Kickmeier-Rust, M. D., Marte, B., Linek, S. B., Lalonde, T., & Albert, D. (2008). The effects of individualized feedback in digital educational games. In T. Conolly & M. Stansfield (Eds.), Proceedings of the 2nd European Conference on Games Based Learning (pp. 227-236), October 16-17, 2008, Barcelona, Spain. Reading, UK: Academic Publishing Limited.

The Effects of Individualized Feedback in Digital Educational Games

Michael D. Kickmeier-Rust[1], Birgit Marte[1], Stephanie Linek[1], Tiphaine Lalonde[2] and Dietrich Albert[1]
[1] University of Graz, Austria
[2] ORT France, Paris, France
michael.kickmeier@uni-graz.at
birgit.marte@uni-graz.at
stephanie.linek@uni-graz.at
tiphaine.lalonde@ort.asso.fr
dietrich.albert@uni-graz.at

Abstract: A crucial factor for successful digital educational games, particularly for older children and adolescents, is an appropriate balance; balance between learning and gaming and balance between challenge and ability. These factors are important to maintain fun, immersion, flow experience, and motivation – the motivation to play and therefore to learn. Moreover, it is important to realize a gaming experience that can compete with that of commercial, non-educational games. A special challenge in this context arises from the need for pedagogical support during learning - and therefore during gaming. At many staves of the learning ladder, from a psych-pedagogical perspective, support and feedback is necessary in order to ensure successful, effective, and complacent learning. Considering the importance of not destroying immersion with the game, the assessment of the learning progress and psycho-pedagogical feedback must occur in a non-invasive way. This, however, requires an intelligent system that is capable of assessing individual competencies and learning progress by observing and interpreting the learner's behaviour in the learning situations within the game. In ELEKTRA, a project funded by the European Commission and aiming at developing a sound psycho-pedagogical framework for immersive educational games, we developed a formal cognitive framework for the non-invasive assessment and interventions within complex learning situations, that is, micro adaptivity. Attuned to the assessed competencies or lack of competencies, meaningful feedback, for example hints, suggestions, reminders, critical questions, or praise, can be triggered, without destroying the gaming experience. Two questions arise with respect to feedback. First, does feedback, although designed to be non-invasive, on educational issues impair gaming experience? Second, can feedback in gaming situations facilitate the learning progress or does it increase the learner's cognitive load, which was suggested be several researchers. In the context of the ELEKTRA project, we implemented the theoretical framework of micro adaptivity in the game demonstrator. This demonstrator is a state-of-the art 3D adventure game teaching physics in relation to national school curricula for the age group of 12 o 14 years. For evaluation purposes, log files of the gaming sessions were recorded and, in addition, questionnaires and performance tests were presented. In this work, we present results from an evaluation session. The results indicate that (micro) adaptive interventions (i.e., appropriate and meaningful interventions/feedback for an individual learner, his/her knowledge and learning progress) are superior to neutral (i.e., non-individualized but semantically correct interventions) and inappropriate interventions (i.e., non-individualized, unsuited interventions) in terms of learning and gaming measures. In addition, we analysed the relationships between learning progress and socio-emotional variables. The results indicate that adaptive feedback not only facilitates learning but also attitude and immersion.

Keywords: Technology-Enhanced Learning, Game-based learning, Feedback, Didactic Interventions, Micro Adaptivity, Evaluation

1. Introduction

A crucial factor for successful digital educational games (DEGs), particularly for those, targeting at older children and adolescents, is an appropriate balance; a balance between learning and gaming and a balance between challenge and ability (in terms of gaming as well as learning. It is important to maintain fun, immersion, flow experience, and motivation – the motivation to play and therefore to learn. Moreover, it is important to realize a gaming experience that can compete with that of commercial, non-educational games.

Thus, successful DEGs must be able to adapt to the learner's knowledge, skills, and abilities, motivation, and also to pedagogical implications. In traditional forms of technology-enhanced learning, concepts of adaptivity, adaptability, and personalization are increasingly important. Generally, adaptive approaches to e-learning contest the one-fits-all approach of traditional learning environments, trying to tailor the learning environment according to individual needs and preferences.

Adaptivity refers to navigation, curriculum sequencing, and presentation. For example, an adaptive system may only provide learning objects which are suitable for an individual's learning progress - learning objects either too difficult or too easy might not be displayed in order to avoid visual and cognitive load and to suggest an appropriate learning path through the learning content.

In the context of immersive digital games, existing approaches to adaptivity must be extended in order to maintain an immersive gaming experience, motivation, and probably flow experience by suitable adaptive interventions (Kickmeier-Rust, Hockemeyer, Albert and Augustin 2007). A special challenge in this context arises from the need for pedagogical support during learning - and therefore embedded in gaming. At many staves of the learning ladder, from a psych-pedagogical perspective, support and feedback is necessary in order to ensure successful, effective, and complacent learning. Considering the importance of not destroying immersion, flow, and engagement in the game, the assessment of the learning progress and psycho-pedagogical feedback must occur in a non-invasive way. This, however, requires an intelligent system that is capable of assessing individual competencies and learning progress by observing and interpreting the learner's behaviour during the learning situations within the game.

The work presented here primarily grounds on the ELEKTRA project (www.elektra-project.org), which was a multi-disciplinary research and development project, running from 2006 to 2008, funded by the European Commission. It had the ambitious goal to utilize the advantages of computer games and their design fundamentals for educational purposes and to address disadvantages of game-based learning as far as possible. Within the project a methodology for successful design of educational games has been established and a game demonstrator was developed based on a state-of-the-art 3D adventure game teaching optics according to national (i.e., French, Belgian, and German) curricula (Kickmeier-Rust, Peirce, Conlan, Schwarz, Verpoorten and Albert 2007). More importantly, ELEKTRA addressed research questions concerning data model design as basis for adaptivity and resource description enabling interoperability of systems as well as the data model itself (Kickmeier-Rust and Albert 2007). In the course of the project, an approach to adaptivity, that is, *micro adaptivity*, was developed that allows assessing learning performance and cognitive states in a non-invasive way by interpreting the learners' behaviour within the game and by responding on the conclusions drawn from their behaviour (Albert, Hockemeyer, Kickmeier-Rust, Peirce and Conlan, 2007). Attuned to the assessed competencies (or lack of competencies), meaningful feedback, for example hints, suggestions, reminders, critical questions, or praise, can be triggered, without destroying the gaming experience.

Two questions arise with respect to interventions and feedback. First, do interventions/feedback, although designed to be non-invasive, on educational issues impair gaming experience? Second, can interventions in gaming situations facilitate the learning progress or do they increase the learner's cognitive load, which was suggested be several researchers. In the context of the ELEKTRA project, we implemented the theoretical framework of micro adaptivity in the game demonstrator. This demonstrator is a state-of-the art 3D adventure game teaching physics in relation to national school curricula for the age group of 12 to 14 years. For evaluation purposes, log files of the gaming sessions were recorded and, in addition, questionnaires and performance tests were presented.

In this paper, we present the results from an evaluation session with French students. The results indicate that (micro) adaptive interventions and feedback (i.e., appropriate and meaningful interventions for an individual learner, his/her knowledge and learning progress) is superior to neutral interventions (i.e., non-individualized but semantically correct interventions) and inappropriate interventions (i.e., non-individualized, unsuited interventions) in terms of learning and gaming measures. In addition, we analysed the relationships between learning progress and socio-emotional variables. The results indicate that adaptive interventions not only facilitate learning but also attitude and immersion.

2. Assessment on a micro level

An essential of current computer games is a meaningful interaction of player and game. This is achieved by something called "Game AI", the artificial intelligence of the game. This AI, for example, makes opponents being a serious challenge within a battle; it makes them behave properly and not always the same. In recreational computer games, the "intelligent" mechanisms are sometimes more, sometimes less satisfying. Educational games that can compete with their non-educational counterparts must not only implement traditional Game AI, but also an AI for educational purposes.

And in educational games, such intelligent mechanisms may be even more important than in non-educational games. The main task for intelligent and adaptive mechanisms is to guide and support the learner in acquiring competencies and skills, for example by teaching the learner, by intervening when misconceptions occur or the learning progress is unsatisfactory, by hinting, or by providing the learner with feedback. In addition, tasks are motivating, maintaining immersion, and personalizing the game according to the preferences and needs of the learner.

To achieve such tasks, a theoretical and technological approach is required that enables the game to assess cognitive states (e.g., competence states or motivational states), learning progress, possible misconceptions, or undirected/unsuccessful problem solving strategies. In contrast to traditional forms of teaching (either in real or virtual environments), where the assessment can occur by test items, questions, or tasks, DEGs require an assessment that does not destroy or impair motivation, immersion, flow experience, or the game's storyline (Kickmeier-Rust and Albert 2007, Kickmeier-Rust, Hockemeyer, Albert and Augustin 2007).

2.1 Merging CbKST and Problem Spaces

To accomplish the ambitious goal of non-invasive assessment and subsequent educational interventions, we combined *Competence-based Knowledge Space Theory* (CbKST), which has been successfully utilized in conventional adaptive, personalized e-learning, and cognitive theories of problem solving. CbKST provides a detailed domain model that includes a set of meaningful competence states as well as a set of possible learning paths.

Very briefly, CbKST is an extension of the originally behavioral *Knowledge Space Theory* (Doignon and Falmagne 1985, 1999), where a knowledge domain Q is characterized by a set of problems. The knowledge state of an individual is identified on the subset of problems this person is capable of solving. Due to mutual dependencies between the problems captured by prerequisite relations, not all potential knowledge states will occur. The collection of all possible states is called a knowledge structure K. To account for the fact that a problem might have several prerequisites (i.e., and/or-type relations) the notion of a *prerequisite function* was introduced. The basic idea of CbKST is to assume a set S of abstract skills underlying a domain of knowledge. The relationships between the skills and problems are established by a *skill function*. Such function assigns a collection of subsets of skills (i.e., *competence states*) to each problem, which are relevant for solving it. By associating skills to the problems of a domain, a *knowledge structure* on the set of problems is induced. The skills, which are not directly observable, can be uncovered on the basis of a person's observable performance. A further extension is to assume prerequisite relationships between the skills, inducing a competence structure C on the set of skills (Korossy 1999). To illustrate this approach, assume that a knowledge domain is represented by $Q=\{a, b, c, d\}$. Consider the set $S=\{V, W, X, Y, Z\}$ of skills that are relevant for solving them. A prerequisite function that might exist among these skills is demonstrated in Figure 1a. For example, this function reads that if a student has skill X we can assume that this student also possesses either skill V or W, or both; the corresponding competence structure is shown in Figure 1b. It includes only 13 possible competence states from a total of $2^5 = 32$ states.

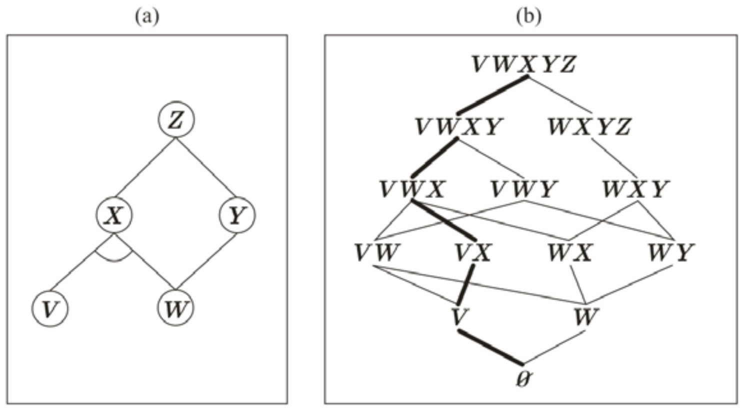

Figure 1: The left panel illustrates a prerequisite function (the bended line below skill X indicates a logical or). The right panel shows the corresponding competence structure. The bolded line indicates one of several meaningful learning paths.

To achieve a non-invasive assessment, we developed a formal model of the problem solving behaviour in game-based learning situations (LeS). Basically, LeS are characterized by a large degree of freedom and complex problem solving demands. The problem solution process is considered to be a meaningful sequence of problem solution states establishing a problem space (Simon 1978). Stochastic process models are applied in order to estimate the probabilities of certain state transitions and to estimate the probability of reaching a solution state (within a specific time interval). In other terms, a LeS is segmented in to a set of possible problem solution states (you may think about all possible states of the *Tower of Hanoi* problem). Each of those problem solution states is mapped, via an ontology, to one of a set of possible competence states (cf. Figure 1b). By this means, the game can interpret the behaviour of the learner in terms of available knowledge, un-activated knowledge, or missing knowledge, by mapping the actions of the learner to competence states (see Kickmeier-Rust, Hockemeyer, Albert and Augustin 2007 for details).

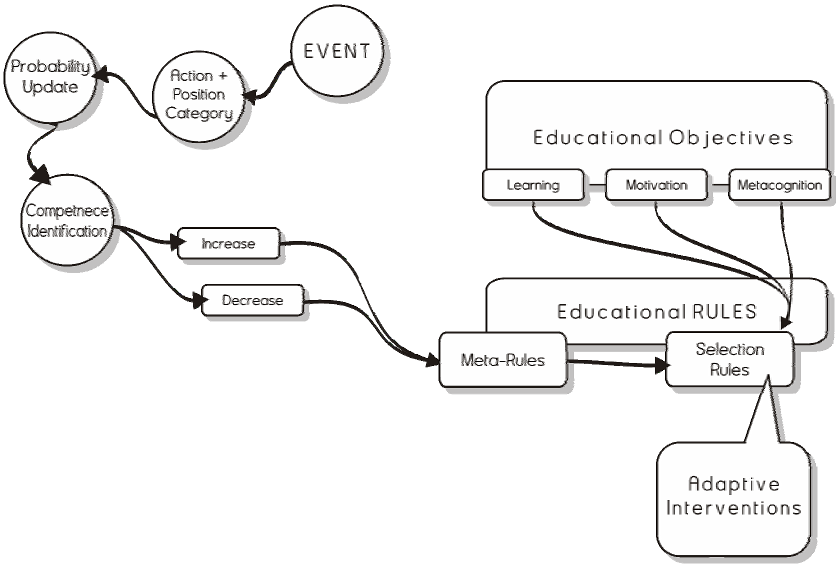

Figure 2: Sketch of the micro adaptive assessment and intervention process for educational games.

2.2 The micro adaptive assessment and intervention process

As summarized in Figure 2, the overall micro adaptive assessment and intervention process is imitated by any action (event) the learner is performing in the game (e.g., by switching on a torch). The situation after such event is analyzed in terms of the given problem solution state and, subsequently, the probability distribution over all competence states is adjusted to the problem solution state. By the probability change of specific competencies involved in a situation (e.g., knowing that the torch's light is necessary), the most relevant/critical competencies can be detected. Depending on an increase (what actually is desired) or a decrease of the probability of specific competencies, pedagogical/didactic meta-rules are utilized to select a specific interventions and feedback (e.g., 'if the probability of a competence v involved in a LeS decreases below a threshold w, and the probability of a competence x is above a value y, then trigger an educational hint z').

2.3 Architecture

From a technical perspective, the architecture consists of four modules or engines (Figure 3). The learner is connected to the system through the game engine (GE). It provides the non-adaptive parts of the game, and as such it is also the user interface to the system. The GE provides information on the learner's action in the game to the skill assessment engine (SAE). The SAE updates the learner model (i.e., the competence state probabilities) according to aforementioned process and additional information from an ontology. The resulting information about the learner's competence state and its changes are then forwarded to the Educational Reasoner (ER), the pedagogical part of micro adaptivity. Based on pedagogical rules (e.g., the diversification principle) and learning objectives (e.g., the straight propagation of light), the ER gives recommendations on adaptive interventions to the adaptation realization module which maps the abstractly formulated educational recommendations onto more concrete game recommendations. In this mapping process, data on game elements and information on previously given recommendations are considered. The game recommendations are then forwarded to the GE which realizes them as concrete adaptive interventions in the game.

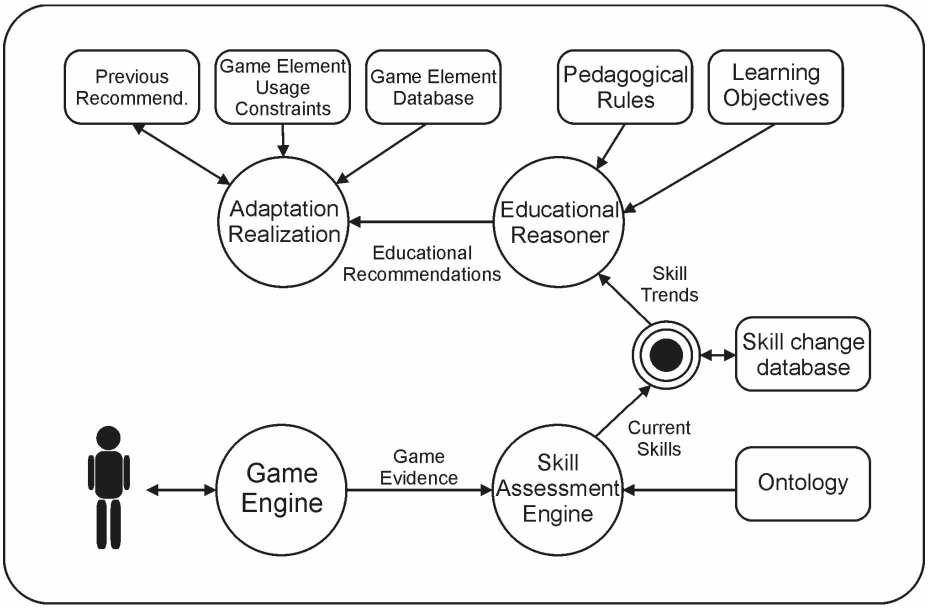

Figure 3: An architecture for micro adaptive assessment and interventions.

2.4 Basic types of interventions and feedback

The conditions under which a certain adaptive intervention is given are developed on the basis of pedagogical rules; however, these rules apply the micro adaptivity framework and utilize the learner model obtained through the assessment within the framework. At the present stage, following basic types of feedback/interventions are implemented:

- *Skill activation interventions* may be applied if a learner gets "stuck" in some area of the problem space and some skills are not used although the user model assumes that the user masters these skills.
- *Skill acquisition interventions* may be applied in a similar situation, however, the user model assumes that the user does not master the unused skill.
- *Motivational interventions* may be applied, for example, if the learner does not act for a certain, unexpected long time.
- *Assessment clarification interventions* may be applied, for example, in form of a query if the learner's actions give contradicting support for and against the assumption of a certain competence state.

3. Interventions and Feedback – Scoring Well or Badly?

Pedagogical interventions and feedback are important aspects in educational settings. They guide the learning process, inform the learner about the learning progress and possible deviations from a planned learning path, and they aim to provide the learner with appropriate information and direct the learner's view on important information. Generally, feedback on learning comes from teachers, fellow students, friends, or from oneself (then it is called reflection). Also in the context of computer-based learning, studies demonstrated the effectiveness of feedback (cf. Azevedo and Bernard 1995). Moreno (2004), for example, used feedback to decrease cognitive load of learners in discovery-based learning environments. Her study compared the effects of explanatory feedback by an artificial agent with corrective feedback. Her results show that explanatory feedback is better in terms of error or learning performance. In their work regarding teachable agents, Leelawong et al. (2002) presented an alternative meaning of feedback in an educational setting. Those authors developed an avatar, *Betty*, that is able to 'learn' and that must be 'taught' by the students. The idea behind grounds on the body of research that suggests that tutors benefit as much from their tutoring as the tutees (e.g., Chi et al. 2001). Based on the feedback Betty gives, the students have to change their tutoring and, as a consequence, their own knowledge. Aleven, McLaren, Roll and Koedinger (2004) demonstrated that in settings were students were required to explain their steps and received feedback on that showed a deeper learning. In a recent study, Tan, Biswas and Schwartz (2006) showed that directed feedback could improve immediate goal achievement, meta-cognitive feedback better prepared students for future/other learning tasks, even in situations without meta-cognitive feedback.

However, immediate feedback in ‚flow' situations, even if it is appropriate from a didactic point of view, may hold a highly disruptive potential and, therefore, may be inimical to learning. A study by Schmidt, Young, Swinnen and Shapiro (1989), which investigated the effects of the timing of feedback in a learning context, revealed that shorter feedback latencies were beneficial for learning performance while feedback was present, in turn, delayed feedback resulted in improved performance once feedback had been withdrawn. Those authors attempted to explain this phenomenon with a *guidance hypothesis*, meaning that feedback guides the learner towards a better performance, however, induces some dependence to feedback by obscuring the need to acquire secondary skills or competencies. An example is the ability to detect and self-correct mistakes or errors in the learning process. Also a study by Lewis and Anderson (1985) within the context of an adventure game provided some support for the guidance hypothesis. Immediate feedback resulted more likely in the selection of appropriate problem solving strategies in the game, whereas delayed feedback resulted in better error detection. Anderson, Conrad and Corbett (1989) suggested that learners receiving immediate feedback go faster through the learning content than did those receiving delayed feedback; however, they did not find significant differences in learning performance. Schooler and Anderson (1990) concluded that there are advantages to immediate feedback, for example, it focuses the learners' attention on important information and assures the relevant information is perceived consciously (therefore being processed in working memory). The disadvantages are basically seen in the decreased necessity of acquiring secondary skills and competencies. From the perspective of cognitive load, processing the feedback information requires (limited) cognitive resources, which may deducted them from the learning process and, as a consequence, impair learning.

4. An empirical investigation of intervention/feedback effects in a DEG

In the context of the ELEKTRA project, we investigated the effects of different types of interventions and feedback with French students playing the demonstrator game. Essentially, the demonstrator is based on a classical 3D adventure game in first-person view. The aim is to salvage the girl Lisa and her uncle Leo who have been kidnapped by the evil Black Galileans; moreover, the learner has to avert that those evil forces possess to entire world. During this journey, the learner needs to acquire specific, curriculum-related knowledge and skills, concretely, the learner learns about 8th grade optics. The learning occurs in different ways, ranging from hearing or reading to freely experimenting. After finding a magic hour glass, the learner is in company of the ghost of Galileo Galilei, who is the learner's (hidden) teacher. In addition, the learner can interact with Lisa via a headset, which is indicated in the upper left corner of the screen. Those non-playing characters also play a significant role for intelligent, non-invasive educational and motivational interventions. For example, Galileo tells the learner specific facts, which are need for certain events in the game, or he intervenes by providing the learner with certain hints or feedback. Figure 4 gives some impressions about the demonstrator game.

In contrast to previous research on the (educational) effects of interventions and feedback, as briefly outlined above, the focus in the present work was on the effects of adaptive, personalized, highly appropriate, and timely interventions in comparison to no interventions, neutral interventions, or even inappropriate and ill-suited interventions.

4.1 Experimental design

The scientifically sound comparison of feedback effects in a highly adaptive and individualize context, as aspired with micro adaptivity, is highly complex. The main idea of adaptation and personalization is that each of the learners/players/participants receives entirely different interventions, tailored to individual knowledge as well as learning and gaming progress. The consequence for the experimental design is that an exact comparison of different participants is not possible.

To address that problem, we utilized a so-called *yoked control* design (Harmatz and Lapuc, 1968) in the demonstrator. Yoked control is based on the idea that comparable pairs of participants are actively generated. A first participant receives a treatment (e.g., an intervention) that is adaptively tailored to his/her behaviour. A second participant is artificially linked to the first by receiving exactly the same treatment as the first one, of course, now the treatment is entirely independent from the participant's behaviour. From the experimental perspective, both participants did now receive identical and, therefore, comparable gaming/learning experiences. The necessary event logging and replay functions have been implemented in the demonstrator game.

Figure 4: These are four screenshots from the ELEKTRA demonstrator game. The game starts outside a villa near a science park (upper left image). In the villa the learner faces, among others, the task to open a solid metal door, which requires some knowledge about the propagation of light (upper right image). The ghost of Galileo Galilei is the learner's accompanying mentor and teacher (lower left image). To acquire knowledge and to precede through the game the learner is doing specific experiments, supported by Galileo and, via a headset indicated in the upper left corner of the screen, Lisa (lower right image).

Yoked control assures that a matched pair of participants received exactly the same interventions and feedback independent from their actions and their progress in the game. However, yoked control cannot avoid that the second participant receives an intervention or feedback that is by chance suitable for the situation. For example, the feedback 'well done' is suitable in many situations. To address this problem, statistical analyses were based on four types of interventions/feedback based on the data gathered by yoked design experiments:

- *Appropriate interventions/feedback* (i.e., statements that are beneficial for the learning or gaming progress, for example, "remember what I told you about the propagation of light" when trying to open a door by hitting a light sensor with a narrow beam of light; Figure 4).
- *Neutral interventions/feedback* (i.e., statements that are always suitable and that do not have much positive or negative effects, for example, "keep a stiff upper lip")
- *Inappropriate interventions/feedback* (i.e., statements that have a negative or at least confusion impact on learning or gaming, for example, "remember what I told you about the propagation of light" when actually solving the Tower of Hanoi problem).
- *No interventions/feedback*

To identify the categories, we performed extensive log file analyses for the yoked control group, comparing the received interventions/feedback with the actual behaviour. The manual classification work was performed by two raters independently; the inter-rater reliability was .81.

4.2 Materials
The basis of the present analyses is the so-called “slope device” situation (Figure 5). In this LeS the students experiment with a machine where several balls of different materials (solid and hollow iron, wood, and plastic) are running down a slope and also a laser can beam down this slope. This machine has a fan and a strong magnet. The learners’ task is to make the balls fall into a hole by setting appropriate values for fan and magnet. In addition they should estimate the trajectory of the laser beam in dependence fan, gravity, and magnetic force. This experiment should visualize the effects of fan, gravity, and magnet on different material and, first of all, that the laser beam is not influenced by such external forces and independently propagates in a straight line. The approach to solution value indicates how fast a learner finds the correct settings of fan and magnet and how well s/he can estimate the trajectory of the laser beam.

Figure 5: The “slope device” experiment.

4.3 Participants
Participants were 40 school students recruited at two schools in Paris; 17 were female, 23 male. The average age was 13.08 years (SD = 1.08). By far the largest group (i.e., 90%) of the children were familiar with computer games, playing about 6.01 hours (SD =8.88) a week.

4.4 Results
In a first step, we analysed how fast and how well the learners could accomplish the slope device experiments, depending on the type of interventions/feedback they received. This is probably the most meaningful perspective to the gaming behaviour since this kind of support is the major purpose of the interventions. In this context we distinguished two measures. First, the so-called ‘approach to solution’ variable, which states how many action were performed following a certain type of intervention/feedback that were (a) closer to the final solution, (b) farther from that, or (c) without an effect. The value of this variable depends on the number of interventions of a type each learner received. Second, we analyzed the response time that is, the time the learners needed after receiving an intervention/feedback to perform their next actions in the experiments. Since this type of analysis compares intervention/feedback types and not participants (each of them got several of different types), the total experimenting time is not a meaningful measure.

The results of these analyses are summarized in Figure 6. Appropriate interventions/feedback resulted in an average approach to the correct solution of 4.95 (SD = 18.37), neutral in an average approach of 3.69 (SD = 16.31), inappropriate in an average approach of 4.00 (SD = 15.21), and not receiving any interventions or feedback resulted in an average approach of 3.76 (SD = 14.30). These differences are statistically not significant. However, they clearly indicate that appropriate interventions/feedback result in a quicker problem solving progress that needs fewer steps. Somewhat different results were found for the response times after each intervention/feedback. Appropriate interventions/feedback resulted in an average response time of 3.90s (SD = 1.16), neutral in an average response time of 4.03s (SD = 1.08), inappropriate in an average response time of 3.94s

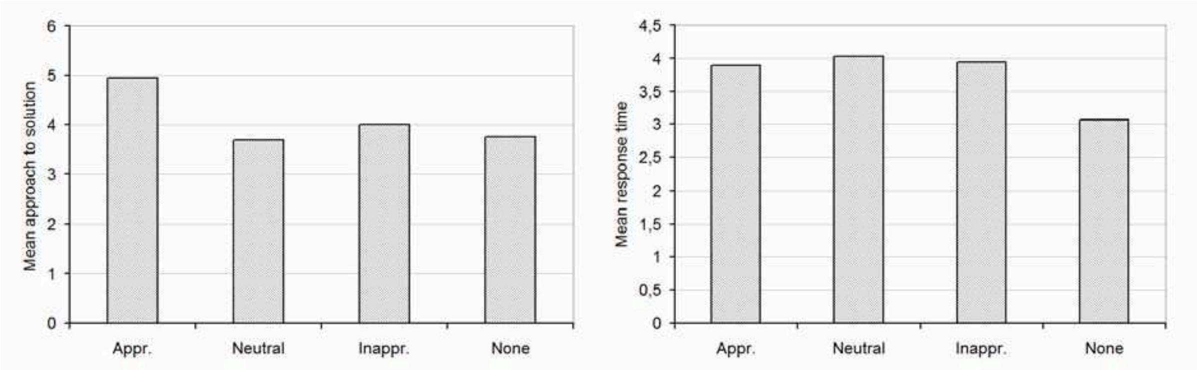

Figure 6: The left panel illustrates the average approach to the correct solution of the slop device experiments. The right panel illustrates the response time after receiving an intervention/feedback.

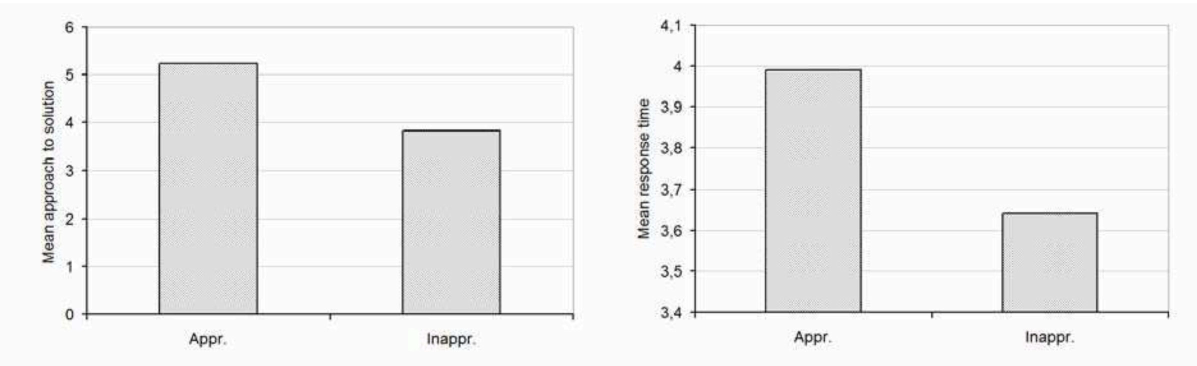

Figure 7: The left panel illustrates the average approach to the correct solution of the slop device experiments for extreme groups with a high and a low amount of inappropriate interventions/feedback. The right panel illustrates the corresponding response time.

(SD = 0.84), and not receiving any interventions or feedback resulted in an average response time of 3.06s (SD = 0.90). An analysis of variance (ANOVA) yielded that receiving no interventions or feedback resulted in statistically significant shorter response times ($F(3)=33{,}86$; $p<01$) than receiving interventions or feedback; the type of feedback, however, did no influence response times.

In addition to the analysis on the intervention type level, we performed analyses on the learner level. We compared the average approach to the correct solution and the average response time for participants who received (almost) no inappropriate interventions and feedback with such participants who received a large portion of inappropriate interventions. The reason for analyzing extreme groups is that, due to the highly adaptive nature of the game and also due to the yoked control design, the distribution of specific types of interventions is flowing and does not allow clearly identifying specific groups on the person level. The extreme groups included 10% of participants who had received the most inappropriate interventions and the least inappropriate interventions respectively. The results are summarized in Figure 7. The average approach to the correct solution was 5.23 (SD = 14.17) in the appropriate intervention extreme group and 3.83 (SD = 16.20) in the inappropriate intervention extreme group. Similarly, the average response times were 3.99s (SD = 0.91) in the appropriate intervention extreme group and 3.64s (SD = 1.02) in the inappropriate intervention extreme group. According to an ANOVA, the differences between the extreme groups were statistically significant for both approach to solution ($F(1)=0{,}31$, $p<0{,}01$) and response time ($F(1)=5{,}05$; $p<0{,}05$).

So far, analyses focussed on the participants' behaviour within the game and on how quickly and how well they could handle the problems of the slope device experiments. In addition to that, a major question is if the participants learned what they were supposed to learn with the slope device LeS and how well they performed. Thus, in a next step we analysed learning performance depending on aforementioned extreme groups. From a general multiple choice learning test, covering the learning objectives of the entire demonstrator game, eight items were related to the slope device experiments

Table 1: Descriptives for the knowledge test regarding the slope device learning objectives.

		Man solution frequency	Std. Deviation	95% Confidence Interval for Mean (lower, upper)	
Q1	Appropriate Interventions EG	,8000	,44721	,2447	1,3553
	Inappropriate Interventions EG	,6000	,54772	-,0801	1,2801
	Total	,7000	,48305	,3544	1,0456
Q2	Appropriate Interventions EG	,6000	,54772	-,0801	1,2801
	Inappropriate Interventions EG	,6000	,54772	-,0801	1,2801
	Total	,6000	,51640	,2306	,9694
Q16a	Appropriate Interventions EG	1,0000	,00000	1,0000	1,0000
	Inappropriate Interventions EG	,4000	,54772	-,2801	1,0801
	Total	,7000	,48305	,3544	1,0456
Q16b	Appropriate Interventions EG	,4000	,54772	-,2801	1,0801
	Inappropriate Interventions EG	,4000	,54772	-,2801	1,0801
	Total	,4000	,51640	,0306	,7694
Q16c	Appropriate Interventions EG	,2000	,44721	-,3553	,7553
	Inappropriate Interventions EG	,4000	,54772	-,2801	1,0801
	Total	,3000	,48305	-,0456	,6456
Q16d	Appropriate Interventions EG	,0000	,00000	,0000	,0000
	Inappropriate Interventions EG	,2000	,44721	-,3553	,7553
	Total	,1000	,31623	-,1262	,3262
Q16e	Appropriate Interventions EG	,2000	,44721	-,3553	,7553
	Inappropriate Interventions EG	,2000	,44721	-,3553	,7553
	Total	,2000	,42164	-,1016	,5016
Q16f	Appropriate Interventions EG	,2000	,44721	-,3553	,7553
	Inappropriate Interventions EG	,0000	,00000	,0000	,0000
	Total	,1000	,31623	-,1262	,3262

(cf. section 4.2). Table 1 lists the descriptive values. Summarized over all items, the appropriate interventions extreme group performed clearly better (45% correctly solved items) than the inappropriate interventions extreme group (35% correctly solved items). This performance was also correlated with gaming duration; the longer the participants played with the demonstrator game, the better was their test performance ($r = .351$; $p = .029$).

Finally, we analyzed overall learning outcomes with the demonstrator game with and without interventions/feedback using the 34 item knowledge test before and after playing the demonstrator. The results are summarized in Table 2. The group with adaptive interventions clearly performed better in the knowledge test than the group without any interventions.

Table 2: Average learning performance (numbers in parentheses indicate the standard deviations).

		Adaptive Interventions	No Interventions
Pre test	Learning outcome	9,60 (3,49)	8,28 (3,07)
	I don't know answers	3,40 (2,22)	4,46 (3,02)
	Incorrect answers	8,00 (3,23)	8,26 (2,72))
Post test	Learning outcome	10,60 (4,04)	9,11 (3,95)
	I don't know answers	3,15 (2,37)	3,68 (3,55)
	Incorrect answers	7,25 (3,56)	8,20 (3,77)
	Learning performance	1,00 (2,45)	0,83 (3,91)

5. Discussion

DEGs that can compete with their commercial, non-educational counterparts in gaming experience and that can effectively teach curriculum-related subject matter are an extremely fascinating idea. To achieve that, psycho-pedagogical background must be developed and a corresponding technology must be realized, that allows a personalization of gaming and learning. Therefore, existing approaches to adaptive educational systems must be extended in order to maintain an immersive gaming experience, motivation, and probably flow experience by suitable adaptive interventions (Kickmeier-Rust, Hockemeyer, Albert and Augustin 2007).

In supporting such 'intelligent' adaptation and personalization, interventions and feedback has a crucial role. At many staves of the learning ladder, from a psych-pedagogical perspective, support and feedback is necessary in order to ensure successful, effective, and complacent learning. Considering the importance of not destroying immersion, flow, and engagement in the game, the assessment of the learning progress and psycho-pedagogical feedback must occur in a non-invasive way. This, however, requires an intelligent system that is capable of assessing individual competencies and learning progress by observing and interpreting the learner's behaviour in the learning situations within the game.

The work presented here has its origin in the ELKTRA project and its demonstrator game. This game technically implements a theoretical approach to non-invasive personalization by interventions and feedback within a complex and fragile game context including all its constraints. To collect empirical evidence on the effects and efficacy of micro adaptive assessment and interventions, we conducted several evaluation sessions with the demonstrator game, focusing on different aspects of game-based learning, assessment, and particularly interventions and feedback.

In general, the literature indicates pros and cons of (immediate) feedback in learning contexts. On the one hand, interventions may have a disruptive influence on concentration and immersion and may decrease the learner's efforts by providing solutions. On the other hand, appropriate interventions an guide the learner in a meaningful way and support learning by providing appropriate information.

The present evaluation the ELEKTRA demonstrator game, and particularly the slope device LeS, provide some evidence that the degree of appropriateness of an intervention is key to its impact and success. We could show that micro adaptive interventions lead to a faster approach to the correct solution, meaning to a faster problem solving process, in problem solving situation than neutral, inappropriate, or no interventions. In addition, we could demonstrate that providing the learner with appropriate, personalized interventions resulted in a better learning performance with the demonstrator game in comparison to providing no interventions at all.

The progress in the state-of-the-art in DEGs and game-based learning made with the ELEKTRA project is taken up by the 80Days project. 80Days (www.eightydays.eu) is a multi-disciplinary research and development project, running from 2008 to 2010, funded by the European Commission. This projects aims to advance the approaches to micro adaptive assessment and interventions and it aims to introduce an approach to macro adaptivity by interactive and personalized storytelling.

6. Acknowledgements

The research and development introduced in this work is funded by the European Commission under the sixth framework programme in the IST research priority, contract number 027986 (ELEKTRA, www.elektra-project.org) as well as under the seventh framework programme in the ICT research priority, contract number 215918 (80Days, www.eightydays.eu).

References

Albert, D., Hockemeyer, C., Kickmeier-Rust, M. D., Peirce, N., & Conlan, O. (2007). *Microadaptivity within complex learning situations –- a personalized approach based on competence structures and problem spaces*. Poster presented at International Conference on Computers in Education (ICCE 2007).

Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2004). Toward tutoring help seeking: Applying cognitive modelling to meta-cognitive skills. In *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, 227-239.

Anderson, J. R., Conrad, C. G., & Corbett, A. T. (1989). Skill acquisition and the LISP Tutor. *Cognitive Science*, 13, 467-505.

Azevedo, R., & Bernard, R.M. (1995). A meta-analysis of the effects of feedback on computer-based instruction. *Journal of Educational Computing Research*, 13, 111-127.
Chi, M. T.H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25(4), 471-533.
Doignon, J., & Falmagne, J. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, 23, 175–196.
Doignon, J., & Falmagne, J. (1999). *Knowledge Spaces*. Berlin: Springer.
Harmatz, M. G., & Lapuc, P. S. (1968). A technique for employing a yoked control in free operant verbal conditioning experiments. *Behavioral Research and Therapy*, 6, 483.
Kickmeier-Rust, M. D., & Albert, D. (2008). The ELEKTRA ontology model: A learner-centered approach to resource description. *Advances in Web Based Learning – ICWL 2007* (pp. 78-89). Lecture Notes in Computer Science, 4823/2008. Springer Berlin / Heidelberg.
Kickmeier-Rust, M.D., Albert, D., Hockemeyer, C., & Augustin, T. (2007). Not breaking the narrative: Individualized competence assessment in educational games. In Proceedings of the *European Conference on Games-based Learning (ECGBL)*, October 25-26, 2007, Paisley, Scotland.
Kickmeier-Rust, M.D., Peirce, N., Conlan, O., Schwarz, D., Verpoorten, D., Albert, D. (2007). Immersive Digital Games: The Interfaces for Next-Generation E-Learning? In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Applications and Services* (pp. 647-656). Lecture Notes in Computer Science, 4556/2007. Berlin: Springer
Korossy, K. (1999). Modeling Knowledge as Competence and Performance. In D. Albert & J. Lukas (Eds.), *Knowledge Spaces: Theories, Empirical Research Applications* (pp. 103–132). Mahwah, NJ: Lawrence Erlbaum Associates.
Leelawong, K., Davis, J., Vye, N., Biswas, G., Schwartz, D., Belynne, T., Katzlberger, T., & Bransford, J. (2002). The effects of feedback in supporting learning by teaching in a teachable agent environment. In P. Bell, R. Stevens, & T. Satwicz (Eds.), *Keeping Learning Complex: The Proceedings of the Fifth International Conference of the Learning Sciences* (ICLS) (pp. 245-252). Mahwah, NJ: Erlbaum.
Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology*, 17, 26-65.
Schmidt, R. A., Young, D. E., Swinnen, S., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 352-359.
Schooler, L. J., & Anderson, J. R. (1990). The disruptive potential of immediate feedback. In Proceedings *of the Twelfth Annual Conference of the Cognitive Science Society*, 702-708, Cambridge, MA.
Simon, H. A. (1978). Information–processing theory of human problem solving. In W. K. Estes (Eds.), *Handbook of Learning and Cognitive Processes, Vol. 5: Human Information Processing*. Hillsdale, NJ: Lawrence Erlbaum.
Tan, J., Biswas, G., & Schwartz, D. (2006). Feedback for metacognitive support in learning by teaching environments. In *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*, Vancouver, Canada, pp. 828-833.