Gamification and Smart, Competence-Centered Feedback: Promising Experiences in the Classroom

[1]Michael D. Kickmeier-Rust, [1]Eva-C. Hillemann, [1]Dietrich Albert
[1,] Cognitive Science Section, Knowledge Technologies Institute
Graz University of Technology, Graz, Austria
{michael.kickmeier-rust; eva.hillemann; dietrich.albert}@tugraz.at

Abstract

Gamification appears being a promising approach to utilize the strong motivational potential of "gaming" in classroom without suffering from shortcomings such as low efficiency, weak pedagogy, or maybe most importantly the high costs. In the context of a European project we developed a rather light weight tool for learning and practicing multiplications. The target age group of the tool is 6 to 8 years. To benefit from the motivational potential of games we used a "gamification" approach. Accordingly we designed and developed a game-like, attractive user interface and integrated aspects of competition. The system is capable of providing students formative, competence-oriented feedback in real-time. Tailored to the age group this feedback is presented in form of a ninja character. For an experimental comparison of the effects of different feedback modes, we realized the conditions (i) no feedback, (ii) written only right/wrong feedback, (iii) audio right/wrong feedback, and (iv) competence-based, smart formative feedback. We applied and evaluated the tool in Austrian classrooms and found some evidence for the motivational aspect of the gamification elements, in particular the scoring. We also found strong positive effects of an individualized and meaningful feedback about achievements and progress.

Keywords*: Gamification, Adaptivity, Formative Feedback, Competence-based Knowledge Space Theory, Math Learning*

1. Introduction

1.1. Gamification in Educational Settings

Computer games in educational settings are becoming more and more popular, driven by the enormous motivational potential that is attributed to such games. The field, however, is tremendously challenging at the same time. Key questions address an effective and efficient adoption of modern computer games in the classrooms: How intensive, fast, or deep can be learned with and in a game? Which games are suitable for which purpose? Is there an added value at all? Such psycho-pedagogical considerations are – in all likelihood, accompanied by rather practical issues, for example concerning the available technical infrastructure (cf. also [6]). Finally, questions about a "return on investment" are crucial aspects of the success of serious games, meaning that the costs of implementing game-based educational activities must be justified by the educational outcomes – perhaps in terms of a deeper understanding, a more formative approach to assessing achievements and providing feedback, or perhaps in terms of reaching learners that are hard to engaging in learning otherwise.

One approach is to make use of the (comparably cheap but) enormous quality and motivational potential of existing computer games (the so-called commercial off-the-shelf games) or simulations for targeted educational purposes. Examples were reported by Kurt Squire [20] in the context of learning from simulation games such as *Civilization* or *Age of Empires*, by Constance Steinkuehler [21] in the context of massively multiplayer games, and for other game genres, for example, Maja Pivec [17], Sara de Freitas9], or David Shaffer [18].

Likewise, a recent trend is the concept of "*gamification*", which refers to the idea of utilizing game characteristics and game features for non-game applications in order to make them more fun, more engaging, and perhaps educationally more effective.

Among the most regularly and successfully utilized gamification features is *goal-setting* including progress paths and badges, awarding the player to identify goal completion. Ling [15] argues that the most motivating goals are those just out of comfortable reach and that this technique is most effective

when users can see their progress toward the end goal. Furthermore, people often increase engagement and efforts when they believe that they are close to a specific goal [11]. A related (or resulting) technique is providing *badges* – little but visible indicators of achievement, success, and ability – perhaps even status. Even if players never earn the badges, through viewing a set reachable and accomplishable challenges they come to understand valued activities within the system. Of course, it needs to be mentioned that tokenizing the achievements of players bears certain downsides, for example the avoidance of competition and fear of failure [16]. Other techniques are *leveling* (i.e., granting players access to new levels of the system – just novel interfaces, in its simplest case), *graphical enhancements* of the system, the use of (visually appealing) avatars, the implementation of additional *challenges and quests*, or the provision of *mini games* (such as board, card, or racing games) for diversion and recreation. Finally, an important aspect of gamification is sweepstakes, lotteries, and "real" *giveaway*s.

There is some evidence that the concept of gamification is a suitable and beneficial approach to foster learning motivation, engagement with a subject, and to improve the attitude towards a subject. On the other hand, the methods of gamification might bear certain risks and downsides. For example, in his book "*Punished by Rewards*" Alfie Kohn [14] argued that such approach to motivation may provoke more (perhaps too much) concentration on game-related achievements while producing lesser quality. In examples he demonstrated that children draw more pictures but in lesser quality when being paid for drawing pictures. More importantly, children did not like drawing pictures as much as before after they are stopped being paid [14]. Despite some risks, there is a clear trend towards the (hopefully cautious and well-planned) application of gamification in educational settings (cf. [3] for a critical review). Future work, however, must increasingly address the actual effects and benefits of gamification, in particular in school settings.

1.2. Formative Assessment and Feedback

The techniques and methods of gamification provide a natural link to the ideas of formative feedback, which is considered a key driver of successful education. Formative assessment and formative feedback focus on the goal of providing learners with constructive and helpful information about achievements, learning paths, learning pace, individual strengths and weaknesses in order to individualize and optimize learning activities. Especially new technologies are important in the field of formative assessment and feedback as they allow gathering large scale data, aggregating and analyzing them, and perhaps most importantly, to visualize and present the outcomes of analyses in a way that is most useful and beneficial for the learners. In this sense, formative assessment means identifying the current differences between current knowledge states and the educational target states of learners with the prime goal of promoting an effective competence development on an individual basis [19]. "Formative" means identifying ways to utilize the value of proper communication between learners and teachers, to strengthen an active role of learners, to optimize teaching/learning on an individual basis, and to acknowledge the psycho-social value of assessment/appraisal [7].

Certainly, such attempts are not new; teachers of all times have focused on supporting their students to the best possible extent and to bring them forward - to identify knowledge/competence gaps to inform learners and to facilitate a deeper understanding. In conjunction with the increasing uptake of new technologies in educational settings (ranging from electronic classroom and learning management systems to the use of mobile devices or electronic whiteboards), the amount of information available about students and their learning progress soars dramatically. Thus it is not surprising that the communities of learning analytics and educational data mining growing hand in hand with the state-of-the-art in formative feedback.

1.3. Formative Assessment and Intelligent Feedback in Gamified Systems

According to Gijbels and Dochy [12], key factors of formative assessment are: (i) appropriate, effective, and tailored feedback to learners, (ii) assigning responsibility for one's own learning, (iii) discovering the need of learners to evaluate and appraise themselves appropriately, (iv) adjusting the teaching activities according to the insights of assessment, and finally (v) acknowledging the motivational aspect inherent to assessment and appraisal and the related impact on self-esteem or perception. In autonomously acting, gamified systems it is necessary to provide learners with direct

and immediate feedback in a non-distracting, intelligent, meaningful, and formative way. This in turn requires equipping systems with the necessary "educational artificial intelligence, AI".

One approach that has been develop in the fields of adaptive tutoring systems and that has been further developed in the context of educationally intelligent serious games is *Competence-based Knowledge Space Theory* (CbKST), originally established by Jean-Paul Doignon and Jean-Claude Falmagne [8][10]. This approach is a well elaborated set-theoretic framework for addressing the relations among problems (e.g., test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. While the original idea considered on performance (the behavior; for example, solving a test item) only, extensions of the approach introduced a separation of observable performance and latent, unobservable competencies which determine the performance [1].

CbKST assumes a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competencies. A prerequisite relation states that competency *a* (e.g., to multiply two positive integers) is a prerequisite to acquire another competency *b* (e.g., to divide two positive integers). If a person holds competency *b*, one can assume that the person also holds competency *a*. To account for the fact that more than one set of competences can be a prerequisite for another competency (e.g., competency *a* or *b* are a prerequisite for acquiring competency *c*), prerequisite functions have been introduced, relying on and/or-type relations. A person's competence state is described by a subset of competencies she holds. Due to the prerequisite relations between the competencies, not all subsets are admissible competence states. By utilizing interpretation and representation functions the latent competencies are mapped to a set of tasks (or test items) covering a given domain. By this means, mastering a task correctly is linked to a set of necessary competencies and, in addition, not mastering a task is linked to a set of lacking competencies. This assignment induces a performance structure, which is the collection of all possible performance states. Recent developments of the conceptual framework are based on a probabilistic mapping of competencies and performance indicators, accounting for making lucky guesses or careless errors. This means, mastering a task correctly provides the evidence for certain competencies and competence states with a certain probability.

In the context of research projects such as *80Days* (focusing on educational games; www.eightydays.eu) or *Next-Tell* (focusing on evidence-centered, formative e-assessment; www.next-tell.eu) software services have been developed to equip autonomously acting, smart, and adaptive educational software with the aforementioned necessary educational AI. A recent development is the service platform *ProNIFA* (which stands for probabilistic, non-invasive, formative assessment engine); the engine retrieves educationally relevant performance data and updates the probabilities of the competencies and competence states in a domain. When a task is mastered, all associated competencies are increased in their probability, vice versa, failing in a task decreases the probabilities of the associated competencies. A distinct feature in the context of formative assessment is the multi-source approach. ProNIFA allows connecting the analysis features to a broad range of sources of evidence. The interpretation of the data occurs depending on a-priori specified and defined conditions, heurists, and rules which associate sets of available and lacking competencies to achievements exhibited in the sources of evidence. Very basically, the idea is to define certain conditions or states in a given environment (no matter if a Moodle test or a status of a problem solving process in a learning game). The specification of such state can occur in multiple forms, ranging from simply listing test items and the correctness of the items to complex heuristics such as the degree to which an activity reduced the 'distance' to the solution in a problem solving process (technically this can be achieved by pseudo code scripting). In essence, this approach equals the conceptual framework of micro adaptivity as, for example, described by Kickmeier-Rust [13].

2. The 1x1 Ninja: Gamifying Maths

In the context of the European research project Next-Tell we developed a rather light weight tool for learning and practicing the multiplication table. This tool, named *1x1 Ninja*, basically generates appropriate multiplication tasks that must be solved by the children. However, it incorporates a set of gamification features such as scoring to increase the children's motivation to practice the multiplication
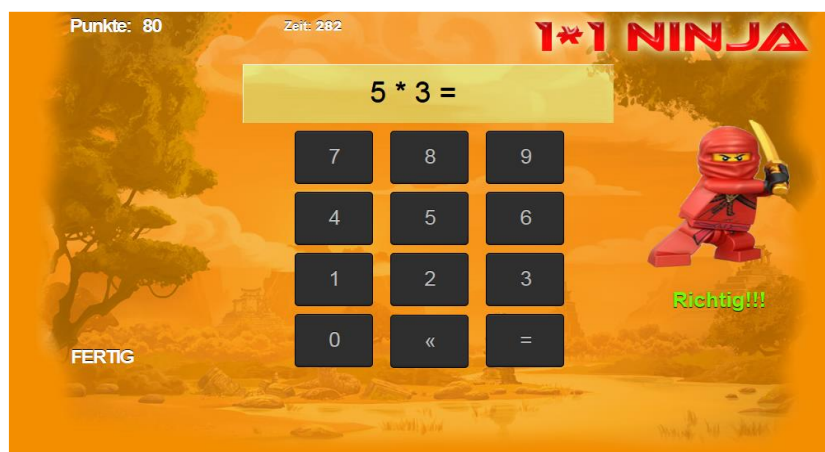
Figure 1. Screenshot of the 1x1 Ninja

table. The focus of this paper is to investigate the effects of gamification, on the one hand, and the effects of different levels of feedback on the other hand.

The tool and its features were developed in cooperation with the practice primary school of the Styrian Teacher Education Academy in Graz, Austria. *1x1 Ninja*, in the first instance, is based on the domain of the basic multiplication table ("*kleines Einmaleins*"). Together with math teachers we analyzed the domain established in total 22 skills involved in the domain and established a competence structure (in the CbKST sense). The skills cover the integers 0 through 10 both as multiplier and multiplicand. The online tool is design for being played using a tablet computer, however, it can also be played using smartphones or regular computers (Figure 1 shows a screenshots of the tool).

The multiplication tasks are generated randomly by the system, however, for educational reasons, each multiplicand (the second integer) is presented with all multipliers. In other terms, each multiplicand is performed for its entire multiplication table. As a means of gamification (in order to motivate the children), *1x1 Ninja* incorporates a scoring and a leveling feature. The scoring appears task wise, whereas the score for each task depends on the difficulty. A teacher can apply arbitrary scoring rules and weights. As an example, for the task 5 * 3, one can assign the value 10 for the multiplier '5' and a value 6 for the multiplicand '3'. The multiplicand might be weighted by a value 2. This scoring rule would result in a score of $10 + 2*6 = 22$ for this task. Depending on the scores and the individual setting various levels can be achieved. The exact scoring and weighting has been developed together with teachers according the curriculum and the teachers' experience. The levels are indicated by changing Ninja images and background images. Finally, a teacher can set session times, the maximum time allowed for multiplying. By default, we set a time slot of 5 minutes. Of course, a child can break a session earlier (using the "Fertig" link at the lower left of the window).

Depending on the individual settings, *1x1 Ninja* includes CbKST-based adaptive features. This means that the difficulty of tasks can be automatically and adaptively adjusted according to the abilities of a child. This means that when a child continuously fails with a certain multiplier or multiplicand, potentially easier tasks can be presented, or, alternatively, the occurrence of exactly such tasks can be increase in order to foster/enforce practicing exactly such integers.

A distinct feature of the tool is the feedback mechanisms. Depending on the settings the following feedback modes ore available:
- No feedback
- Visual correct/incorrect feedback (displayed as text below the image of the ninja)
- Visual and audio feedback (displayed below the image of the ninja and inform of a spoken statement such as "super", "correct", or "ups", "no", etc. The number of different feedback audio files is arbitrary and can be set by the teacher. Also, to the correct as well as the incorrect audio feedback, a probability can be assigned with which it occurs. This ensures that not each task results in a (maybe annoying or disrupting) feedback.

- Visual and audio feedback (as described above) and a CbKST-based formative feedback (providing certain summaries/explanations to the child about achievements and difficulties in case the likelihood of having a certain skill drops or passes a certain threshold; the feedback text can be set by the teacher).

Aim of this work is, as mentioned before, to look into the effects of gamification – even on a superficial level though and, more importantly, into the effects of the four different feedback levels. We were particularly interested in the effects on learning performance.

2. Method

2.1. Participants

The experimental study with school children comprised in total 58 persons from an Austrian primary school. The sample consisted of two classes of school children aged about 7 to 8 years.

The experiment was carried out in the context of class hours within a time period of two weeks following a task-based approach. First, pupils were introduced to *1x1 Ninja* and its functionalities. Subsequently, they were asked to accomplish five five-minute sessions of multiplication tasks for which the *1x1 Ninja* tool was used. Pupils' interactions with the tool were automatically logged and analyzed by the tool.

2.2. Materials

The software used for the experimental study described in this paper was the 1x1 Ninja, as described above. The multiplication tasks were generated randomly and adaptively (CbKST-based) by the system whereby each multiplicand is presented for its entire multiplication table. After solving a specific task, the scoring appears. This score depends on the difficulty of the task and is the basis for achieving various levels. provided to the student depends on the setting and is available in the following modes: i) no feedback, ii) visual correct/incorrect feedback, iii) visual and audio feedback, and iv)visual and audio feedback and a CbKST-based formative feedback.

The formative feedback, in this context, comprised informing the pupils about the progress and in particular about informing them about potential weaknesses. For example, if the tool detected that a specific child performed well for the number dimensions 4, 5and 6 but had a low performance for the number dimensions 7 and 9, the child received the feedback "Hey, you are doing well! But I noticed some difficulties for the number 7. Let's practice them a bit!" (said by the ninja character; cf. Figure 1).

As described in the previous section, the assessment is probability based. This means that not each incorrect answer triggers feedback. Only if the probability of certain competencies sinks below a given critical threshold, the systems provides the feedback and recommends actions (e.g., in our example the system might increasingly display tasks with the number 7). In addition, the system provides praise and cheer for the competencies with the highest likelihoods.

2.3. Study Design

In order to examine the effects of different feedback types on learning outcome, a 4x5 mixed design was employed consisting of the between-subject variable feedback type (no feedback vs. visual feedback vs. audio and visual feedback vs. audio, visual, and CbKST-based feedback) and the within-subject variable time of measurement (5 sessions/measuring times). In fact, the children used the *1x1 Ninja* freely and voluntarily during in total five free periods over three weeks as part of their regular school activities. There was no specific time limit; although the tool stops after a given limit of 5 minutes, the children could start again. Table 1 illustrates the experimental design and the number of subjects.

Table 1. A 4x5 mixed factorial design with the variations of the two variables
'Type of feedback' and 'Time of measurement'

Time of measurement

			Session 1	Session 2	Session 3	Session 4	Session 5
Type of feedback	No feedback	n	9	10	11 **(7)**	9	10
	Visual feedback	n	10	10	9 **(6)**	10	11
	Audio & visual feedback	n	11	16	11 **(7)**	17	9
	Audio, visual, CbKST-based Feedback	n	10	9	9 **(3)**	3	12

Note: values in parentheses refer to the number of those pupils, for which data sets were available
from all session phases. N refers to the number of pupils in a session.

3. Results

3.1. Tool Usage – Descriptive Results

Pupils completed on average 25 multiplication tasks ($M = 24.80$, $SD = 9.70$) in any session independent from the feedback they received. A high variation in the total amount of task solutions could be identified on the one hand with pupils only working on 10 tasks and on the other with students making quite extensive usage of the system with more than 40 task completions. In session 1 to 3 they worked on nearly 24 tasks ($M = 23.88$, $SD = 10.07$ for session 1; $M = 24.40$, $SD = 9.14$ for session 2; $M = 24.45$, $SD = 10.10$ for session 3). Regarding session 4 and session 5, participants completed on average 26 tasks ($SD = 9.00$, session 4) and 25 tasks ($M = 25.07$, $SD = 10.36$, session 5). With regard to the different feedback nodes presented to the learner, the absolute number of completed tasks per session ranges from 23.49 ($SD = 9.50$) for the group receiving no feedback and 26.93 ($SD = 9.50$) for students receiving audio feedback. Students getting a combination of audio and visual feedback completed on average 23.49 tasks ($SD = 10.42$). Presenting audio, visual and competence-based feedback resulted in nearly 25 tasks ($M = 24.56$, $SD = 8.59$) completed by pupils.

In each session, pupils spend on average 294.79 seconds ($SD = 21.08$) with a minimum duration length of 107 and a maximum length of 300 seconds. Having a look at the individual sessions and different feedback modes, only slightly variations in time spent could be identified with mean values ranging from 292.71 ($SD = 27.96$ for session 2) to 300 seconds ($SD = 0.00$ for session 4 and 5) and 292.23 ($SD = 31.97$ for feedback 3) to 198.47 ($SD = 10.06$ for feedback 4). A detailed overview of descriptive statistics is given in Table 2.

Table 2. Descriptive statistics (mean values and standard deviation) for time duration
and completed tasks for each session and for each feedback type

	Session 1	Session 2	Session 3	Session 4	Session 5
Time	293.55 (21.08)	292.71 (27.96)	300 (0.00)	300 (0.00)	288.38 (37.65)
Tasks completed	23.88 (10.07)	24.42 (9.14)	24.45 (10.10)	26.00 (8.98)	25.07 (10.36)

|	Feedback 1	Feedback 2	Feedback 3	Feedback 4	
Time	293.16 (24.50)	296.48 (17.65)	292.23 (31.97)	298.47 (10.06)	
Tasks completed	24.38 (9.85)	26.93 (9.50)	23.49 (10.42)	24.56 (8.59)	

After completing each session, pupils were asked to point out on a five-point-rating scale (ranging from 1 "strongly dislike" to 5 "strongly like") whether they like using *Ninja 1x1* or not. The likeability of the tool was judged with 3.96 (SD = 1.11; Median = 4), which indicates a good result. Half of the users scored the tool with on average 4.00 or better, on a rating scale ranging from 1-5 with higher values indicating a better result. This means that pupils like learning with this tool.

3.2. Learning Performance

A main aim of this study was to investigate whether pupils benefit from playing and working with the *Ninja 1x1* Tool. Additionally, on a more detailed level, the effect of different feedback types on learning performance was inquired.

In a first step, the means of correctly solved problems were calculated in order to make results of different pupils in different sessions comparable. A comparison of the first session and the last session, without taking into account the feedback type, showed that the total amount of correctly solved problems, the learning performance (M_{first} = 0.71, SD_{first} = 0.11; M_{last} = 0.83, SD_{last} = 0.09) increases over the different sessions (the t-test yielded t_{80} = -5.138, p = .000). When considering the results for the different feedback groups separately, a similar picture emerged: all mean values for all types of feedback are lower in the first session compared to mean values obtained in the last session. For feedback type 1 'no feedback' (M_{first} = 0.69, SD_{first} = 0.08; M_{last} = 0.82, SD_{last} = 0.07), 2 'visual feedback' (M_{first} = 0.69, SD_{first} = 0.08; M_{last} = 0.78, SD_{last} = 0.08), and 3 'CbKST-based feedback'(M_{first} = 0.71, SD_{first} = 0.19; M_{last} = 0.90, SD_{last} = 0.04) comparisons performance scores increased significantly (group 1: t_{17} = -3.646, p = .002; group 2: t_{19} = -2.978, p = .0008; group 4: t_{20} = -3.422, p = 0.003). For group 3 'auditive and visual feedback' (M_{first} = 0.75, SD_{first} = 0.08; M_{last} = 0.79, SD_{last} = 0.11), however, a non-significant tendency for better learning performance could be identified.

In order to find out more about differences between different types of feedback, classes, and measurement points, a repeated-measures analysis of variance was calculated. For these statistical analyses only those students, for which response sets were available for all 5 sessions, could be taken into account; concretely, in feedback condition 1 we had 7 pupils, in condition 2 were 6 pupils, in condition 3 were 7 and in feedback condition 4 were 3 pupils. The Analysis of Variance revealed a main effect of the session number with $F_{2.76}$ = 5.308, p = .005, that was not qualified by type of feedback ($F_{8.18}$ = 0.510, p = .845) and class ($F_{2.73}$ = 0.886, p = .448). There was also a main effect of type of feedback with F_3=7.00 (p = .04). These results are also illustrated in Figure 2.
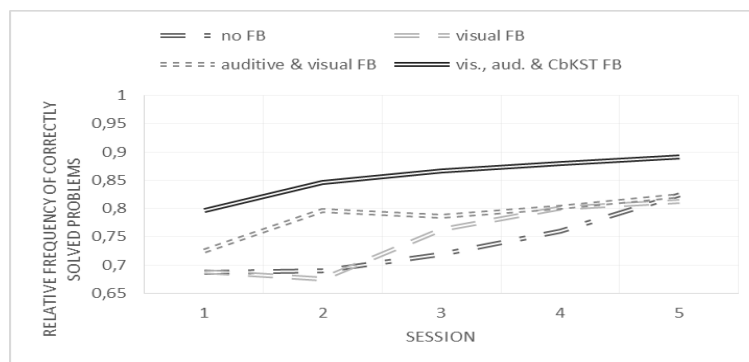
Figure 2. Profile diagram of the results for those students, for which values were available from all measuring points

Post hoc analyses using Tamhane's T2 indicated that the average value of learning performance was significantly higher in the feedback 4 condition with M = 0.86 (SD = 0.06) than were those in the feedback 1 condition with M = 0.74 (SD = 0.09; p = .013) and in the feedback 2 condition with M = 0.74 (SD = 0.08; p = .025). All other comparisons were non-significant.

4. Conclusion

The results of this study and the experiences made in the school classes indicate that the children much appreciated using the *1x1 Ninja* for practicing the multiplication table. Informal discussions with the children revealed that using the tool was more attractive and motivating than regular work on paper. This is remarkable to a certain extent because in fact the tool is not a game but incorporates very basic gamification elements such as scoring and the feedback by the ninja figure. One reason for these findings might also be the fact that children were not evaluated or monitored by a human teacher but got some performance feedback directly by the system. All in all, boys did rate the tool and the feedback features slightly better than the girls. One distinct difference was the rating of the scoring feature. The possibility to obtain high scores was much more liked by the boy – which confirms a gender cliché to a certain extend. We also observed that boys immediately started comparing the scores among them without being told to do so or without even mentioning the possibility to do so. Future work should gather more systematically information about the gamification features and their effects.

In the context of feedback, this study revealed very clear results in line with the findings described in the literature on the efficacy of providing feedback to improve learning performance and outcome (e.g., [2], [4], [22]). The current investigation sought to answer the question whether the type of feedback delivered is related to subsequent student performance. Bangert-Drowns and colleagues [4] stated based on a meta-analyses they conducted that the type of feedback is strongly related to effect size: "When feedback merely indicated that the response was correct or incorrect; it resulted in a lower effect than when the feedback in some way informed the learner of the correct answer" (p.232). Results of the study presented in this section support this assumption.

Overall, results of this experimental study indicate that children can benefit from using and playing with the tool as they can acquire knowledge regardless of the type of feedback delivered. In all four feedback conditions positive effects could be suggested. That is, improvement was observed on subsequent attempts regardless of the feedback type received which consequently suggest possible practice effects due to playing the game. However, significant differences could be identified in the learning process and outcome for individuals who received CbKST-based formative feedback relative to the other feedback types (i.e., no feedback, visual correct/incorrect feedback).

In general, it seems that providing some form of feedback is better than no feedback at all and students benefit more from receiving elaborate CbKST-based formative feedback relative to standard or regular feedback. This type of feedback can have a significant impact on learning performance as it allows for providing every student with immediate, individualized and elaborate feedback regarding their performance. Additionally to this, playing with the *1x1 Ninja* tool itself seems to improve students' learning performance. Thus, in any case, students benefit from "playing the game" or working with this tool.

In conclusion, we can emphasize that even minor and cost effective form of gamifying learning tools can boost motivation and engagement and that the described approach of providing smart feedback might provoke superior learning performance. This is in accordance with other similar studies, e.g. that of Bellotti and colleagues [5] Future classroom studies will broaden the basis of evidence and will allow more in-depth insights into the effects of gamification features and feedback.

5. Acknowledgements

6. References

[1] Albert, D., & Lukas, J. (Eds.). (1999). Knowledge spaces: Theories, empirical research, and applications. Mahwah, NJ: Lawrence Erlbaum Associates.

[2] Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). Journal of Organizational Behavior Management, 21, 3–29. http://dx.doi.org/10.1300/J075v21n01_02

[3] Anderson, J., Q., & Rainie, L. (May 18, 2012). Gamification: Experts expect 'game layers' to expand in the future, with positive and negative results: Online: http://www.pewinternet.org/~/media/Files/ Reports/2012/PIP_Future_of_Internet_2012_Gamification.pdf

[4] Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effects of feedback in test-like events. Review of Educational Research, 61, 213–238. http://dx.doi.org/10.3102/00346543061002213

[5] Bellotti, F., et al. (2013). A Gamified Short Course for Promoting Entrepreneurship among ICT Engineering Students. In Proceedings of Advanced Learning Technologies (ICALT) 2013. http://dx.doi.org/10.1109/ICALT.2013.14

[6] Bellotti, F., et al. (2013). Assessment in and of serious games: an overview. *Advances in Human-Computer Interaction*. http://dx.doi.org/10.1155/2013/136864

[7] Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), Assessment and learning (pp. 81-100). London: Sage.

[8] Doignon, J., & Falmagne, J. (1999). Knowledge Spaces. Berlin: Springer. http://dx.doi.org/10.1007/978-3-642-58625-5

[9] de Freitas, S., & Maharg, P. (Eds) (2010). Digital games and learning. London & New York: Continuum Press.

[10] Falmagne, J., & Cosyn, E., Doignon, J., & Thiéry, N. (2003). The Assessment of Knowledge, in Theory and in Practice. Institute for Mathematical Behavioral Sciences, UC Irvine, USA. Available online at http://www.scribd.com/doc/3155044/Science-Behind-ALEKS.

[11] Fox, S., & Hoffman, M. (2002). Escalation behavior as a specific case of goal-directed activity: A persistence paradigm. Basic and Applied Social Psychology, 24, 273–285. http://dx.doi.org/10.1207/S15324834BASP2404_3

[12] Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference? Educational Studies, 32(4), 399-409. http://dx.doi.org/10.1080/03055690600850354

[13] Kickmeier-Rust, M. D., & Albert, D. (2011). Micro adaptivity: Protecting immersion in didactically adaptive digital educational games. Journal of Computer Assisted Learning, 26, 95-105. http://dx.doi.org/10.1111/j.1365-2729.2009.00332.x

[14] Kohn, A. (1999). Punished by Rewards. The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes. Boston: Houghton Mifflin.

[15] Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., Cosley, D., Frankowski, D., Terveen, L., & Rashid, A.M. (2005). Using social psychology to motivate contributions to online communities. Journal of Computer-Mediated Communication 10, 1-10. http://dx.doi.org/10.1111/j.1083-6101.2005.tb00273.x

[16] Montola, M., Nummenmaa, T., Lucero, A., Boberg, M., Korhonen, H. (2009). Applying Game Achievement Systems to Enhance User Experience in a Photo Sharing Service. Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era on - MindTrek '09, 94. http://dx.doi.org/10.1145/1621841.1621859

[17] Pivec, M., & Kearney, P. (2007). Games for Learning and Learning from Games. Informatica, 31(4), 419-423.

[18] Shaffer, D. W. (2006). How computer games help children learn. New York: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230601994

[19] Smit, R. (2009). Formative Beurteilung und ihr Nutzen für die Entwicklung von Lernkompetenz [Formative assessment and ist value fort he development of learning competence]. Hohengehren: Schneider.

[20] Squire, K. (2003). Video games in education. International Journal of Intelligent Simulations and Gaming (2), 1.

[21] Steinkuehler, C., & Johnson, B. Z. (2009).Computational literacy in online games: The social life of a mod. The International Journal of Gaming and Computer Mediated Simulations, 1 (1), 53-65. http://dx.doi.org/10.4018/jgcms.2009010104

[22] Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. Learning and Instruction, 15, 589–602.http://dx.doi.org/10.1016/j.learninstruc.2005.08.001