vol. 20, no. 4, pp. 358–378, 1998.

[40] C.-C. Wang and M.-C. Hsu, "An exploratory study using inexpensive electroencephalography (eeg) to understand flow experience in computer-based instruction," *Information & Management*, vol. 51, no. 7, pp. 912–923, 2014.

[41] D. Akcan, "Flow experience and achievement of advergames on software developers: An eye-tracking analysis," Master's thesis, Çankaya University, 2019.

[42] S.-F. Wu, Y.-L. Lu, and C.-J. Lien, "Detecting students' flow states and their construct through electroencephalogram: Reflective flow experiences, balance of challenge and skill, and sense of control," *Journal of Educational Computing Research*, vol. 58, no. 8, pp. 1515–1540, 2021.

[43] W. Oliveira, O. Pastushenko, J. Hamari, S. Isotani, and T. Hruška, "Modeling students' flow experience through data logs in gamified educational systems," in *2021 IEEE 21st International Conference on Advanced Learning Technologies (ICALT)*, pp. 1–5, IEEE, 2021.

[44] W. Oliveira, K. Tenório, J. Hamari, O. Pastushenko, and S. Isotani, "Predicting students' flow experience through behavior data in gamified educational systems," *Smart Learning Environments*, vol. 8, no. 1, pp. 1–18, 2021.

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

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

[47] R. Likert, "A technique for the measurement of attitudes.," *Archives of psychology*, 1932.

[48] S. S. Gomes, *Evidências psicométricas da escala de predisposição ao fluxo (DFS-2) para brasileiros*. PhD thesis, Universidade São Judas Tadeu, 2014.

[49] R. Orji, G. F. Tondello, and L. E. Nacke, "Personalizing persuasive strategies in gameful systems to gamification user types," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018.

[50] S. Hallifax, A. Serna, J.-C. Marty, G. Lavoué, and E. Lavoué, "Factors to consider for tailored gamification," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 559–572, 2019.

[51] A. C. G. Santos, W. Oliveira, J. Hamari, and S. Isonati, "Do people's user types change over time? an exploratory study," in *Proceedings of the 5nd International GamiFIN Conference, GamiFIN 2021*, CEUR-WS, 2021.

[52] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.

[53] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.

[54] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[55] J. F. Hair Jr, G. T. M. Hult, C. Ringle, and M. Sarstedt, *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications, 2016.

[56] J. F. Hair Jr, M. Sarstedt, C. M. Ringle, and S. P. Gudergan, *Advanced issues in partial least squares structural equation modeling*. saGe publications, 2017.

[57] K. K.-K. Wong, "Partial least squares structural equation modeling (pls-sem) techniques using smartpls," *Marketing Bulletin*, vol. 24, no. 1, pp. 1–32, 2013.

[58] G. C. Challco, F. R. Andrade, S. S. Borges, I. I. Bittencourt, and S. Isotani, "Toward a unified modeling of learner's growth process and flow theory," *Journal of Educational Technology & Society*, vol. 19, no. 2, 2016.