Relations between actions performed by users and their engagement

Ana Paula O. Bertholdo[1], Claudia de O. Melo[2], Artur S. Rozestraten[3], and Marco Aurelio Gerosa[1,4]

[1] Department of Computer Science - University of Sao Paulo, Brazil
ana@ime.usp.br
[2] Department of Computer Science - University of Brasilia, Brazil
claudiam@unb.br
[3] Faculty of Architecture and Urbanism - University of Sao Paulo, Brazil
artur.rozestraten@usp.br
[4] Northern Arizona University, USA
Marco.Gerosa@nau.edu

Abstract. Although Galleries, Libraries, Archives, and Museums (GLAMs) increasingly encourage users to assist in the curation of online collections through open collaboration systems, measuring users' engagement in these systems is a dynamic and complex challenge. We analyzed 18 user's actions over 20 days according to the User Engagement Scale (UES) and based on Maximal Repeating Patterns (MRPs) and correlations between user interaction elements and dimensions of user engagement (focused attention, perceived usability, aesthetics, and reward). Our results show differences in usage tactics for users with high, medium, and low scores from UES, and monotonically increasing moderate correlations between perceived usability scores and game design elements. Additionally, we found that the longer the mean time interval between two consecutive user actions during a usage period lasted, the higher the UES score was. These results help to understand what influences user engagement, isolating the effects of user interaction elements.

Keywords: User engagement · user action · user interaction element · gamification · UES · MRP · Open collaboration community · GLAM.

1 Introduction

Galleries, Libraries, Archives, and Museums (or GLAMs) have been struggling to engage users in the selection, cataloging, contextualization, and curation of collections [20, 26] through crowdsourcing in open collaboration systems [26]. This new mode of interaction surpasses passive access and can lead to a deeper level of engagement with collections [15, 26].[5] Since user participation is key to success in this context [26], GLAMs need to

[5] This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq, proc. 465446/2014-0, CAPES proc. 88887.136422/2017-00, and FAPESP, proc. 2014/50937-1.

create and maintain an open collaboration system that fosters a sense of community around artifacts [25]. Put simply, communities that support open collaboration [9] must engage people [17]. As users become engaged, certain behaviors should increase, such as click frequency [7]. However, with search actions, which are common in GLAMs, there is evidence that the most engaged users have the least amount of search interaction [24] and exhibit more search behaviors when they are frustrated [7, 24, 8]. It is necessary to distinguish users' recurring actions in terms of whether they cause engagement or frustration. Furthermore, Lalmas *et al.* [19] report that in user engagement measurement there is "less emphasis on the role of the task (i.e., what the user is doing), device (desktop versus mobile), and context (e.g., quickly checking something or browsing leisurely), (...) and more work is needed to see how measures from one type of approaches align with that of another one."

Our goal is to analyze the relation between users' recurring actions, based on Maximal Repeating Patterns (MRPs) and the User Engagement Scale (UES), to understand whether recurring actions are related to higher engagement in an online open collaboration community in the context of GLAMs. We begin by introducing background about UES and theories about user attention and task reaction time. Next, we present related work in the context of MRPs and describe the context of the current study, the research questions, and the method. Lastly, we show our results, summarize our findings, and discuss the limitations and implications of this work.

2 Background and Related Work

User engagement is "a quality of user experience characterized by the depth of an actor's investment when interacting with a digital system" [27]. O'Brien & Toms [25] consider engagement as a process with four distinct stages: point of engagement, period of sustained engagement, disengagement, and reengagement. This process is characterized by attributes of engagement from the user, system, and user-system interaction.

O'Brien *et al.* define the **User Engagement Scale (UES**) [28, 25] as a tool to measure user engagement. The original UES consists of 31-items to measure six dimensions of engagement. Recent research from O'Brien *et al.* [28] proposed a shorter version with 12 items to measure a four-factor structure: focused attention (FA), perceived usability (PU), aesthetic appeal (AE), and reward (RW). UES can be analyzed by subscales or dimensions or aggregated as an overall engagement score.

Participants' scores on each of the UES subscales can be calculated by summing individual responses to items within each subscale and dividing this total by the number of items. Total scores for the UES are calculated by adding the averages of each subscale and dividing by the number of subscales. The scores can be then divided according to percentiles to create three groups (low, medium, and high) based on the median [24]. Table 1 summarizes the self-reported engagement metrics according to the shorter version of the UES [28].

Table 1. Self-reported engagement metrics from the shorter version of the UES[28].

Engagement Metric	Description
Focused attention (FA)	Refers to feelings absorbed in the interaction and losing track of time
Perceived usability (PU)	Refers to negative effect experienced as a result of the interaction and the degree of control and effort expended
Aesthetic appeal (AE)	Refers to the attractiveness and visual appeal of the interface
Reward (RW)	Refers to endurability (or the overall success of the interaction), novelty, and felt involvement
UES total score	Overall self-reported engagement score

User actions relate to user interaction elements. Categorizing user actions according to accessed user interaction elements can help us to understand user engagement. According to Harnad [12], "cognition is categorization;" consequently, "assigning terms to categories plays a major role in communication" [38]. There are many user interaction elements according to the domain of a software system. We focus on user interaction elements related to online open collaboration communities, specifically collaborative and functional elements. Since gamification
- the use of game design elements in non-game contexts [5] - has been defined as a way to foster greater user engagement in online communities [2], our study analyzes game design elements [5]. Table 2 provides a summary of the definition of each user interaction element we address in this study.

Table 2. Definition of each user interaction element addressed in this study.

User interaction elements	Definition
Game design elements (gamification)	Elements that belongs to one of the five levels of the game design elements [5].
Collaborative elements (collaboration)	Elements that give support for collaboration [10, 11].
Functional elements	Elements related to functional requirements [29].

As defined by Deterding *et al.* [5], game design elements consider: 1) game interface design patterns, 2) game design patterns and mechanics, 3) game design principles and heuristics, 4) game models, and 5) game design methods. As argued by Fuks *et al.* [10], collaborative elements can be related to one of the following dimensions: 1) communication, 2) coordination, and 3) cooperation (3C collaboration Model). Finally, a functional requirement specifies functions that systems must be able to perform [29]. In our study, game design and collaborative elements are not defined by functional elements, although they have functional requirements. For game design elements, this study only considers elements from "game interface design patterns."

2.1 User attention, reaction time to tasks, and usage tactics

According to Manly & Robertson [21], "action doesn't necessarily stop when our mind is elsewhere." Theorists have investigated people's attentional lapse over decades [21, 22, 30]. The Sustained Attention to Response Test (SART) [21] was designed to measure attentional lapses; it is a laboratory test in which participants view a computer monitor and are tasked to press a response key after each presentation, except for a "no-go digit," to which no response should be performed. Performance on SART was predictive of action slips and everyday attentional failures in participants [21].

Smallwood *et al.* [33] reported task engagement and disengagement during the SART. They performed experiments to investigate the relationship between subjective experience and attention lapses. The results suggest that during sustained attention people experience task unrelated thought (TUT), which corresponds to an absent-minded disengagement from the task [33, 34]. TUT [34] and attentional lapses are attributed to situations of boredom and worry [30, 33]. In the context of SART and under conditions of low target probability, shorter reaction times were related to more significant distraction and insensitivity to the task. Robertson *et al.* [30] support this claim when reporting that the "oops" phe- nomenon associated with errors suggests that error detection tends to redirect attention towards the task, resulting in slower emergence of the alternative or correct answer.

On a task that is repetitive in nature, an engaged user's thinking usually strays from information visible in the current environment, and in this context, the user's attention becomes more focused on the task [33]. Performance is crucially determined by the duration of time over which attention must be maintained on one's actions [22]. Maximal Repeating Patterns (MRPs) are used to extract recurring user action patterns (or usage tactics) from a user session transcript [13]. MRPs identify the longest string of actions that is repeated at least once in the entire data set [32]. According to Siochi [32], a repeating pattern is a substring that occurs at more than one position in a string. Substrings of longer patterns are also considered MRPs if they occur independently. MRPs may also overlap, as in "abcabcabc," where "abcabc" is the MRP.

In the context of a video retrieval system, Wildemuth *et al.* [37] collected users' search tactics to understand how people search databases for videos, and how the medium of the object can influence the search behaviors. Each search move was coded, and the data was examined for MRPs. Tactics were mainly characterized by the addition of concepts, and frequent display and browsing of the search results. A previous study by Wildemuth [36] found that the search tactics changed over time as the participants' domain knowledge changed. Edwards and Kelly [7] examined the differences in the search behaviors and physiological characteristics of people who were engaged or frustrated during searches. Users engaged more with search results for tasks they found interesting. Their results support the idea that task interest is an important component of engagement. The authors demonstrated that increased search behavior was a stronger indicator of frustration than interest.

In the context of usability evaluations, analyzing maximum repeating patterns might reveal interesting information on user interface usability [13]. Several researchers have used MRPs in the context of search tactics [36, 37, 3]. A search tactic comprises several individual moves, where a move is a single step in executing a search tactic, such as "deleting a concept to increase the size of the result set" [37]. In this study, we applied MRPs to examine users' usage tactics. Compared to search tactics, **usage tactics** comprise one or more individual moves, which involve **user recurring actions** in general. Additionally, we analyzed correlations between user actions and user engagement collected through users' self-reports according to the UES scale.

3 Research Method

This study was conducted in the context of the Arquigrafia online community. Arquigrafia is a public, nonprofit digital collaborative community dedicated to disseminating architectural images, with special attention to Brazilian architecture (www.arquigrafia.org.br). The main objective of the community is to contribute to the teaching, research, and diffusion of architectural and urban culture by promoting collaborative interactions among people and institutions. Arquigrafia is still small and needs to foster a community around architecture images and information. Participants include professional architects, architecture and urbanism students, architecture and urbanism professors, librarians, library science students, and professional photographers; whose ages range between 20-68 years old; and of which 61.11% are male (11 out of 18).

We designed this study as a correlational research, which aims to discover the existence of a relationship, association, or interdependence between two or more aspects of a situation [18]. This study comprises an analysis from another point of view for an experiment described in [1]. In the current study, our goal was to understand the relationship between users' actions and their engagement for 18 users that used the system for 20 days (November 16, 2017 to December 5, 2017) and answered an online questionnaire based on the **User Engagement Scale (UES)** [25, 28], described in Section 2.

The online questionnaire was applied according to a UES-translated version for Portuguese. Users actions were collected from logs inserted in the system or user session transcripts, which represent the time-ordered sequence of actions users performed [32]. We then investigated what user actions and usage tactics - one or more user recurring actions - contributed to high, medium, and low user engagement, as defined in Section 2. We also analyzed whether there are correlations between the actions performed by each group and their score in the UES scale. We decomposed user actions according to user interaction elements, defined in Section 2.

Therefore, each user action was classified by the number of collaborative, functional, and game design elements accessed by 18 users when performing the actions. This number was correlated to UES. We analyzed reaction time [22, 33, 30] in the context of Arquigrafia as the interval between two consecutive actions, since the user performed an action, received a response from the system, and performed another action as a reaction to system response. In this context, we answer the following research questions:

RQI How do different usage tactics relate to high, medium, and low user engagement?
RQII How do users' level of engagement relate to their reaction times?
RQIII Are the user interaction elements correlated to users' engagement?

3.1 Data Preparation and Analysis

To evaluate the reliability of the UES, we calculated Cronbach's alpha. The goal was to examine the internal consistency of subscales based on DeVellis's guidelines (0.7-0.9 is optimal) [6]. As shown in Table 3, the UES was highly reliable. Initially, 12 items were considered in the UES. After the analysis of Cronbach's alpha, one item was dismissed to improve the value of Cronbach's alpha for the AE (from 0.65 to 0.84) subscale.

Table 3. Cronbach's alpha and descriptive statistics of UES.

UES Subscales	N items	Mean	Std	alpha
Focused attention (FA)	3	2.7	1.2	0.95
Perceived usability (PU)	3	3.7	0.89	0.86
Aesthetic appeal (AE)	2	3.6	0.61	0.84
Reward (RW)	3	4	0.83	0.91
Total Engagement	11	3.5	0.74	0.92

Table 4 presents users' actions classified according to functional, game design, and collaborative elements. The research was applied in a naturalistic setting. Although there are other user interaction elements available in the system, the focus was on the elements accessed by users. Therefore, the number of each element described in Table 4 refers to how many user interaction elements of each type were effectively involved in user interaction with the system. The number of collaborative and game design elements accessed was much smaller than the number of functional elements accessed, which demonstrates the challenge of engaging users to collaborate in GLAMs, in which users consume information rather than collaborate to produce it.

Table 4. Users actions classified by the user interaction element (functional, game design, or collaborative).

User interaction elements	ACTION	High n	Medium n	Low n
Functional elements	Home page	7	18	33
	Image Searching	27	80	6
	Download	82	36	0
	User Selection	4	5	5
	User Edition	1	1	1
	Access to similar evaluations	0	0	2
	Image Selection (institution)	126	62	6
	Image Selection (no Progress Bar)	0	1	15
	Logout	1	4	4
Collaborative elements	Image Evaluation	0	0	2
	User Following	0	1	1
	Chat	0	1	0
	Access to Upload	1	0	0
	Access to Chat page	0	0	1
	Access to public Albuns creation	0	1	0
	Sharing by Facebook	0	0	1
	Sharing by Google+	3	0	0
	Access to Data Review (no Points)	0	0	3
	Completed Data Review	0	0	1
Game design elements	Image Selection (Progress Bar view)	4	1	5
	Access to Data Review (Points view)	0	0	1
TOTAL	Functional elements	248 (96.88%)	207 (98.1%)	72 (82.76%)
	Collaborative elements	4 (1.56%)	3 (1.42%)	9 (10.34%)
	Game design elements	4 (1.56%)	1 (0.47%)	6 (6.9%)
	All user interaction elements accessed	256	211	87

The analysis of MRPs was performed by manually looking for the MRPs in each group of users according to UES scores (low, medium, and high), and by comparing usage tactics based on MRPs from different user groups. The Shapiro-Wilk normality test rejected the hypothesis that UES scores and amount of user interaction elements for each user come from a normal population. Therefore, we used the Spearman's rank correlation coefficient to measure the strength of the relationship between the number of each user interaction element and each one of 4 subscales, as well as the overall score from UES; the Kruskal-Wallis rank sum test was used to compare user interaction elements from high-UES-score, medium-UES-score, and low-UES-score groups. To define an effect size for this comparison, the Epsilon-squared was measured. We used guidelines to interpret Spearman's correlation coefficient [23, 14] and effect sizes [4, 16].

4 Results

4.1 How do different usage tactics relate to high, medium, and low user engagement? (RQI)

Out of the 18 participants, 6, 7, and 5 participants are in low-UES-score, medium-UES-score, and high-UES-score groups, respectively. High-UES-score users had scores above 4.02; medium-UES-score users had scores between 3.27 and 4.02, and low-UES-score users had scores below 3.27. The minimum overall UES score was 1.72, and the maximum score was 4.45. The 5-item Likert-scale score was used, so the highest overall score for any subscale was 5. Figure 1 presents the main usage tactics based on MRPs during the analyzed period for each group.

User recurring actions were represented by action name multiplied by (*) the number of consecutive invocations. The number of MRPs for each group was represented by n(High), n(Medium), and n(Low). For example, the MRP "A" comprises three consecutive invocations of the action download, (i.e., download * 3), and the number of MRPs "A" found for high-UES-score users is 21, and for medium and low UES-score users, the MRP "A" was not found.

For high score users, there was a predominance of single recurring actions, especially actions of image selection and download from the institutional collection. This behavior occurred because users first selected images and then downloaded them. For medium score users, there was a predominance of the usage tactic of image search, image selection from the institutional collection, and download, and single recurring actions of image search and image selection that were not followed by downloading. For low score users, the usage tactics of accessing the homepage and selecting an image from the institutional collection were most recurrent.

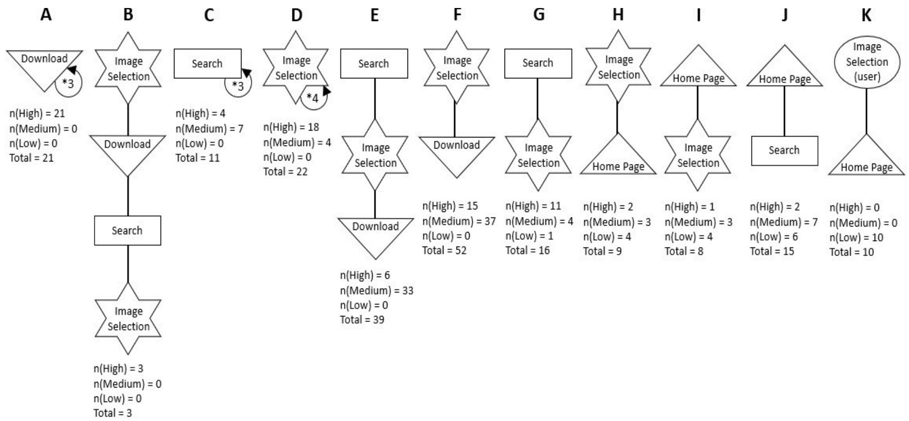

Fig. 1. Main usage tactics during the analyzed period based on MRPs. The representation of tactics is based on Wildemuth *et al.* [37].

Figure 1 distinguishes image selection and image selection (user); the first is the image selection from the institutional collection, and the second is the image selection from private users' collections. The last can be presented with or without progress bar view. The progress bar is a game design element intended to help users check the completeness of the image data, according to a previous experiment described in [1]. For low-UES-score users, the action of image selection (without progress bar view) was the second most accessed, after only access to the homepage. These actions are presented in usage tactic K. In summary, usage tactics differed among groups of users with high, medium, and low user engagement, but involved the same group of actions: accessing the homepage, image searching, downloading, and image selecting. These actions are related to functional elements.

4.2 How do users' level of engagement relate to their reaction time? (RQII)

Consecutive actions from low-UES-score users occurred in an interval between 0 and 49 seconds, with one user (outlier) reaching 8 minutes and 17 seconds. The mean was 28 seconds with a standard deviation of 35 seconds. For medium-UES-score users, consecutive actions occurred in an interval between 0 and 6 minutes and 49 seconds, with one user reaching 8 minutes. The mean was 1 minute and 15 seconds with a standard deviation of 2 minutes and 06 seconds. For high-UES-score users, consecutive actions occurred in an interval between 0 and 37 minutes and 18 seconds. The mean was 7 minutes and 37 seconds with a standard deviation of 12 minutes and 58 seconds. Therefore, in this study with 18 users, the longer the mean time interval between two consecutive actions (i.e., user reaction time) during a usage period, the higher the user's UES score.

4.3 Are the user interaction elements correlated to users' engagement? (RQIII)

Table 5 presents the correlation between the number of user interaction elements accessed by 18 users and the user's engagement according to Spearman's rank correlation rho. Table 6 presents the same correlation classified by high, medium, and low scores based on UES. The main results can be summarized as follows:

1. **A monotonically increasing moderate correlation between the number of game design elements in general and the perceived usability (PU) score** (see Table 5). The PU scores varied between 1.66 and 5 for users who have not accessed game design elements (mean 3.48, sd 0.49); whereas PU scores varied between 3.66 and 5 for users who accessed game design elements (mean 4.39, sd 0.88). Ten actions (90.9%) were related to the progress bar element, and only one action was related to the points view.

Table 5. Correlations between user interaction elements and the user's engagement (Spearman's rank correlation rho).

UES score	Functional - rho (p)	Collaborative - rho (p)	Game design - rho (p)
Overall	0.17 (0.4993)	-0.02 (0.9185)	0.28 (0.2581)
FA	0.20 (0.403)	-0.16 (0.5024)	0.18 (0.4531)
PU	-0.009 (0.9687)	0.30 (0.2125)	**0.55 (0.01733)**
AE	-0.025 (0.9203)	0.28 (0.2494)	0.016 (0.9483)
RW	0.29 (0.234)	0.15 (0.526)	0.37 (0.1199)

Table 6. Correlations between user interaction elements and the user's engagement classified by High, Medium, and Low scores (Spearman's rank correlation rho).

User interaction elements	Score	UES rho (p)	FA rho (p)	PU rho (p)	AE rho (p)	RW rho (p)
Functional	High	0.44 (0.4502)	0.3 (0.6833)	-0.66 (0.2189)	0.35 (0.5594)	0 (1)
	Medium	-0.19 (0.6701)	0.13 (0.7752)	-0.30 (0.5007)	**-0.79 (0.03432)**	0.31 (0.4869)
	Low	0.04 (0.9339)	-0.12 (0.8131)	0.08 (0.8699)	-0.31 (0.5452)	0.36 (0.4734)
Collaborative	High	0.12 (0.8413)	-0.44 (0.4502)	0.34 (0.5707)	0.39 (0.5101)	0.39 (0.5101)
	Medium	-0.15 (0.7362)	0 (1)	-0.09 (0.8374)	0.39 (0.3813)	-0.42 (0.3481)
	Low	0.73 (0.09312)	0.09 (0.8529)	0.75 (0.08014)	0.37 (0.4636)	**0.95 (0.00301)**
Game Design	High	0.17 (0.7761)	0.21 (0.7336)	0.10 (0.8626)	-0.18 (0.7641)	-0.18 (0.7641)
	Medium	-0.61 (0.1392)	**-0.76 (0.04566)**	0 (1)	-0.64 (0.1174)	0 (1)
	Low	0.53 (0.278)	-0.42 (0.4018)	0.65 (0.1583)	0.53 (0.2694)	0.39 (0.4339)

2. **A monotonically decreasing strong correlation between the number of Functional elements from medium score users and the aesthetic appeal (AE) score** (see Table 6). From 7 users with medium scores from overall UES, 5 users who accessed few functional elements (at most n=9) classified with 4 as their AE score; and 2 users that accessed 48 and 134 functional elements, respectively, classified with 3.5 as their AE score. For both users, there were few downloads when compared to the number of image searches and selections. This result indicates that search results were unsatisfactory or that they faced difficulties with the search process.

3. **A monotonically increasing very strong correlation between the number of collaborative elements from low score users and the reward (RW) score** (see Table 6). From 6 users with low score from

overall UES, highest RW scores users accessed collaborative elements (with RW scores of 3.33, 4, and 4.33). For users who did not access collaborative elements, scores varied between 2 and 3.

4. **A monotonically decreasing strong correlation between the number of game design elements from medium score users and the focused attention (FA) score** (see Table 6). Out of the 7 users with medium score from overall UES, the only user who accessed game design elements was the one with the lower score for FA (2.33).

Figure 2 presents results for the Kruskal Test among functional, collaborative, and game design elements, respectively, classified by high, low, and medium scores from overall UES. Although Figure 2 does not present statistically significant results, it helps to understand that the number of users who accessed each user interaction element differed between groups.

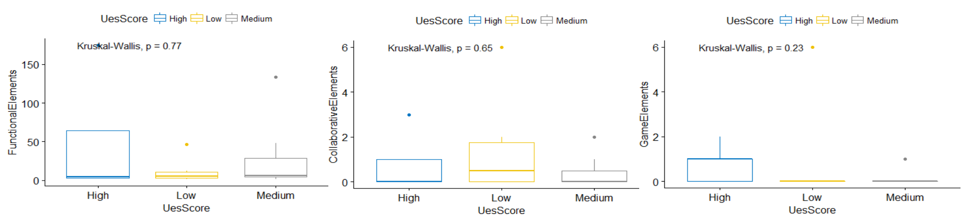

Fig. 2. Kruskal Test for Functional, Collaborative, and Game design elements, comparing High, Medium, and Low scores from UES.

Table 7. Effect sizes for Functional, Collaborative, and Game design elements compared according to High, Medium, and Low scores from UES.

User interaction elements	Kruskal-Wallis chi-squared	Epsilon-squared
Functional	0.51182 (p=0.77)	0.0301
Collaborative	0.84853 (p=0.65)	0.0499
Game design	2.9423 (p=0.23)	0.173

Table 7 presents Epsilon-squared effect sizes for the Kruskal Test with high, medium, and low scores from overall UES for each user interaction element. Epsilon-squared presented relations of 0.03, 0.04, and 0.17 in the UES overall score for functional, collaborative, and game design elements, respectively. Only game design elements present large effect sizes, according to the guidelines [4, 16]. Functional and collaborative elements show small effect sizes. However, the Kruskal Test did not present statistically significant results, which indicates the need to analyze a higher sample size because there are indications of strong effects between the most-accessed game design element (Progress Bar) and the user's UES score.

5 Discussion

5.1 Main MRPs, reaction times and their meanings

Siochi [32] classified Type 1 MRPs, i.e., consecutive invocations of the same command, as a behavior that may indicate that a user needs to perform the same command on several objects, or that the user is ”fine-tuning” a single object. In line with Siochi’s studies, we found MRPs consisting of consecutive invocations of the same action or Type 1 MRP. Main consecutive actions were: image searching, downloading, and image selection from the institutional collection. Information consumption behavior (passive participation) is the most common behavior in the context of GLAMs.

A sequence of image searches indicates that users reformulated their keywords several times before a relevant result appeared. A series of downloads that occurred after a sequence of actions of image selection from the institutional collection indicates that users opened many tabs, one for each selected image, and then returned to each image for their analyses and the decision to download images. Only for high-UES-score users were there consecutive sequences of actions of download after actions of image selection from the institutional collection.

However, a sequence of image searches occurred for high and medium-UES-score users, or a possible difficulty to find an object. Finding an interesting image to download after fine-tuning produced a positive impression, which is presented by users for high-UES-score. The opposite occurred with medium-UES-score users, for whom fine-tuning resulted in downloading few of the selected images. This result aligns with previous studies [7], wherein participants presented with poor result quality submitted more queries and saved fewer documents.

Edwards & Kelly [7] reported a number of significant differences in participant search behaviors based on the quality of search results. Studies from Smallwood *et al.* [33], described in Section 2.1, can explain why Type 1 MRPs were not found for low-UES-score users. Consecutive invocations of the same MRP represent a level of attention from users to maintain the same response.

Additionally, most MRPs from low-UES-score users involved accessing the homepage. This behavior can indicate that: 1) users in this group did not understand how the system worked, 2) they were exploring the system without a specific goal, or 3) they were concurrently executing other tasks. The first case can indicate a usability problem, while the second can indicate a lack of prior motivation to use the system, in which users tried to understand what value the system might have for them. The third case can indicate Task Unrelated Thought (TUT) [34]. TUT and attentional lapses are attributed to situations of boredom and worry [30, 33]. This may be why low-UES-score users self-reported their engagement as low and why users in this group overlooked other options for executing actions, which led them to return to the homepage.

The mean time between two consecutive actions differed among groups. Although there were users in all groups with similar time intervals between successive actions (i.e., reaction time), significant differences appeared in the mean time between actions. We found as the mean time interval between two consecutive actions during a usage period increased, so did the higher the user’s score in the UES. This behavior is in accordance to the literature [30, 33].

In the context of Arquigrafia, we did not find periods of high search activity for low-UES-score users. This behavior was found only for high and medium-

UES-score users. Between these two, high-UES-score users presented the lowest browsing times when compared to medium-UES-score users. However, the time of each user session and the reaction time were higher for high-UES-score users. This behavior was in line with results from the literature [22, 33].

5.2 Relations among user interaction element and the UES

Spearman's rank correlation presented statistically significant results, varying from moderate to very strong correlations. By decomposing the accessed user interaction elements and correlating them with UES, there is a statistically significant positive relationship between PU and the number of game elements accessed. This behavior may have occurred because the progress bar only presents the information and does not allow the user to perform an active action. Therefore, it is worth distinguishing between elements that can passively influence users' actions and those that allow them to act.

The monotonically decreasing strong correlation between the number of functional elements from Medium score users and the AE score could be explained by the frustration in locating images in the desired subject, which may have led the user to have a lower aesthetic quality impression in this group. The monotonically increasing very strong correlation between the number of collaborative elements from low score users and the RW score could indicate that collaborative elements did not influence the lowest UES score.

5.3 Limitations

The limitations of this study derive from the number of users (n=18) who answered the questionnaire, which resulted in 5 High-UES-Score users, 7 Medium-UES-Score users, and 6 Low-UES-Score users. However, the study was performed in a naturalistic configuration in a system with engagement problems, allowing for evaluation of actual use rather than a simulated laboratory environment. The study was short-term and one-off in nature [31]. One-off studies need to be replicated and comparative and longitudinal designs employed to draw stronger, more generalizable conclusions.

5.4 Implications for future studies

To understand the effects of a system intervention, it is necessary to evaluate which actions users accessed and how they correlate with subjective and objective engagement measures. The effects of one user interaction element can be compared to the effects of others. It is necessary to compare effects obtained during access to a given element and to a set of elements to evaluate whether the element alone or the interaction of a set of elements produced greater engagement. In this case, what are the involved user interaction elements?

By classifying user interaction elements as game design or collaborative elements, we intended to evaluate, respectively, the effect of gamification or collaboration on user engagement. In future studies, the analysis of user actions can be decomposed, for example, by correlating between characteristics from collaboration - or user interaction elements related to communication, coordination, and cooperation - and each dimension of user engagement. The same behavior can be performed for game design elements and functional elements, i.e., decomposing them to specific characteristics of each user interaction element. The benefit of this decomposition is that it isolates the effects of a system intervention.

Frustration is a function not only of the current interaction but of the previous state of frustration [8]. This behavior may imply that users who experienced frustration in a system during previous access are prone to a frustrating experience in new accesses. Future studies can analyze the correlation between each user action and each dimension of engagement, distinguishing actions from newcomers and existing users. For the latter, researchers can compare actions from previous access to current access to better understand users' reported engagement or frustration. From these results, researchers can design appropriate support for newcomers [35].

6 Conclusions

Our findings revealed that the recurring user actions set combined with the mean reaction time can inform user engagement more than the frequency of actions alone. User actions can be active, such as performing image evaluations or uploads, or passive, such as performing searches or downloads. Engaged users contribute to the system community (active access) and/or to meeting their own goals (passive access). Both are guided by a specific goal that implies a pre-existing motivation, reinforcing the idea that goal-oriented motivation guides engagement more than the exploitation of the system without specific goals, which may explain why a high amount of repeated actions do not indicate that the users were engaged when they performed them.

Additionally, results show differences in usage tactics and in the distribution of access to each user interaction element for users with high, medium, and low UES scores. Functional elements contributed to greater engagement for users with high UES scores, and monotonically increasing moderate correlations were found between perceived usability score and game design elements. Our results are useful for further analyses of user actions to better understand what determines user engagement and to isolate the effects of each user interaction element.

Acknowledgments

We would like to thank to Dr. Heather L. OBrien from iSchool, University of British Columbia, for her valuable inputs on this study.

References

1. Bertholdo, A.P.O., Melo, C.d.O., Rozestraten, A.S., Gerosa, M.A., O'Brien, H.L.: User engagement in an open collaboration community after the insertion of a game design element: An online field experiment. In: Proceeding of the 24th Americas Conference on Information Systems (AMCIS 2018) (2018)
2. Bista, S.K., Nepal, S., Colineau, N., Paris, C.: Using gamification in an online community. In: Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on. pp. 611–618. IEEE (2012)
3. Bron, M., Van Gorp, J., Nack, F., de Rijke, M., Vishneuski, A., de Leeuw, S.: A subjunctive exploratory search interface to support media studies researchers. In: Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval. pp. 425–434. ACM (2012)
4. Cohen, J.: Statistical power analysis for the behavioral sciences 2nd edn (1988)

5. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: Proc. of the 15th international academic MindTrek conference: Envisioning future media environments. pp. 9–15. ACM (2011)
6. DeVellis, R.F.: Scale development: Theory and applications (2003)
7. Edwards, A., Kelly, D.: Engaged or frustrated?: Disambiguating emotional state in search. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 125–134. ACM (2017)
8. Feild, H., Allan, J.: Modeling searcher frustration. Proceedings from HCIR (2009)
9. Forte, A., Lampe, C.: Defining, Understanding, and Supporting Open Collaboration: Lessons From the Literature. AMERICAN BEHAVIORAL SCIENTIST **57**(5, SI), 535–547 (MAY 2013)
10. Fuks, H., Raposo, A., Gerosa, M.A., et al.: The 3c collaboration model. In: Encyclopedia of E-collaboration, pp. 637–644. IGI Global (2008)
11. Gerosa, M.A., Fuks, H., Lucena, C.: Analysis and design of awareness elements in collaborative digital environments: A case study in the aulanet learning environment. Journal of Interactive Learning Research **14**(3), 315 (2003)
12. Harnad, S.: To cognize is to categorize: Cognition is categorization. In: Handbook of Categorization in Cognitive Science (Second Edition), pp. 21–54. Elsevier (2017)
13. Hilbert, D.M., Redmiles, D.F.: Extracting usability information from user interface events. ACM Computing Surveys (CSUR) **32**(4), 384–421 (2000)
14. Hinkle, D.E., Wiersma, W., Jurs, S.G.: Applied statistics for the behavioral sciences, vol. 663. Houghton Mifflin College Division (2003)
15. Huvila, I.: Participatory archive: towards decentralised curation, radical user orientation, and broader contextualisation of records management. Archival Science **8**(1), 15–36 (2008)
16. Keppel, G., Wickens, T.: Effect size, power, and sample size. Design and Analysis. A Researchers Handbook, ed **4**, 159–801 (2004)
17. Kraut, R.E., Resnick, P.: Encouraging contribution to online communities. Building successful online communities: Evidence-based social design pp. 21–76 (2011)
18. Kumar, R.: Research methodologies: a step-by-step guide for beginners. 3rd (2011)
19. Lalmas, M., O'Brien, H., Yom-Tov, E.: Measuring user engagement. Synthesis Lec- tures on Information Concepts, Retrieval, and Services **6**(4), 1–132 (2014)
20. Lankes, R.D., Silverstein, J., Nicholson, S.: Participatory networks: the library as conversation. Information technology and libraries **26**(4), 17 (2007)
21. Manly, T., Robertson, I.H.: The sustained attention to response test (sart). In: Neurobiology of attention, pp. 337–338. Elsevier (2005)
22. Manly, T., Robertson, I.H., Galloway, M., Hawkins, K.: The absent mind:: further investigations of sustained attention to response. Neuropsychologia **37**(6), 661–670 (1999)
23. Mukaka, M.M.: A guide to appropriate use of correlation coefficient in medical research. Malawi Medical Journal **24**(3), 69–71 (2012)
24. O'Brien, H.L., Lebow, M.: Mixed-methods approach to measuring user experience in online news interactions. Journal of the Association for Information Science and Technology **64**(8), 1543–1556 (2013)
25. O'Brien, H.L., Toms, E.G.: What is user engagement? a conceptual framework for defining user engagement with technology. Journal of the American Society for Information Science and Technology **59**(6), 938–955 (2008)
26. Oomen, J., Aroyo, L.: Crowdsourcing in the cultural heritage domain: opportunities and challenges. In: Proceedings of the 5th International Conference on Communi- ties and Technologies. pp. 138–149. ACM (2011)
27. OBrien, H.: Theoretical perspectives on user engagement. In: Why Engagement Matters, pp. 1–26. Springer (2016)
28. OBrien, H., Cairns, P., Hall, M.: A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. International Journal of Human-Computer Studies (2018)
29. Radatz, J., Geraci, A., Katki, F.: Ieee standard glossary of software engineering terminology. IEEE Std **610121990**(121990), 3 (1990)

30. Robertson, I.H., Manly, T., Andrade, J., Baddeley, B.T., Yiend, J.: Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia **35**(6), 747–758 (1997)
31. Seaborn, K., Fels, D.I.: Gamification in theory and action: A survey. International Journal of human-computer studies **74**, 14–31 (2015)
32. Siochi, A.C., Ehrich, R.W.: Computer analysis of user interfaces based on repetition in transcripts of user sessions. ACM Transactions on Information Systems (TOIS) **9**(4), 309–335 (1991)
33. Smallwood, J., Davies, J.B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., Obonsawin, M.: Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. Consciousness and cognition **13**(4), 657–690 (2004)
34. Smallwood, J., Obonsawin, M., Heim, D.: Task unrelated thought: The role of distributed processing. Consciousness and cognition **12**(2), 169–189 (2003)
35. Steinmacher, I., Conte, T.U., Treude, C., Gerosa, M.A.: Overcoming open source project entry barriers with a portal for newcomers. In: Proceedings of the 38th International Conference on Software Engineering. pp. 273–284. ACM (2016)
36. Wildemuth, B.M.: The effects of domain knowledge on search tactic formulation. Journal of the Association for Information Science and Technology **55**(3), 246–258 (2004)
37. Wildemuth, B.M., Oh, J.S., Marchionini, G.: Tactics used when searching for digital video. In: Proceedings of the third symposium on Information interaction in context. pp. 255–264. ACM (2010)
38. Ye, M., Janowicz, K., Mu¨lligann, C., Lee, W.C.: What you are is when you are: the temporal dimension of feature types in location-based social networks. In: Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. pp. 102–111. ACM (2011)