

Mathematics Trails: Shallow and Deep Gamification

Iwan Gurjanow¹, Miguel Oliveira², Joerg Zender³, Pedro A. Santos⁴, Matthias Ludwig⁵

^{1,5} Institute of Mathematics Education, University of Frankfurt, Frankfurt/Main, Germany
{gurjanow, ludwig}@math.uni-frankfurt.de

^{2,4} INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Portugal
{miguel.oliveira.f, pedro.santos}@tecnico.ulisboa.pt

³ University of Applied Science RheinMain, Wiesbaden, Germany
Joerg.Zender@hs-rm.de

Abstract

Mobile Math Trails for Europe (MoMaTrE) is an ongoing project to conceptualize and develop a fully gamified platform for creating, organizing and executing mathematics trails. We present some early experimental results of an empirical study from Frankfurt am Main with 218 pupils concerning the introduction of shallow gamification techniques in the platform. The study was based on the self-determination theory and therefore the Intrinsic Motivation Inventory was used. First results are mostly in line with other findings of the previous research literature, but also some interesting effects could be shown. The article concludes with a discussion of our plans for adding other (deeper) gamification elements.

Keywords: *Mathematics Trails, Gamification, Mobile, Students' Motivation and Education;*

1 Introduction

A maths trail is a trail where the participants can discover mathematics or solve mathematical problems in the environment at predefined stations [1]. Being introduced more than forty years ago, mathematics trails have been mainly serious pedagogical activities.

Along with the possibilities of mobile devices came the opportunity not only to create a smartphone application for walking maths trails but also to gamify this pedagogical activity. Shallow gamification elements have been implemented rather soon in the project, while deeper gamification elements need careful planning and entail more work in terms of its implementation. Nevertheless, we raise the question of the effects of gamification in mathematics trails, a field quite untouched by previous research. To answer the question, we have started to conduct a series of studies, of which one is presented in this paper.

1.1 Shallow and deep gamification

The most basic distinction between types of motivation as given by self-determination theory was made by Ryan & Deci [2] by their subdivision of motivation into intrinsic and extrinsic forms. Many school activities, due to their obligatory character, require an external reason for the engagement of the student. However, even if an activity is initially externally motivated, inherently interesting properties of it can be experienced, leading to a motivation shift [2].



Gamification describes various techniques for controlling the behaviour of users through game elements towards a specific goal, that is, the application of game elements in a possibly non-gaming context [3]. The overriding goal of gamification is to modify the affected activity, which was originally designed for a specific purpose, so that the user feels intrinsically more engaged, thus increasing their intrinsic motivation and commitment.

There are several experiences of gamification in education, using different approaches [4, 5, 6]. Techniques can be thought at two different levels. In a *shallow* level or *thin layer* of gamification, the core teaching and learning processes are not substantially changed. There are still lectures, exercises, homework. But the language changes to making quests, crafting items, defeating bosses with the grade given in Experience Points. An example of such an approach applied to undergraduate studies is given in [5]. Another possible shallow gamification technique is to give stars, badges and prizes for activities in the course, use leaderboards, or yet use game-like interface components. Shallow gamification has been the target of some criticism because it can be seen as manipulative and making excessive use of external motivation [7]. Shallow gamification can be seen as a layer that is put above and on top of the core processes, without changing their essence.

In contrast to shallow gamification, deep gamification can be defined as introducing game elements that change the core processes of the activity [8]. A seminal example of that approach was given by Quest to Learn, an innovative school for grades 6 to 12 that started in 2009 in New York City [6], where the whole curriculum was planned using game design techniques. While shallow gamification needs mainly programming and visual design skills, deep gamification uses mainly game design skills, because it is necessary to rethink the activity and design game mechanics at its core.

1.2 Mathematics Trails, MathCityMap and the MoMaTrE project

Blane & Clarke [9] were among the first to present the maths trail idea to a broad, scientific audience. Their concern at the time was the popularization of mathematics. Today, we see the benefits above all in the application of mathematics in real, authentic situations, as well as the modelling that precedes the calculations. Gamification of maths trails can be found even at the early days of maths trails. The first element of Gamification reported is storytelling. In the '80s Blane [10] mentioned a maths trail with role-playing in an old gold mining town, which wraps a story around the user to be one of the gold miners. Also in the '90s Muller [11, 12] blazed two trails in Canada which take the user on a walk together with a group of three, who are discussing certain phenomena along the way and at some points asks the user to do a certain calculation to fill in the missing numbers. Another good example is reported in [13], where students created a maths trail as an adventure game, where the users have to stop the evil Dr Weed by solving several tasks in his lab. Furthermore, one can raise the question if a maths trail in itself is a possible gamification of a maths lesson. In a Finnish study, pupils were highly motivated and had lots of fun by running a maths trail with textbook tasks placed as QR Codes inside the school building [14]. The authors highlighted, that the pupils were happy to leave the classroom and move around, although the tasks they have completed were similar to the ones they were used to in the classroom.

The MathCityMap (MCM) project of the Goethe University in Frankfurt combines the idea of mathematics trails with the possibilities of smartphones [15]. The main idea of the project is to give the participants another perception and experience of mathematics. Maths takes place mainly indoors inside the school, inside the classroom. However, all human concepts, including mathematical concepts, are based on the perceptual-motor system experiences we have while interacting with the world around us [16]. Furthermore, the project addresses the advantages of Web 2.0, to have users create share- and reusable content. The MathCityMap website (www.mathcitymap.eu) is a portal, where (teacher) users can create tasks everywhere on the world map. They can ask for the



publication of their tasks and if successful, other users can combine these tasks with their tasks for a maths trail (which can also be published).

The objective is to bring mathematics outdoors more often, for students as well as every citizen, the same goals that have been shared from the beginning of the maths trail concept [10]. We seek for new forms of getting a mathematical view of one's neighbourhood, of one's environment, see the Math questions and problems which are everywhere and, on top of that, do it digitally with a smartphone.

In order to leave the borders of Germany and to further develop MathCityMap, a consortium was formed and an Erasmus+ Grant obtained the MoMaTrE (Mobile Math Trails for Europe) project. The main target groups for this project are student teachers, in-service teachers and the public. Our approach contains a platform and a mobile application to create tasks and with these tasks, mathematical trails can be built by everybody, especially teachers. New ways of collaboration should be possible and we want to build a community of active maths trailers, who share their work and help each other to develop things further. Some shallow gamification elements are part of the app today. In this paper, we will present and discuss a study done to measure their effects and we will present ideas how to apply other gamification elements (including some deep gamification) to the concept of mathematics trails.

Since debriefing is crucial for learning after a gamified pedagogical experience [17], during the MoMaTrE lifetime a digital classroom option was developed for the MathCityMap project. In alignment with privacy and data protection, it allows the teacher to see the actions of the pupils during the maths trail activity and get a report right after it is finished. This way, the teacher is provided with material for a debriefing session after the maths trail activity, whether it follows up directly or during the next maths lesson.

1.3 Literature results on the impact of gamification

Dicheva & Dichev [18] conducted a meta-analysis on gamification in education. Compared to the period of over four years (from January 2010 to June 2014), where 34 papers related to this topic were published, and the time from July 2014 to June 2015 (one year), the number of published papers has increased to 41. These findings indicate a growing interest in gamification in the area of education. Nevertheless, only about half of the publications can draw a positive conclusion [18].

The intrinsic motivation and experience associated with playing video games can be seen in gamified activities as well [19]. This is why gamification can benefit from a large array of areas when used correctly as is shown in several studies.

For instance, tourism [20] where the presence of game mechanics in a travel-related application and website resulted in an improvement in terms of engagement and made the experience more social and interactive. Another example is the introduction of gamification in higher education classrooms. One of these cases resulted in an increase in approval rating, interaction and attention in the classroom [21]. Other cases indicated positive effects on the engagement of students and a moderate improvement in learning outcomes, see for instance Ibáñez et al. [22] and Santos [8].

The gamification of a training module with fiction also showed a significant improvement in the satisfaction of the trainees. This same study [23] also shows that the declarative knowledge did not differ between the control condition and the gamified condition, but the procedural knowledge scores were higher in the control condition.

2 The impact of shallow gamification elements

2.1 Introducing gamification elements into MathCityMap



Besides many positive observations of students walking a maths trail, two negative observations lead to the addition of gamification elements to the MathCityMap app. Firstly, students often tried to guess answers, if their first attempt was incorrect. Secondly, there seems to be a motivational obstacle to begin working on the tasks, which is expressed by walking slowly to the first task and thus leading to a low ratio of doing mathematics to the time spent on the trail. Before a decision on the type of gamification was taken, we have analyzed the project as suggested by Morschheuser et al. [24] and defined gamification goals.

The mathematics trail activity as supported by the MathCityMap app focuses on secondary school students, who are familiar with using smartphones and apps. A maths trail is usually carried out on an irregular basis e.g. day's hike or project days. In our approach, students collaborate in groups of three (one is using the MCM app, one is responsible for measuring and the last person has the task to take notes) and walk the trail independently. To complete a maths trail students have to complete each task that is part of the trail. The activity to complete a task is divided into seven sub-activities: (1) finding the task's location; (2) reading the task's description; (3) choosing a mathematical model; (4) collecting data; (5) calculating the answer; (6) entering answer into the app and getting feedback; (7) optionally, taking hints and retry. During the steps (1), (2), (6) and (7) students have to use the application. Finally, two gamification goals were defined.

- Prevent students from guessing answers.
- Increase intrinsic motivation for working on maths trail tasks (increase the number of tasks completed per hour).

Based on Lieberoth [25], who found shallow gamification in the form of a game-like design to have an impact on the participants' motivation, the first gamification elements added to the application were simple and fall into the category of shallow gamification. Following his suggestion to "break clusters of game elements down to individual functional units" [25], we have decided to create three different versions of the app. Each version adds further game elements to the application. To prevent guessing, a points system (G1) was added to the task activity, rewarding the students with up to 100 points for a correct answer. Additionally, each wrong answer after the first one decreases the value of the task by ten points. The second gamification approach is the local leaderboard gamification (G2). It augments the points gamification with the possibility to see the score of the user in front and behind you so that users get into competition with each other. To not frustrate the last ranked group, we have added a computer player, who will be always ranked last.



Table 1. *Types of gamification in MathCityMap.*

G0: No gamification	G1: Points	G2: Local leaderboard
		

2.2 The study

In a study with ninth-graders (15 years old students), the following research questions were studied in the summer semester 2017:

- How does the gamification system influence the motivation of the participants?
- What effects does the gamification system have on the performance parameters (completed tasks per hour, incorrect entries per task) of the 9th graders when working on a maths trail?

Methodology. 218 ninth-grade students (111 female, 107 male) from 14 different school classes took part in the tests. They were randomly assigned to the control group G0 (no gamification; N=84; 44 female, 40 male), experimental group G1 (points gamification; N=60; 31 female, 29 male) or experimental group G2 (local leaderboard gamification; N=74; 36 female, 38 male). After a short introduction to the MathCityMap application and the use of measuring tools, the students were handed out a set of materials (smartphone, measuring tools and a paper trail guide) and walked a mathematics trail in groups of three (in some cases of two). Overall, the 218 students formed N = 96 groups. The trail's tasks focused on cylinder problems. Finally, they were asked to fill in a standardized questionnaire on intrinsic motivation. We designed the intervention as close to real (German) mathematics classes as possible so that findings could be easily applied to the school context. For instance, teachers divided their students into groups of three themselves. Furthermore, we invited teachers and their students for a frame of 90 minutes that is a typical duration of mathematics classes (two classes with 45 minutes each). Subtracting the time for the introduction and the time for answering the questionnaire, about 70 minutes were left to walk the trail. The motivation questionnaire contained a subset of questions of the Intrinsic Motivation Inventory [21]. The questionnaire comprises 21 items that could be answered on a 7-point Likert scale. The questions belong to one of four subscales: (1) enjoyment, (2) competence, (3) usefulness and (4) pressure. The first subscale aims to measure directly the intrinsic motivation. Questions on this subscale ask if the activity was interesting or enjoyable. The subscales (2) competence and (3) usefulness can be considered as indicators to classify the activity as intrinsically motivated. The fourth subscale pressure, measures if the participant carried

out the activity on a voluntary basis. Higher scores in the first three subscales and lower scores in the last subscale correspond to a participant whose motivation for performing an action is located intrinsically.

The second data source consists of a technical log. The provided smartphones came with a special MCM app version that saves information on each event the students perform. Interesting events for the data analysis of this article are the answer to validate events that are accompanied by a timestamp and geolocation. These events provide us with information on how many tasks the students solved, how much time it took them to do so if they were near the task when entering the answer and finally, how many wrong answers they gave per task.

During the field studies, we could observe students copying results from each other. This unwanted behaviour would distort the data analysis that is why the raw data was filtered before analyzing it. The filter rule took the time-on-task (the time from the first opening of the task until the validation of the right answer) as well as the geolocation into account. Therefore, a task was classified as "cheated" when the time-on-task was below 60 seconds and the geolocation indicated that students were not near the task's object when validating the answer. In case that the time on task was below 60 seconds but the position seemed legit, we verified students' notes and sketches manually to decide on the state of the particular task. The following results use filtered values of tasks per hour.

The motivational scores are reported individually with $N = 218$. The performance parameters (tasks per hour and incorrect entries per task) that were extracted from the technical logs are reported per group of students with $N = 96$.

Results. Table 2 displays the overall scores of the questionnaire as well as the performance parameters. With mean values of 4.7 and 4.5, the subscales (1) enjoyment and (3) are significantly higher than the theoretical average value of 4 ($p < .01$). The perceived competence (2) mean score of 3.9 can be considered average. The pressure subscale (4) mean value (2.7) is lower than average. The subscales indicate that the overall experience of walking a maths trail supported by a smartphone application was perceived positively regarding motivational aspects. Students enjoyed the activity and had the feeling that it was useful to them. In average, a group of students managed to complete about three tasks and made about 3.1 incorrect entries per task. The minimum and maximum of both scores indicate big differences between the groups.

Table 2. Overall scores IMI and performance parameters.

	N	Mean	SD	Min	Max
Enjoyment	218	4.7	1.3	1.1	7.0
Competence	218	3.9	1.4	1.0	7.0
Usefulness	218	4.5	1.4	1.0	7.0
Pressure	218	2.7	1.1	1.0	5.8
Tasks per hour	96	2.9	1.6	0.0	7.1
Incorrect entries per task	96	3.1	3.2	0.0	15.9

Figure 1 shows boxplots for the four motivational subscales divided by gamification. At the first glance, the subscales (1) enjoyment and (3) usefulness appear on a similar level throughout the gamification types, whereas the subscales (2) competence and (4) pressure differ. In particular, the scores of the local leaderboard gamification (G2) seem



above the others. Multiple two-way analysis of variance was conducted on the influence of one independent variable (gamification) on (a) the enjoyment score, (b) competence score, (c) usefulness score and (d) the pressure score. Gamification included three levels (G0 – no gamification, G1 – points gamification, G2 – leaderboard gamification).

- The effect was not statistically significant ($F(2, 215) = 1.908, p = .151$).
- The effect was statistically significant at the .01 significance level. The main effect for gamification yielded an F ratio of $F(2, 215) = 4.954, p < .01$, partial $\eta^2 = .044$, indicating a significant difference between G0 ($M = 3.9, SD = 1.4$), G1 ($M = 3.5, SD = 1.5$) and G2 ($M = 4.3, SD = 1.2$). Post-hoc multiple comparisons with LSD revealed a statistically significant difference between the G1 and G2 group with $p < .01$.
- The effect was not statistically significant ($F(2, 215) = 0.982, p = .376$).
- The effect was statistically significant at the .05 significance level. The main effect for gamification yielded an F ratio of $F(2, 215) = 4.022, p < .05$, partial $\eta^2 = .036$, indicating a significant difference between G0 ($M = 2.5, SD = 1.1$), G1 ($M = 2.8, SD = 1.1$) and G2 ($M = 3.0, SD = 1.1$). Post-hoc multiple comparisons with LSD revealed a statistically significant difference between the G0 and G2 group with $p < .01$.

The conducted ANOVAs show that gamification had a significant influence on competence and pressure subscale. In particular, the G2 group reported higher scores in the competence subscale as well as in the pressure subscale. Whereas the effect on enjoyment and usefulness subscales was not statistically significant.

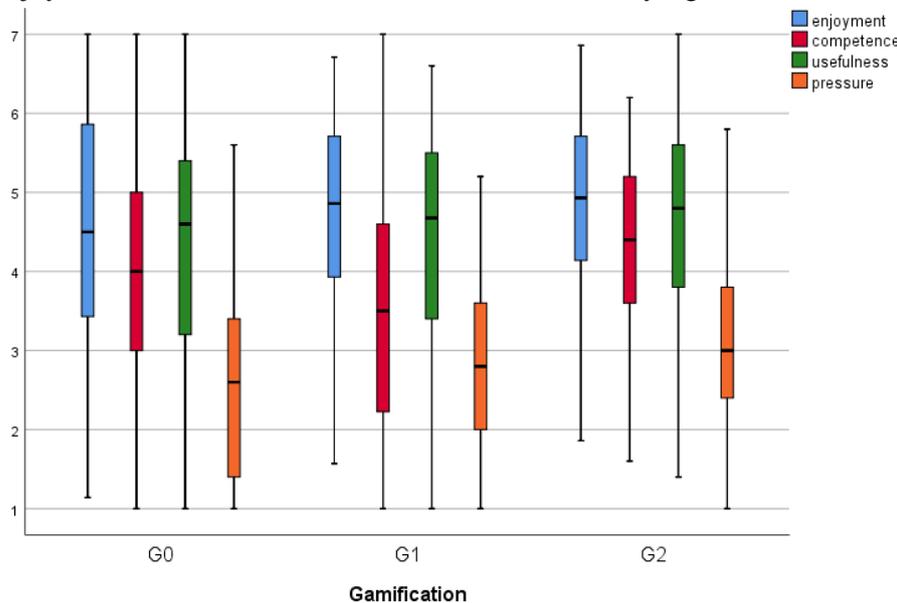


Figure 1. Box plots of the motivational subscales.

Figure 2 shows boxplots for the performance parameters tasks per hour and incorrect tries per task divided by gamification. At first glance, the values of G0 and G1 hardly differ, whereas the G2 gamification shows a higher number of tasks per hour and a lower number of incorrect entries per task. Two two-way analysis of variance was conducted on the influence of one independent variable (gamification) on (e) the tasks per hour and (f) the incorrect entries per task.

- The effect was not statistically significant ($F(2, 93) = 2.189, p = .118$). Nevertheless, posthoc multiple comparisons with LSD revealed a statistically significant difference between the G0 and G2 group with $p < .05$.
- The effect was statistically significant at the .05 significance level. The main effect for gamification yielded an F ratio of $F(2, 93) = 3.345, p < .05$, partial $\eta^2 = .067$,



indicating a significant difference between G0 (N = 36, M = 3.8, SD = 3.2), G1 (N = 26, M = 3.6, SD = 4.1) and G2 (N = 34, M = 2.0, SD = 2.0). Post-hoc multiple comparisons with LSD revealed a statistically significant difference between the G0 and G2 group with $p < .05$ and the G1 and G2 group with $p = .05$.

Although gamification had no statistically significant effect on tasks per hour in the ANOVA analysis, a posthoc comparison revealed that the G2 group had completed significantly more tasks than the G0 group. Regarding the incorrect entries per task, gamification had a statistically significant effect on the model. Especially the G2 group had significantly lower values (fewer mistakes) than both other groups.

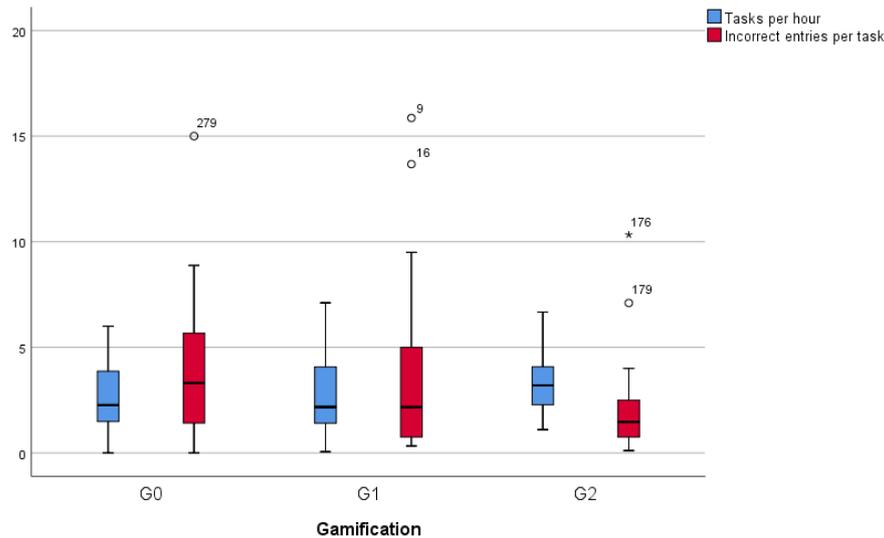


Figure 2. Box plots of the performance parameters.

An earlier study indicated that the influence of gamification might be dependent on the gender of a person [27]. That is the reason why, besides the gamification, in the further process also the gender of the students was taken into consideration. Table 3 displays the mean and SD values for the four subscales divided by gamification and by gender.

Table 3. Mean and SD values of motivational subscales divided by gamification and gender.

	N	G0 (N=84)		G1(N=61)		G2(N=73)	
		M N=40	F N=44	M N=30	F N=31	M N=37	F N=36
Enjoyment	218	4.6 (±1.5)	4.3 (±1.4)	4.7 (±1.1)	4.8 (±1.3)	4.8 (±1.2)	4.9 (±1.3)
Competence	218	4.3 (±1.5)	3.6 (±1.2)	3.4 (±1.7)	3.6 (±1.3)	4.5 (±1.1)	4.0 (±1.4)
Usefulness	218	4.6 (±1.4)	4.0 (±1.5)	4.4 (±1.3)	4.4 (±1.5)	4.7 (±1.2)	4.6 (±1.5)
Pressure	218	2.5 (±1.3)	2.5 (±1.0)	2.7 (±1.0)	2.8 (±1.2)	3.0 (±1.2)	3.0 (±1.1)



Multiple two-way analysis of variance was conducted on the influence of two independent variables and their interaction effect (gamification, gender, gamification * gender) on (g) the enjoyment score, (h) the competence score, (i) the usefulness score and (j) the pressure score. Gamification included three levels (G0 – no gamification, G1 – points gamification, G2 – leaderboard gamification) and gender consisted of two levels (M – male, F – female).

- g) None of the effects were statistically significant ($F(2, 212) = 1.817, p = .165$).
- h) One effect was statistically significant at the .01 significance level. The main effect for gamification yielded an F ratio of $F(2, 212) = 5.066, p < .01$, partial $\eta^2 = .046$, indicating a significant difference between G0, G1 and G2. The main effect for gender yielded an F ratio of $F(1, 212) = 3.281, p = .071$, partial $\eta^2 = .015$, indicating that the effect for gender was not significant, female ($M = 3.7, SD = 1.3$) and male ($M = 4.1, SD = 1.5$). The interaction effect was not significant, $F(2, 212) = 2.273, p = .105$, partial $\eta^2 = .021$.
- i) None of the effects were statistically significant ($F(2, 212) = 0.897, p = .409$).
- j) One effect was statistically significant at the .05 significance level. The main effect for gamification yielded an F ratio of $F(2, 212) = 3.942, p < .05$, partial $\eta^2 = .036$, indicating a significant difference between G0, G1 and G2. The main effect for gender yielded an F ratio of $F(1, 212) = 0.008, p = .931$, partial $\eta^2 = .000$, indicating that the effect for gender was not significant, female ($M = 2.7, SD = 1.1$) and male ($M = 2.8, SD = 1.2$). The interaction effect was not significant, $F(2, 212) = 0.062, p = .940$, partial $\eta^2 = .001$.

In none of the ANOVAs gender or the interaction effect of gamification * gender could explain the differences in variances of the motivational subscales significantly.

In previous publications (e.g. [27]), we, unfortunately, made a logical mistake when analyzing the tasks per hours and the incorrect entries per task divided by gamification and by gender. In contrast to the motivational subscales, it does not make sense to look at these parameters individually, since they were measured per group (triad). The following analysis uses a different approach. Hence, a group has no gender attribute, we classified groups depending on their constitution in one of three categories: female (short F; group consists of girls only), male (short M; the group consists of boys only) and mixed (short X; the group consists of boys and girls). The corresponding teacher of a class constituted the groups, therefore this variable can be considered as not controlled by the researchers. Table 4 shows the mean and SD values of the performance parameters divided by gamification and group constitution.

Table 4. Mean and SD values of performance parameters divided by gamification and group constitution.

	F (N = 28)			M (N = 36)			X (N = 28)		
	G0 N=10	G1 N=5	G2 N=13	G0 N=19	G1 N=9	G2 N=8	G0 N=4	G1 N=12	G2 N=12
Tasks per hour	2.7 (±1.7)	3.6 (±2.8)	3.2 (±1.3)	3.0 (±1.4)	2.9 (±1.4)	2.8 (±0.9)	2.0 (±0.8)	2.4 (±1.9)	3.9 (±1.6)
Incorrect entries per task	3.9 (±2.3)	2.1 (±2.2)	1.6 (±1.0)	3.5 (±2.9)	1.8 (±1.5)	2.1 (±2.3)	2.6 (±1.7)	5.7 (±5.2)	2.4 (±2.8)

Two two-way analysis of variance were conducted on the influence of two independent variables and their interaction effect (gamification, group constitution, gamification * group constitution) on (k) tasks per hour and (l) incorrect entries per task.



Gamification included three levels (G0 – no gamification, G1 – points gamification, G2 – leaderboard gamification) and group constitution consisted of three levels (M – male, F – female, X - mixed).

k) None of the effects were statistically significant ($F(2, 83) = 0.435, p = .649$).

l) None of the effects were statistically significant ($F(2, 83) = 1.130, p = .328$).

2.3 Conclusions

The motivation of the groups that used a gamified app version tends to be higher but not significantly so. This result is in line with the analyzes of different studies on gamification by Dicheva & Dichev [18] and suggests that the actual activity is crucial for the motivational expression and the introduction of shallow game elements has at most a small impact on this. This is also in line with the concept of flow [28], as it requires, for example, a challenging activity with clear goals, which comes from the activity itself and not from the gamification. Furthermore, it is important to get immediate feedback, which is related to the app itself, but not specifically to the gamification. So we can assume, that the flow is supported mainly by the activity and the app, but not by the shallow gamification.

On the other hand, gamification significantly influenced performance parameters. Groups that used the leaderboard gamification G2 outperformed groups without gamification (G0) regarding the number of completed tasks per hour. Furthermore, participants of the G2 group made significantly fewer mistakes per task than participants of both other groups (G0 and G1). The G2 group reports also a significantly higher value in the pressure subscale. It seems that, in our case, only the leaderboard gamification had a remarkable influence on the group activity by transforming the maths trail into a mathematical competition or race. The transformation adds pressure to achieve a good ranking that engages all group members. Even though the pressure is usually perceived as a contra-indicator for intrinsic motivation, the actual enjoyment scores of the G2 group do not differ significantly. Through the introduction of a visible ranking, the leaderboard gamification manages to influence the students to carry out the activity in a serious way (although a teacher does not supervise them) without decreasing the fun. Unlike earlier findings [27], the new classification regarding the constitution of the groups (female, male or mixed) revealed that gender made no significant differences within the motivational subscales as well as the performance parameters.

Introducing gamification elements into a complex educational situation through an application is challenging. In particular, the cooperative group working method, in which the only one of three participants uses the gamified app, may hinder the gamification from unfolding its psychological potential on all participants. For the gamification to work, the active user must influence other team members through his actions and his speech. Considering this challenge and the results that the points gamification (G1) did not significantly differ from the non-gamified version, allows us to draw an alternative conclusion. Points may hold psychological power on an individual level, but not enough to become visible in the group setting.

In the broader context of gamification research, we followed the suggestions of Lieberoth [25] to examine the impact of individual functional units on motivational subscales. In particular, the impact of points (G1) and a local leaderboard (G2) on motivation and performance in an educational setting were studied. In our case, only the leaderboard gamification could improve performance parameters. However, one has to consider the methodical particularities of mathematics trails: group work, where only one member uses the gamified app. Further research is required to answer the question if these findings can be applied to other settings as well (e.g. individual settings).

Zender & Ludwig [29] discuss further results on learning performance of students using the MathCityMap app.



3 MoMaTrE and deeper gamification

Video Games are valued by their potential to provide a flow experience [29]. Since we can see that the mobile supported maths trail activity with MathCityMap fulfils many of the flow criteria as immediate feedback, sense of control, concentration on the task at hand, clear goals, being challenging and requiring skills, it is possible to turn the whole activity into a game. Deeper gamification techniques such as a narrative approach may turn MathCityMap into a valued game, thus hopefully increasing the motivation further.

Indeed, the above results support the idea of implementing and testing more complex gamification elements. These might not be, by definition, deep gamification elements but they are objectively richer and more elaborate than the ones presented so far. In this section, we propose three different solutions that would flesh out the interaction between the students and the MathCityMap app. These solutions, although independent, are not exclusive, meaning that they could, in theory, work together and are designed to even be better if implemented side by side.

These features should be created one by one and their implementation and design should accommodate the other features, in order for the user experience be as seamless as possible.

3.1 Narrative approach

The first proposed feature is the implementation of narrative elements in the routes.

A study performed to 858 secondary school students [31] shows that their preference for video games in the classroom is affected directly by, amongst other things, their perception regarding the usefulness of videogames and their previous experiences with them. This could be favourable due to the demographic consisting of students below 18 years old, which recent studies show is more propitious to playing video games [32].

The introduction of narrative elements to the MathCityMap app brings some challenges. Several maths trails which utilize fiction [10, 11, 12, 13] are very strict when it comes to the presentation of that fiction, being limited to one story, very contextualized within the target theme of the trail. This inflexibility, however, has the advantage of being more appropriate for the trail.



Figure 3. First visual draft of a non-contextualized Pirates narrative.

The MathCityMap app allows users to create and participate in several routes. As such, creating a narrative contextualized for each one is a challenge due to the necessary

authoring effort. It is not our goal to force the user who creates a route to also create a narrative that fits said route. Therefore, two possible solutions to cover different approaches for this problem are being worked on. Both solutions are based on the creation of narrative nodes. Narrative nodes are pieces of information that contains the text to be displayed, as well as meta-data relative to when and how this text will be displayed:

Non-contextualized narrative nodes: By non-contextualized narrative nodes we mean a set of fiction nodes that between themselves are part of a bigger narrative but that is non-connected with the tasks or the route itself. The nodes have different purposes:

- To introduce the narrative (these nodes would appear at the start of a route);
- To introduce each task;
- To replace the message of a correct or incorrect answer;
- To make the bridge between tasks;
- To conclude the narrative (these nodes would appear after the completion of the last task of the route);

A system to avoid repetition of text is planned. In order to avoid the same narrative to become stale and repetitive, several nodes that serve the same purpose should be created, so that a function that organizes and chooses nodes based on frequency can be used.

Overall, this solution is fairly light in terms of authoring effort, the tradeoff being a possible disconnection between the trail and the narrative in terms of flow.

Contextualized narrative nodes: This approach is similar to the one stated above but the nodes are specific to each task type. The narrative would make references to the problems present in the tasks themselves instead of just presenting itself without mentioning the current task objective. This means that more specified nodes would have to be created to introduce a set of types of tasks. The number of nodes created would be proportional to the number of task types. As an example, nodes could reference some tasks available in the task wizard like:

- Determination of slopes of ramps;
- Number of stones in a rectangular wall;
- The volume of a cylindrical pool;
- Walk a certain distance or the shape of a geometrical object.

This solution requires a higher authoring effort when compared to the non-contextualized one, due to the fact that, as previously mentioned each task needs to have specific narrative nodes tailored to it, which implies a larger number of nodes. Also, it may be difficult to create a route where all the tasks have narrative nodes specific to their type. The higher authoring effort can be mitigated with the use of Procedural Content Generation techniques, in a similar way to what has been proposed for game level generation [33].

In contrast with the previous approach, this one has the possibility to create more engaging experiences by blending the narrative elements more deeply within the activities of the route. Also, the same approach taken by the previous proposal can be taken here regarding repetition, of course, this would mean even more authoring effort, but it is something to take into account.

To summarize, both solutions have pros and cons, and possibly the best solution is an implementation that balances both approaches, so it uses contextualized nodes whenever possible, but always measuring the authoring effort in order to use un-contextualized nodes whenever it is deemed to be too much to author every specific task type.

In terms of how the narrative approach would be implemented, first, the idea is to provide teachers with a set of predefined narratives that they can attach to their created routes, but in the future, the possibility for route authors to generate their narrative is something that should be possible.



3.2 *Gamified Activities*

On top of the narrative approach, we propose the addition of objectives to tasks and routes. This also aims to improve the experience of the users and cooperation between them.

As proposed by Hauge et al. [34] context-aware activities can be implemented in such a system, for example, a Treasure Hunt, or a Conquer activity. Changes on the routes themselves can be made, like limiting the number of tasks showed, having a system where the completion of a task unlocks the next one, connecting task activities, either by having two tasks in the same place that refer one another in some way (for example, having a task become easier after completing another because one of the steps to solve it was already done in the previous) or by, for example, having a task placed at the end of another (one task may be “walk 50 meters south”, and upon arrival at the solution, the next task would be unlocked)

This helps to connect the several tasks and making them less individual. On top of that, it also allows for a bigger pool of options in the integration of the narrative elements.

Different ways to gamify routes are also possible, the simple implementation of time as a factor could enable a more competitive side of the routes by the using races as a variant of a normal route. Another example is to add a cooperative element to the interactions by creating routes where teams receive different tasks, and the reward of those tasks is a clue, after solving their respective task, teams should come together with their clues in order to complete a final challenge.

These are two examples of new activities that could be implemented within any route to tailor the types of interactions between the participants. Ideally, the author of a route could just pick the variant of interaction he wants the participants to have and apply that to his route.

3.3 *Global teams and challenges*

Games are seen as a social activity by most teenagers [35] and this can be used to enhance the experience and engagement of a gamified application.

We propose a feature that creates a light social component, inspired by recent video games like PokemonGO that gives the player the possibility to choose one of three teams to be a part of. These teams do not affect the gameplay in a major way but give the players something to argue about.

Another example is the Splatfest game mode in Splatoon and Splatoon 2. This mode allows each player to choose one of two options, dividing the community into two teams. At the end of each month, a team win based on the number of players that voted on the team and the performance of those players in the game. The rewards given to the winning team are not game-changing but that does not hinder this event's popularity.

The proposed feature is similar to Splatfest in the sense that each user of the MathCityMap app could choose a team and from time to time a team would be chosen as the winner based on the number of completed routes of the members of said team.

This feature intends to not interfere with the normal use of the application while doing routes, but instead, to create a new layer of gamification, more passive that has the possibility to reach a lot of users and create a unique experience within the app.

4 *Conclusions*

In this paper, we presented the results of a study on the effects of shallow gamification in the mathematical trails application MathCityMap, which is being developed as part of the MoMaTrE Erasmus+ project. The study showed that the introduction of shallow gamification elements has at most a small impact on motivation, as measured by motivation questionnaires, but successfully influenced performance parameters.



Especially the leaderboard gamification increased the speed of task resolution and lowered the number of incorrect answers per task. So, while the full promised potential of gamification was not reached, there remains the question if deeper gamification can improve this type of results. With the goal of shifting motivation to a more intrinsic nature, we propose to introduce both narrative arcs associated and interlaced with the trails and meta-team and challenge creation. The implementation of narrative elements is currently nearing completion. The effects of those gamification techniques in the MathCityMap application will then be tested.

References

- [1] M. Shoaf, H. Pollak and J. Schneider, "Math trails", Lexington: COMAP, 2004.
- [2] R. M. Ryan and E. L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions", *Contemporary educational psychology*, 25(1), pp. 54-67, 2000. <https://doi.org/10.1006/ceps.1999.1020>
- [3] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification", In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pp. 9-15, ACM, 2011. <https://doi.org/10.1145/2181037.2181040>
- [4] I. S. Labouriau, "Instrução Personalizada na Matemática Universitária", *Boletim da SPM* 64, Maio 2011, pp. 55-65, 2011.
- [5] L. Sheldon, "The multiplayer classroom: Designing coursework as a game", Cengage Learning, 2011.
- [6] K.S. Tekinbas, R. Torres, L. Wolozin, R. Rufo-Tepper, and A. Shapiro, "Quest to learn: Developing the school for digital kids", MIT Press, 2010. <https://doi.org/10.7551/mitpress/8909.001.0001>
- [7] I. Bogost, "Why Gamification is Bullshit", In: *The Gameful World - Approaches, Issues, Applications* pp. 65-80, MIT Press, 2014.
- [8] P.A. Santos, "Gamification of a University Course", In: *Proceedings of SciTecIN'15, Sciences and Technologies of Interaction*, Coimbra, 2015.
- [9] D. C. Blane and D. Clarke, "A mathematics trail around the city of Melbourne", Monash Mathematics Education Centre, Monash University, 1984.
- [10] D. C. Blane, "Mathematics Trails", *ICMI Papers on The Popularization of Mathematics*, 1989.
- [11] E. Muller, "Niagara Falls Math Trail", Department of Mathematics, Brock University, 1993.
- [12] E. Muller, "Welland Canal Math Trail", 1993. <http://spartan.ac.brocku.ca/~emuller/mathtrail/wcmt/wcmtstart.html>
- [13] K.S. Quek, "Context for mathematical problem posing", In J. Ee., Kaur Berinderjeet, N. H. Lee. & B. H. Yeap (Eds.), *New Literacies: Educational Response to a Knowledge-based Society*, Vol 1: Education pp. 612-620. Singapore: Educational Research Association of Singapore, 2000.
- [14] J. Rikala and M. Kankaanranta, "Blending Classroom Teaching and Learning with QR Codes", *International Association for Development of the Information Society*, pp. 141-148, 2014.
- [15] M. Ludwig, J. Jesberg and D. Weiss, "MathCityMap - a smartphone project to do math", University Library Dortmund, 2013.
- [16] M. C. Wittmann, V. J. Flood and K. E. Black, "Algebraic manipulation as motion within a landscape", *Educational Studies in Mathematics*, 82(2), pp. 169-181, 2013. <https://doi.org/10.1007/s10649-012-9428-4>
- [17] L. C. Lederman, "Debriefing: Toward a systematic assessment of theory and practice", In: *Simulation & gaming* 23.2, pp. 145-160, 1992. <https://doi.org/10.1177/1046878192232003>
- [18] D. Dicheva and C. Dichev, "Gamification in Education: Where Are We in 2015?", In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1445-1454). Association for the Advancement of Computing in Education (AACE), 2015.
- [19] M. D. Dickey, "Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment", *British Journal of Educational Technology*, 42(3), pp. 456-469, 2011. <https://doi.org/10.1111/j.1467-8535.2009.01032.x>
- [20] M. Sigala, "The application and impact of gamification funware on trip planning and experiences: The case of TripAdvisor's funware", *Electronic Markets*, 25(3), pp. 189-209, 2015. <https://doi.org/10.1007/s12525-014-0179-1>



- [21] A. Iosup and D. Epema, "An experience report on using gamification in technical higher education", In Proceedings of the 45th ACM technical symposium on Computer science education pp. 27-32. ACM, 2014. <https://doi.org/10.1145/2538862.2538899>
- [22] M. B. Ibáñez, A. Di-Serio and C. Delgado-Kloos, "Gamification for engaging computer science students in learning activities: A case study", IEEE Transactions on learning technologies, 7(3), pp. 291-301, 2014. <https://doi.org/10.1109/TLT.2014.2329293>
- [23] M. B. Armstrong and R. N. Landers, "An evaluation of gamified training: Using narrative to improve reactions and learning", Simulation & Gaming, 48(4), pp. 513-538, 2017. <https://doi.org/10.1177/1046878117703749>
- [24] B. Morschheuser, J. Hamari, K. Werder and J. Abe, "How to gamify? A method for designing gamification" In Proceedings of the 50th Hawaii International Conference on System Sciences 2017. University of Hawai'i at Manoa, 2017. <https://doi.org/10.24251/HICSS.2017.155>
- [25] A. Lieberoth, "Shallow Gamification - Testing Psychological Effects of Framing an Activity as a Game" In: Games and Culture, 10, pp. 229-248, 2015. <https://doi.org/10.1177/1555412014559978>
- [26] Intrinsic Motivation Inventory (IMI): "The Intrinsic Motivation Inventory", Scale description, <http://selfdeterminationtheory.org/intrinsic-motivation-inventory/>, last accessed 2018/10/01.
- [27] I. Gurjanow and M. Ludwig, "Gamifying math trails with the Mathcitymap app: impact of points and leaderboard on intrinsic motivation", In: Aldon, G. & Trgalova, J. (Eds.): Proceedings of the 13th International Conference on Technology in Mathematics Teaching (ICTMT 13), pp. 105-112, Lyon, France, 2017.
- [28] J. Nakamura and M. Csikszentmihalyi, "The concept of flow. In Flow and the foundations of positive psychology", pp. 239-263, Springer, Dordrecht, 2014. https://doi.org/10.1007/978-94-017-9088-8_16
- [29] J. Zender and M. Ludwig, "The long-term effects of MathCityMap on the performance of German 15 year old students concerning cylindrical tasks", In: Proceedings of the eleventh congress of the European society for research in mathematics education (CERME11, February 6-10, 2019). Utrecht (in Press), 2019.
- [30] J. Chen, "Flow in games (and everything else)", In: Communications of the ACM 50.4, pp. 31-34, 2007. <https://doi.org/10.1145/1232743.1232769>
- [31] J. Bourgonjon, M. Valcke, R. Soetaert and T. Schellens, "Students' perceptions about the use of video games in the classroom", Computers & Education, 54(4), pp. 1145-1156, 2010. <https://doi.org/10.1016/j.compedu.2009.10.022>
- [32] Essential video gamenews http://www.sell.fr/sites/default/files/essential_video_game_news_sell_2018-eng-hr.pdf, p. 34, last accessed 2018/10/01.
- [33] P. Lucas and C. Martinho, "Stay Awhile and Listen to 3Buddy, a Co-creative Level Design Support Tool", In Eighth International Conference on Computational Creativity, ICC3, Atlanta, 2017.
- [34] J. B. Hauge, I. A. Stefan, A. Stefan, M. Cazzaniga, P. Yanez, T. Skupinski and F. Mohier, "Exploring Context-Aware Activities to Enhance the Learning Experience", In International Conference on Games and Learning Alliance pp. 238-247. Springer, Cham, 2017. https://doi.org/10.1007/978-3-319-71940-5_22
- [35] A. Lenhart, J. Kahne, E. Middaugh, A. R. Macgill, C. Evans and J. Vitak, "Teens, Video Games, and Civics: Teens' Gaming Experiences Are Diverse and Include Significant Social Interaction and Civic Engagement", Pew Internet & American life project, 2008.

Acknowledgements: This work was partially supported by the Erasmus+ European project MoMaTrE, EC Project Number: 2017-1-DE01-KA203-003577 and Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019.

