

Interaction and Reflection with Quantified Self and Gamification: an Experimental Study

BENEDIKT S. MORSCHHEUSER

E-Mail: Benedikt.Morschheuser@uni-leipzig.de

Current position: Research Scientist at University of Leipzig

VERÓNICA RIVERA-PELAYO

E-Mail: rivera@fzi.de

Current position: Research scientist at FZI Forschungszentrum Informatik

ATHANASIOS MAZARAKIS

E-Mail: mazarakis@fzi.de

Current position: Research scientist at FZI Forschungszentrum Informatik

VALENTIN ZACHARIAS

E-Mail: zach@fzi.de

Current position: Division Manager at FZI Forschungszentrum Informatik

Abstract

In this paper, we present our research on the impact of gamification – ‘the use of game design elements in non-game contexts’ – to increase the motivation of students to use PLEs (Personal Learning Environments) that enhance interaction and support reflection in lectures. To examine this, we conducted an experiment with the Live-Interest-Meter (LIM), a Quantified Self (QS) application which allows capturing, sharing and visualizing several types of feedback with the aim of improving the learning experience during and after lectures. The results show that perceived fun has a positive effect on the motivation to use the LIM and the motivation to use the application with gamification is significantly higher than for the application without it. Therefore, gamification seems to be an appropriate enabler to engage people in using QS approaches as PLEs for improving their learning experiences.

Introduction

Recently there has been a growing interest in the impact of gamification, i.e. ‘the use of game design elements in non-game contexts’ (Deterding et al., 2011), on motivation in several contexts, including business and education (e.g. Thom et al., 2012; Lee & Hammer, 2011). In the context of learning, gamification may contribute to increase the motivation of students to use tools for optimizing their personal learning environment (PLE) and knowledge about their own learning behaviour in the future⁴.

In the case of informal learning, several tools have been developed and tested within the EU-Project MIRROR – Reflective Learning at Work⁵ –, to support reflective learning in

4 This paper is an extension of work originally reported in the PLE Conference 2013 by B. S. Morschheuser, V. Rivera-Pelayo, A. Mazarakis, V. Zacharias (2013). Gamifying Quantified Self Approaches for Learning: An experiment with the Live Interest Meter. *In Proceedings of the PLE Conference 2013*. (in press)

5 ¹ <http://www.mirror-project.eu>

work environments. Some of these tools are self-tracking applications, also known as Quantified Self (QS) tools, i.e. tools that collect personally relevant information to gain self-knowledge about one's own behaviours, habits and thoughts⁶. Rivera-Pelayo et al. (2012) defined an integrated model which describes how QS tools can support reflective learning. This support is divided in three possible dimensions, namely (i) tracking cues, (ii) triggering reflection and (iii) recalling and revisiting experiences, with the aim of guiding learners to achieve their desired outcomes. One of these MIRROR QS tools is the Live-Interest-Meter (LIM) (Rivera-Pelayo et al., 2013), which allows capturing, aggregating and visualizing feedback given to the lecturer with the aim of improving interaction and supporting reflective learning for both speaker and audience. In this concrete scenario, the desired reflection outcomes may be improvements of the presenter's skills and performance when addressing an audience. In order to achieve this, the LIM implements the three dimensions defined in the integrated model.

According to Warschauer (2006), 'the intersection between interaction and reflection is of critical importance in education'. In his article, Warschauer discusses the relationship of technology to literacy by focusing on computer-mediated texts, but the importance of enhancing the interaction and reflection is also applicable to data in other formats (Warschauer, 2006).

This is the case of the LIM, which aims at capturing several interaction aspects during the lecture (e.g. feedback), making this data available retrospectively, and thereby improving reflective practices. Both students and lecturers can benefit from the LIM, as students are provided with an easy means to assess their learning (e.g. understanding of the lecture's

6 ¹ <http://www.quantifiedself.com>

content) and lecturers can gain insights about their performance and the students' perspective, by analysing the captured feedback. The LIM has a uni-dimensional configurable meter that allows students to track their feedback and evaluate the lecture (e.g. speed of speech or comprehension). During a lecture and afterwards, the individual learning experience is supported by the visualization of this tracked feedback data, which is also aggregated and analysed, and by peer comparison. Therefore, this contributes to the learners own learning process. Studies with the LIM and other MIRROR apps have shown that the use of such QS reflective learning applications in educational and working contexts faces a lack of motivation. Concretely, the results of several studies conducted with the LIM showed concerns regarding the students' voluntary participation to give feedback and actively be involved in the lectures (Rivera-Pelayo et al., 2013).

With the goal to enhance the user's engagement, we examined whether gamification can increase motivation to use QS tools like the Live-Interest-Meter. We conducted an extended literature review on gamification, the QS community and learning through reflection in order to create a theoretical framework. We also analysed successes and failures of existing gamification approaches. Following, we conducted an experiment to analyse the users' intention to use an adapted version of the Live-Interest-Meter with and without gamification.

In the following, we present the background of our work, including gamification, its role in Technology Enhanced Learning (TEL) and related approaches. In Section 3, the presented case study will be explained in detail. Following, the conducted experiment will be described in Section 4. Finally, we will outline our findings (Section 5) before concluding this paper.

Theoretical background

The use of game design elements in non-game contexts (Deterding et al., 2011), also known as gamification, represents a huge trend in Human-Computer-Interaction (HCI), marketing (Zichermann & Cunningham, 2011), enterprise (Schacht & Schacht, 2012) and education (Lee & Hammer, 2011). Already in the 1980s, Malon et al. (1982) researched the positive impact of game elements in interfaces and suggested to use video game elements to enhance the interest, joy and satisfaction of computer systems. Following the predictions of analysts like Gartner (2011) and gamification visionaries (Schell, 2010; McGonigal, 2011), it is likely that gamification will play an important role in future urban spaces, including new forms of gamified education (Charles et al., 2011). Since the rise of gamification, education is a popular application field of this new motivation method. Charles et al. (2011) examined that gamification in education and TEL can increase the learners' engagement, strengthen the social relations, raise satisfaction, help to identify personal strengths and weaknesses and give a more detailed personal feedback.

The use of gamification to support learning through self-reflection with QS tools has not been previously studied in detail. However, this combination has been successfully applied in many popular applications like Nike+⁷, HealthMonth⁸ or Mint⁹. All these examples motivate people with gamification to collect personal information about their behaviour. The target of our research is to transfer this approach to the education context and to examine, if

7 <http://nikeplus.nike.com>

8 ¹ <http://healthmonth.com>

9 <https://www.mint.com>

gamification can improve the motivation to collect data and reflect on it with technology enhanced PLEs in order to improve personal learning.

Case study

The object of our research is an adapted version of the Live-Interest-Meter, a Quantified Self application and PLE that supports the reflective learning process for presenters and listeners. This version of the LIM, which was developed based on Rivera-Pelayo et al. (2013), consists of two main components: 1) the *Meter*, a mobile app that allows capturing and visualizing live feedback in lectures and 2) the *LIM-Community*, a web platform to review past presentations, analyse the personal learning behaviour and interact with other users.



Figure 1. Prototype of the LIM meter as Android App

The Meter (Figure 1) was designed to quantify and track the performance of the presenter as well as the context of the students during a lecture in order to improve the individual learning process of both students and presenter. The tool allows listeners to evaluate a lecture in real time on a uni-dimensional meter, whose scale can be chosen from a

preconfigured set (speed of speech, interest, difficulty and comprehension). The gathered data is aggregated and visualized to the users. If a certain threshold value is exceeded, the presenter will see a discreet hint and can react on it, thereby receiving information about the presentation and adapting it to the audience. This live feedback loop is illustrated in Figure 2. Participants can also compare their individual learning situation with peers. Therefore the Meter supports learners who can reflect on their own performance and improve their behaviour in comparison to peers but also the presenter, who gets real-time feedback about the lecture.

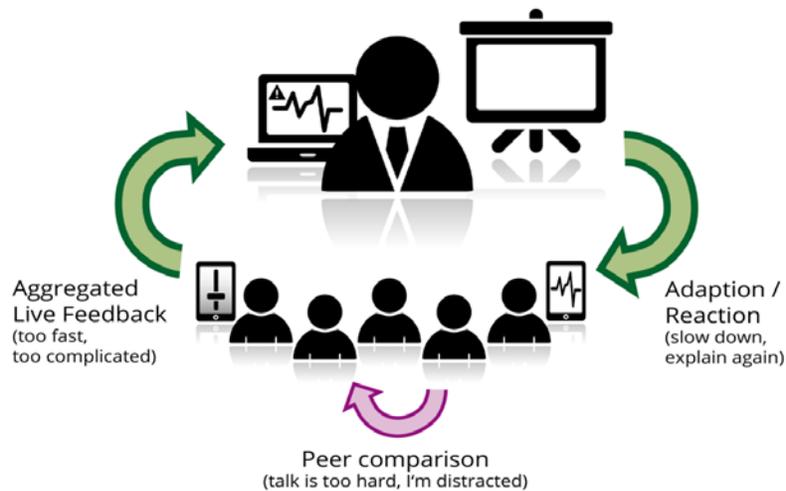


Figure 2. Schematic illustration of the LIM scenario

Later, participants can reflect on a past lecture, by logging in into the LIM-Community. This web platform allows reviewing captured lectures, visualized and aggregated in graphs and enriched with collected metadata like date, time, topic and participants, and also context information (notes added by the audience). The students can discuss about lectures in forums and can evaluate their collected data. Thus, the LIM-Community can be used to recall information from past lectures combined with the associated events in the audience. The data from each lecture is also available for the lecturer, who can review in the LIM-Community

the feedback received in each session and reflect on it in order to improve future presentations. The combination of the Meter and the LIM-Community provides a personal learning environment that helps presenters as well as regular attendees of lectures to visualize, understand and improve their personal learning process and behaviour.

Experiment: The Gamified LIM

We conducted an online experiment with a 2x2 Latin Square design (Hicks, 1973) to analyse the impact of gamification on the motivation to use the LIM, by examining the users' intention to use a gamified and a not gamified version of the tool.

Gamification of the LIM

In order to keep the distraction of the students to a minimum during the lectures, gamification was only applied to the LIM-Community to foster the motivation to collect quantitative and qualitative data during lectures. In a software design process, based on Radoff's (2010) player-centred design model, we selected multiple game mechanics and interface elements for the gamified version of the LIM, based on analysed needs of our target audience. We developed personas, derived from the surveyed participants of Rivera-Pelayo et al. (2013) and matched them with the player types of Bartle (1996) and interviews/profiles of Mayer (2009).

Central element of the gamified version is the personal *Knowledge Tree*. This narrative element stands for the personal learning progress and grows with each lecture in which the user collects data with the LIM. Engagement in using the LIM or the accomplishment of tasks are rewarded with badges and points. The animated badges, e.g. little birds or squirrels (see Figure 3 right), can be decorated in a tree branch, which represents a lecture. The overall personal progress is indicated in points and can be compared with other players in global leaderboards.

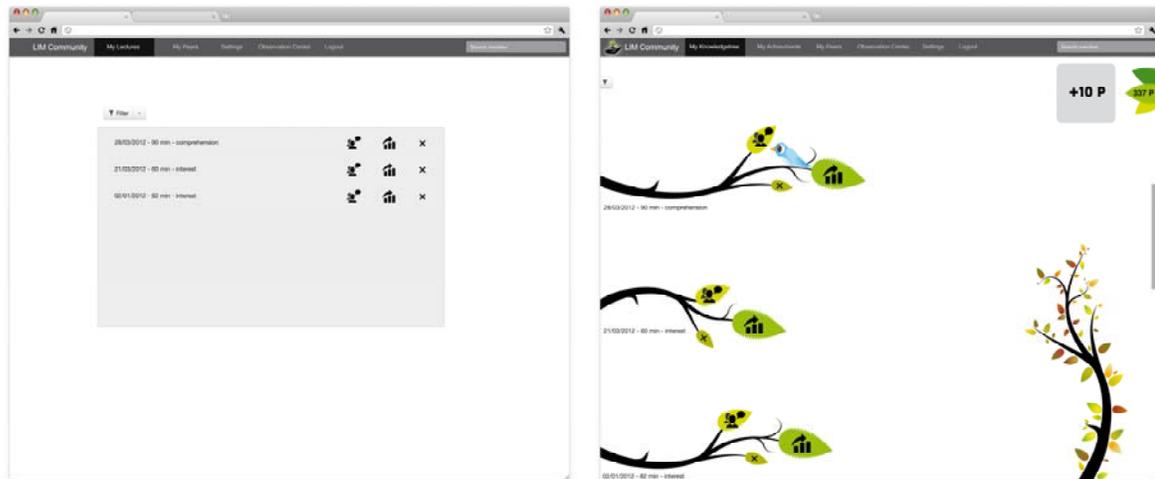


Figure 3. Prototypes of the LIM-Community. Left without gamification, right with gamification.

Experiment

Considering successful gamified QS examples like Nike+⁴ or HealthMonth⁵, it seems that QS approaches can benefit from gamification. Therefore, gamification may be also an appropriate enabler to engage people in using QS approaches as PLEs for improving their personal learning experiences. Concerning our experiment, our first hypothesis was:

H1: *The intention to use the LIM with gamification (game elements, game mechanics, storytelling and playful design) is higher than without gamification.*

Consequently, if gamification can increase the motivation to use the LIM in general, we also believe that gamification can support the motivation to keep tracking, according to Quantified Self, over a long-term period. Li et al. (2010) points out that a lack in motivation is one important barrier in the QS process and motivation to keep collecting personal data must be

raised regularly for long-term success. We believe that suitable gamification elements are able to counteract that problem in a systematic and methodical manner. This leads to the following hypothesis concerning the LIM:

H2: *The intention to use the QS application LIM long-term and to visit the Community regularly is higher with gamification than without.*

In general, it is assumed that gamification can enhance intrinsic motivation. Igbaria et al (1994) showed that ‘system usage is affected by both extrinsic motivation (usefulness) and intrinsic motivation (fun). Both are important in affecting the individual decision whether to accept or reject a new technology’. Based on this, we believe that this is also true for gamified QS-applications:

H3: *The perceived fun during the usage has a positive correlation on the usage intention of QS applications like the LIM.*

We conducted an online experiment which was 20 days active and allowed us to reach the appropriate target group. Each participant was randomly assigned to one of two groups. Group 1 (G1) evaluated first the non-gamified version of the LIM and then the gamified version, whereas group 2 (G2) evaluated the versions the other way around. After a short video introduction and verifying the role of the participant in lectures, the gamified (G) or non-gamified version (O), depending on the group, was presented to the subjects (see Figure 4). Subsequently, we asked the participants a set of questions. Afterwards, we presented them the other version of the LIM-Community and asked them the same questions. Finally, we

asked demographical data and supplementary questions, like the interest to learn from personal QS data. The experimental design allowed us to perform two different analyses between the gamified and non-gamified version: the independent differences between the randomized and independent groups (between-subject) as well as the responses at the individual level (within-subject).

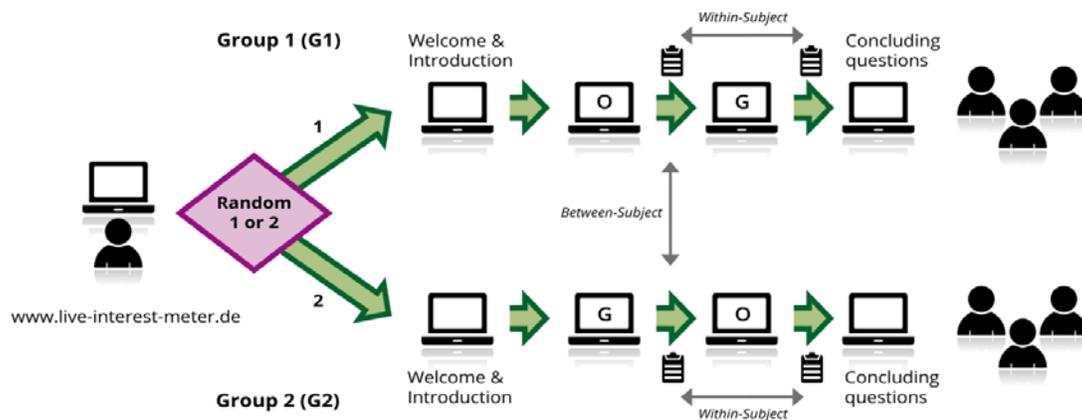


Figure 4. Structure of the experiment

The questions that were asked after each presentation were divided in four sections (see Table 1). The intention to use the software - behavioural intention (BI1), which based on Fishbein & Ajzen (1975) is an indicator for the real usage - was derived from successful TAM studies of Venkatesh & Davis (2000) and Davis (1989). The questions about the long-term usage (BI2) were inspired by Igbaria et al. (1994). Perceived usefulness (PU) was operationalized with five items oriented at Venkatesh & Davis (2000) and Davis (1989). The questions were adapted to the LIM needs. To measure FUN, we used the proven construct from Igbaria et al. (1994) consisted of a 7-point semantic differential with six pairs.

Table 1. Research factors, questions and reliability assessment

Research Construct	Questions	Cronbach's alpha				Scale
		Group 1		Group 2		
		G	O	G	O	
PU	General usefulness (i), usefulness to improve effectiveness (ii) and performance (iv), usefulness for self reflection (iii) and self improvement (vi)	0.898	0.897	0.917	0.945	7-point Likert- scale
	FUN	rewarding /unrewarding (i), pleasant/unpleasant (ii), fun/frustrating (iii), enjoyable/unenjoyable (iv), positive/negative (v), interesting/ uninteresting (vi)	0.925	0.868	0.910	
BI1	General intention to test (i) and use the meter(ii), the community (iii), the collected data (iv)	0.912	0.902	0.872	0.908	7-point Likert- scale
	BI2	Intention to use the meter (i), the community (ii) and the collected data (iii) regularly	0.876	0.837	0.792	

Data analysis and findings

During the online experiment, the website was visited by 607 unique visitors and 14% of them participated in the study. At the end, 70 complete valid data sets (according to control variables and having specified that the participant takes part in lectures regularly) could be used for the analysis. The automatic randomization resulted in 35 valid data records in G1 and 35 in G2. The distribution was homogeneous (Female G1: 13, G2: 13; male G1: 22, G2: 22; median age G1: 26, G2: 24; attend lectures regularly as student G1: 29, G2: 28; attend lectures regularly at work G1: 10, G2: 10; QS interest G1: 22, G2: 19). Therefore, the application of Pearson Chi-Square tests did not show significant differences in the groups, concerning demographic data and QS interest. We assessed the internal consistency of the measurement model by computing Cronbach's alpha coefficients for each of the four constructs PU, FUN, BI1 and BI2 in both groups and both versions (gamified (G) and not gamified (O)). All 16 were between 0,775 and 0,945 and showed a high reliability (see Table 1).

Within-subject analysis

Figure 5 visualizes a descriptive analysis of the individual answers to the two LIM versions (G and O). Comparing the BI1 and BI2 item sums of each participant showed that both kinds of usage intention were higher in both groups with gamification than in the group without gamification. We used non-parametric tests because the application of Kolmogorov-Smirnov-Tests showed that it is possible that BI1(G) ($p=0.045$) and PU(O) ($p=0.046$) are not normal distributed. For this reason, we verified our hypotheses 1 and 2 by using Wilcoxon signed-rank tests. The analysis shows that the intention to use the LIM (BI1) and the long-term and regularly use of the LIM and the LIM-Community (BI2) is with gamification significantly higher than without. These results support our hypotheses 1 and 2.

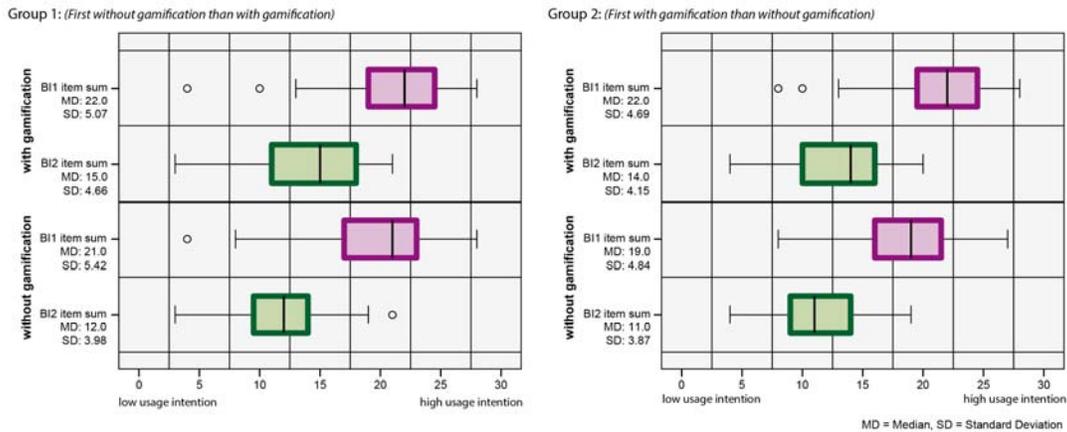


Figure 5. Boxplot of all BI item sums to compare within-subject differences in each group.

Furthermore it was shown that also perceived fun is with gamification significantly higher (see Table 2).

Table 2. Results of the within-subject analysis, Wilcoxon signed-rank test

Comparison			N	Sum of ranks	p (one-tailed)
PU	G - O	Negative ranks	17	478.00	0.025*
		Positive ranks	35	900.00	
		Ties	18		
		Total	70		
FUN	G - O	Negative ranks	12	298.50	0.00**
		Positive ranks	46	1412.50	
		Ties	12		
		Total	70		
B11	G - O	Negative ranks	10	280.00	0.00**

		Positive ranks	46	1316.00	
		Ties	14		
		Total	70		
BI2	G - O	Negative ranks	12	239.00	0.00**
		Positive ranks	43	1301.00	
		Ties	15		
		Total	70		

* is one-tailed significant at 0.05 level.

** is one-tailed significant at 0.01 level.

Between-subject analysis

The within-subject analysis is based on the individual comparison of both versions of the LIM. With the between-subject analysis we tried to determine whether the results are comparable, even if the participants only know one alternative. A tendency for the gamified version was also recognized in this analysis, in which we examined only the independent first answers in each group i.e. the answers to the not gamified version (O) in G1 and the answers to the gamified one (G) in G2. The intention to use the LIM (BI1 & BI2) was with gamification higher than without gamification (Figure 6). However, our hypotheses could not be verified with tests due to the small sample in the between-subject observation.

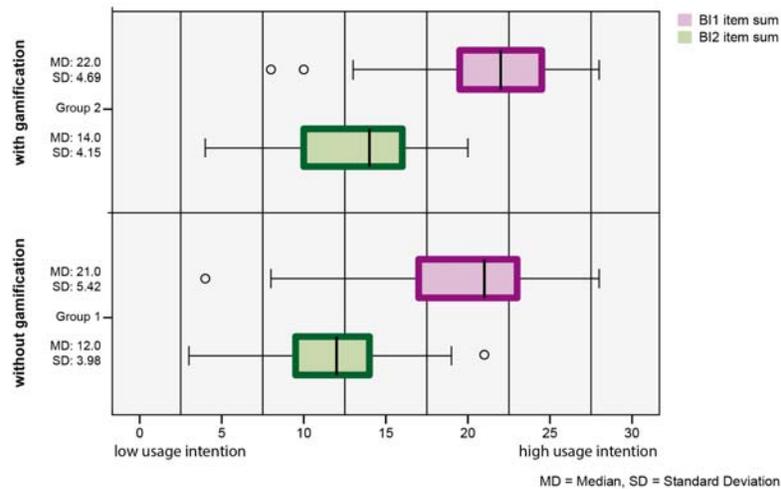


Figure 6. Boxplot with the BI item sums of the first answers in each group to compare between-subject differences.

Correlation analysis

To test H3, Spearman's rank correlation coefficients for FUN, BI1, BI2 and PU were calculated (see Table 3). All group-spanning correlation coefficients were one-tailed positive significant at $\alpha=0.001$ (with Bonferroni correction). The results show a positive correlation between FUN and BI1 & BI2 in both cases. It can be concluded that perceived fun during the usage (FUN) may have a positive impact on the usage intention (BI) of QS applications like the LIM - in general (BI1) higher than in long term (BI2). This finding supports our third hypothesis and compared with the results from the within-subject analysis, it can be said that gamification can increase perceived fun (see Table 2), which have a direct influence on the intention to use an application like the LIM.

Additionally, the correlations show that also PU may have a positive significant impact on BI with higher correlation coefficients than between FUN and BI (see Table 3). These findings follow the results of Igbaria et al. (1994), who measured a stronger influence of PU on BI compared to FUN on BI.

Table 3. Group-spanning Spearman-Rho correlations between FUN, PU, BI1 and BI2.

with gamification		group-spanning	
Sums of		BI1	BI2
FUN	correlation coefficient	.691***	.561***
PU	correlation coefficient	.743***	.679***
without gamification		group-spanning	
Sums of		BI1	BI2
FUN	correlation coefficient	.732***	.572***
PU	correlation coefficient	.773***	.569***

*** $p < 0.001$ (with Bonferroni correction). (N=70)

In addition to the analysis presented above, we conducted a group-spanning comparison of the individual responses. This analysis yielded that 62 out of 70 participants (88.57%) are willing to test the LIM (independent from gamification) and 64.14% (47 out of 70) recognized clear benefits in using the LIM at study or work. Additionally, 45 out of 70 (64.28%) replied to the question ‘If you would use the Meter in a lecture once a week, how often would you login into the LIM-Community’ that they would login once a week or more frequently.

Conclusion

In this study, we showed that gamification can increase the motivation to use QS applications, like the LIM, to collect personal data about the own learning and improve the learning process. The hypotheses H1-H3 were supported by the statistical analysis of the experimental results. The general as well as the long-term intention to use the LIM were with

gamification higher than without. According to Fishbein & Ajzen (1975) and Davis (1989) it can be argued that these intentions have a direct impact on the actual usage. Further, it could be shown that perceived fun has a positive effect on the motivation to use the examined PLE. Together with the finding that perceived fun is with gamification higher than without, we can conclude that gamification can increase the motivation for using the examined application.

Considering the gamification findings and the result that nearly 2/3 of the respondents see clear benefits in using the LIM to improve their personal learning process, gamification seems to be an appropriate enabler to engage people in using QS approaches as PLEs for improving their learning experiences.

Regarding the limitations of this work, the sample size was insufficient to validate the hypothesis with a between-subject test. However, meaningful tendency for gamification was indicated between-subject and the performed within-subject analysis showed high significant results. Furthermore, the measurement was hypothetical and self-reported. Larger experiments in real settings are planned to validate our results. However, this study provides a first important contribution to the successful use of gamification approaches for improving interaction and supporting individual reflective learning with QS tools as PLEs.

Acknowledgement

The project "MIRROR - Reflective learning at work" is funded under the FP7 of the European Commission (project number 257617).

References

- Bartle, R. (1996). Hearts, Clubs, Diamonds, Spades: Players who suit MUDs. *Journal of MUD research*, 1(1), 19.
- Charles, T., Bustard, D. & Black, M. (2011). Experiences of Promoting Student Engagement Through Game-Enhanced Learning. In Ma, M., Oikonomou, A., Jain, L. (Eds.), *Serious Games and Edutainment Applications* (pp. 425-445). London: Springer.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Deterding, S., Dixon, D., Khaled, R. & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining “gamification”. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9-15). New York: ACM.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, Massachusetts: Addison-Wesley.
- Gartner (2011). *Gartner Predicts Over 70 Percent of Global 2000 Organisations Will Have at Least One Gamified Application by 2014*. Retrieved April 20, 2013 from <http://www.gartner.com/it/page.jsp?id=1844115>.
- Hicks, C. R. (1973). *Fundamental concepts in the design of experiments*. New York: Holt, Rinehart and Winston.
- Igbaria, M., Schiffman, S. J., & Wieckowski, T. J. (1994). The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology. *Behaviour & Information Technology*, 13(6), 349–361.

- Lee, J. & Hammer, J. (2011). Gamification in Education: What, How, Why Bother?
Academic Exchange Quarterly, 15(2).
- Li, I., Dey, A., & Forlizzi, J. (2010). A Stage-Based Model of Personal Informatics Systems.
In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
(pp. 557-566). New York: ACM.
- Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from
computer games. In *Proceedings of the 1982 conference on Human factors in
computing systems* (pp. 63-68). New York: ACM.
- Mayer, M. (2009). *Warum Leben, Wenn Man Stattdessen Spielen Kann? - Kognition,
Motivation Und Emotion Am Beispiel Digitaler Spiele*. Boizenburg: Werner
Hülsbusch.
- McGonigal, J. (2011). *Reality is broken*. New York: The Penguin Press.
- Radoff, J. (2011). *Game On: Energize Your Business with Social Media Games*.
Indianapolis: Wiley Publishing, Inc.
- Rivera-Pelayo, V., Zacharias, V., Müller, L., Braun, S. (2012). A Framework for Applying
Quantified Self Approaches to Support Reflective Learning. In I. A. Sánchez & P.
Isaías (Hrsg.), *Proceedings of the IADIS International Conference Mobile Learning
2012* (pp. 123-131). Berlin: IADIS Press.
- Rivera-Pelayo, V., Munk, J., Zacharias, V., Braun, S. (2013). Live interest meter: learning
from quantified feedback in mass lectures. In *Proceedings of the Third International
Conference on Learning Analytics and Knowledge (LAK '13)* (pp. 23-27). New York:
ACM.

- Schacht, M. & Schacht, S. (2012). Start the Game: Increasing User Experience of Enterprise Systems Following a Gamification Mechanism. In Maedche, A., Botzenhardt, A. & Neer, L. (Eds.), *Software for People* (pp. 181-199). Heidelberg: Springer.
- Schell, J. (2010). *When games invade real life*. Video. Retrieved May 3, 2012 from http://www.ted.com/talks/jesse_schell_when_games_invade_real_life.html
- Thom, J., Millen, D. & DiMicco, J. (2012). Removing Gamification from an Enterprise SNS. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 1067-1070). New York: ACM.
- Venkatesh, V. & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, 46(2), 186–204.
- Warschauer, M. (2006). Literacy and technology: Bridging the divide. In D. Gibbs & K.-L. Krause (Eds.), *Cyberlines 2: Languages and cultures of the Internet* (pp. 163-174). Albert Park: James Nicholas.
- Zichermann, G. & Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. New York: O'Reilly Media.