Designing for self-efficacy in a driving simulator:
A pilot study

Björn Sjödén

VT 2007

Designing for self-efficacy in a driving simulator:
A pilot study

Björn Sjödén

ABSTRACT

Virtual Reality and advanced, interactive simulators illustrate the potentials of modern technology for improving the ecological validity of learning experiments. Yet much more preliminary work is needed to appreciate how these growingly complex means can be utilized to promote high-order emotional and motivational constructs of the user, such as self-efficacy. Drawing from the frameworks of serious games and social cognitive theory, the present study presents a tentative design for promoting the learner's self-efficacy for driving, as represented in a full-scale game-based car simulator. In a pilot experiment, 30 inexperienced drivers assessed their own driving self-efficacy with regard to three specific traffic safety variables before and after they were randomly assigned to either a new, adapted version of the game, aimed at promoting the player's self-efficacy through positive feedback and matching challenge (experimental condition) or a standardized game version developed for previous studies (control condition). The results indicate a higher increase in self-efficacy, a stronger over-all experience and a greater improvement in keeping the speed for participants in the experimental condition than in the control condition. All participants improved their driving skills, but the results were less clear between the two groups with regard to changing lanes and keeping distance to other vehicles. The results encourage further research into how serious game-based simulations should be designed, specifically and generally, to incorporate self-efficacy theory for promoting the learner's sense of being able to manage the learning task(s) at hand.

Introduction

Computerized and game-based simulations have been used for training and education purposes already for decades, from the early applications in primarily scientific and military settings in the 1970's, to the "edutainment" era of the 1980's and 90's and the vast range of "e-learning" products and 3D simulators available today (for an overview, see Susi, Johannesson & Backlund, 2007). The unparalleled development of technology during this time has allowed continuously more advanced and complex artificial environments to emerge as models of real-life settings or systems, characterized in terms such as "Virtual Reality", "Interactive Simulations" and, generally with regard to software for educational purposes, "Virtual Learning Environments" (VLEs) (www.wikipedia.org).

Still, the effectiveness of computers and simulations for learning continues to be heavily debated, suggesting that learning research has not quite kept pace with the technological advancement. Issues of controversy, addressed in several recent reviews and meta-analyses, have concerned not only what is more beneficial from the aspect of cognitive gain – for example whether one learns more from using a fully computerized VLE compared to traditional teaching methods with a human tutor (Mitchell & Savill-Smith, 2004; Vogel et al, 2006) – but also the roles of specific emotional, motivational and social factors for the

human-computer interaction, such as the feeling and need of control (Kay, 2001), fun and engagement (Quinn, 2005; Zyda, 2005). In addition, in an extensive review of the literature on emotion and motivation in learning, Olsson (2003) has pointed to the importance of integrating basic, theoretical research with practical learning applications. In short, researchers seem to agree that when studying VLEs in their actual, social context, much more preliminary work is needed to find out what makes a specific application effective and why.

Although researchers, such as those mentioned above, seem to call in unison for an integrative framework, in which to tie basic, theoretical domains (such as cognitive learning science and neuroscience) and applied domains (such as usability and interaction design) together, there are relatively few studies aimed at using the VLEs to gain knowledge of the learning process itself through active and experimental manipulation of the user's cognitive states. Many designs are restricted to assessing specific features of the software and/or administering pretest/posttest knowledge tests to evaluate the effectiveness of the application used. Meta-reviews and analyses of existing applications may be theoretically valuable, but are not sufficient to make progress, whereas technical guides may be useful for producing new designs, but are as such not sufficient to increase our knowledge of how they work and thus motivate their use in education. In the words of Quinn (2005, p. 1), himself a learning systems designer, "There are too many design opportunities and too little understanding.". [1]

This ubiquitous problem of "man versus machine" becomes especially conspicuous when approaching the complexity of motivation and learning not from the micro-level of connections between synapses in the brain or simplistic reward-and-punishment paradigms, but from the level of the typically human, "global concepts" (cf. Olsson, 2003, p. 28), such as those relating to higher cognitive strategies, self-control, autonomy, intrinsic and extrinsic motivation, interest and empathy – all undeniable influences on learning effectiveness, yet unsatisfactorily understood and explored in technological contexts. The ever-increasing use of computers notwithstanding, human beings remain social beings, a fact which inevitably affects their interaction also with artefacts (a point demonstrated more at length in the work by Reeves & Nass, 1996). The question arises as to how the features of future VLEs should be designed to acknowledge more of the learner's *human* characteristics, as reflected not only in the facts learned, but in how the learning process is approached and undertaken.

The present study takes its theoretical starting point in the concept of self-efficacy, which has been introduced and established through the immensely influential work by social psychologist Albert Bandura (e.g. 1977, 1997). In popular terms, self-efficacy has been described as "the power of believing you can" (Maddux, 2005), though it is formally defined as "… people's beliefs in their capabilities to produce desired effects by their own actions" (Bandura, 1997, p. vii). Importantly, self-efficacy does not refer to specific skills or self value (although there are strong associations), but to the notion of how one perceives and manages one's personal resources to attain desired goals. This emphasis on personal agency and first-hand experience would seemingly stand in conflict with 'lending' oneself to the artificial aid of computers, especially if the artefacts are to dictate the conditions for complex tasks demanding personal effort, such as learning new skills. Perhaps it is unarticulated views like these that lie behind the observations that, in spite of their central position in psychology, self-efficacy theories "… have been largely unnoticed by the Human-Computer Interaction (HCI) community" (Middup & Johnson, 2006, p. 1).

---

[1] It should be noted that also Quinn's work is essentially a pragmatic design guide, and does not dive into the details of specific theoretical constructs, as will be dealt with here.

I will argue, however, that there are many reasons for aiming to integrate self-efficacy theory with interaction design and, particularly, learning simulation games. Firstly, the virtual environment offers unique opportunities for altering and adjusting factors which have a direct effect on the perception of task difficulty, motivation, and locus of control, all of which closely relate to self-efficacy. Secondly, according to Olsson (2003), a notorious weakness of learning research to-date has been the "tremendous lack of ecological validity in the experimental conditions used" (p. 28). It does not require a lot of imagination to realize that precisely issues of ecological validity may be rectified by conducting more learning research in realistic simulations and virtual environments. This echoes the notions of situated and distributed cognition (e.g. Gee, 2004), which view learning in a meaningful and supportive context. Similarly, Van Merriënboer (1997) argues that complex skills are better accomplished in a simulation, where the learner practices the actions needed in real life. In the words of self-efficacy theory, genuine learning experiences correspond to "enactive mastery" (i.e. successful performance) of challenging tasks, which Bandura (1997) considers the main influential source of self-efficacy.

The main point pursued here is that games, in general, become effective for learning when the gap between game-specific skills and real-life applications is bridged by the ability to generalize – comprehend and re-apply – one's knowledge fitting the situation. Because self-efficacy is task-specific, but not skill-specific (that is, one task can be performed by coordinating and orchestrating different skills in a number of ways), its role in learning is not restricted to the ability of operating trivial and specific instruments in the VLE, but rather to how one approaches and adopts to the means available, in challenging and changing situations. For instance, the task of driving a car is not exactly the same in a driving simulator as in real life, since a simulation is only one representation of what it is like to drive a car, under more or less constrained conditions. But if self-efficacy for driving a car in real life is promoted by driving in a simulator, by making the driver more attentive, judicious, etc, as reflected in an actual improvement of performance, then there is learning above the limitations of the simulator, which calls for a further investigation of how specific factors influencing self-efficacy can be applied and guide the design of the simulation environment.

Unfortunately, high ecological validity is not sufficient to make a learning system effective: it still needs to attract users. One salient strategy for meeting this need is reflected in the increasingly popular concept of "serious games", or "(digital) games used for purposes other than mere entertainment" (Susi, Johannesson & Backlund, 2007, p. 1). There are numerous similar and related terms and concepts to serious games, and it is not my purpose to give a thorough account here (see ibid for an overview or www.seriousgames.org), but the basic approach is simple: learning appeals more to the learner's interest and motivation in the form of a challenging and entertaining activity, that is, a game. As Zyda (2005) puts it, serious games "use pedagogy to infuse instruction into the game play experience" (p. 26).

In the present study, efforts were directed to a serious game-based car simulator, partly due to the large-scale on-going research project on computer gaming and traffic safety at the University of Skövde (see www.his.se/seriousgames). Another reason is that self-efficacy theory has recently been brought to the fore by traffic researchers, in the context of formulating new guidelines for Swedish driving education, which emphasize the need of taking self-assessments of specific driver competencies into account when learning how to drive a car and behave in traffic (Sundström, 2006). In addition, the present study addresses some of the issues highlighted by previous experiments with the driving simulator (Backlund et al, 2007), for example narrowing down the previously large number of multiple analyses to

test more specific hypotheses, and focusing more on the gaming element, challenge and feedback, by virtue of the solid theoretical framework chiselled out from self-efficacy research.

The driving simulator and previous studies

The technical setup of the simulator environment has been described in detail by Lebram, Engström and Gustavsson (2006) and with specific regard to traffic safety research by Backlund, Engström and Johannesson (2006) and Backlund et al (2007). Here, the setup, as employed in the present study, is described in summary on the basis of these previous accounts.

The simulator consisted of a real car (a Volvo), which has been set up as a game control, surrounded by seven screens. Some views of the simulator environment are given in figure 1. In addition to the large external screens, there was a small monitor inside the car, which in the present study displayed additional information, such as performance feedback, between driving sessions (but never while driving). The car was equipped with internal speakers which allow for information to be read aloud directly to the driver, occasionally including instructions while driving (for the tasks requiring the driver to change lanes on command). There was also a small camera inside the car, pointed towards the driver's face while driving, permitting subsequent analysis of the driver's use of the rear-view mirrors.

The gear change was fully automatic, which is why participants needed only familiarize themselves with the most basic controls for driving (brake, accelerator, steering wheel and turn-signals). While driving, an illusion of physical movement was created through the use of motor sounds, vibrations and the car's fan. The vibrations grew stronger with increased speed, providing some sensory feedback.

The game itself was introduced by a short narrative, leading to the task of having to follow an ambulance on a four-laned road, without letting it pass from sight. In the standard game version, the player received some automated feedback and grades concerning his/her performance concerning violations of common traffic rules (e.g. speed violations). In case of failure, for example crashing or losing the ambulance from sight, the level was restarted. The game version for this study comprised 10 regular game levels and one "instructional level" (changing lanes on command) used for reference testing the driver's lane-changing skills, including the use of turn-signals, before and after practice in the simulator. The difficulty of each level was manipulated by a number of independent variables, such as the sight and weather conditions (e.g. light, fog), the volume of traffic and the traffic composition, including the behaviour of individual computer-steered cars and lorries. For example, whereas level 1 of the game was characterized by little traffic and good sight conditions, the player of level 9 needed to deal with fog, jammed traffic lanes and an increasing number of large lorries blocking the road view, which made overtaking more difficult. On the high number levels, the driver also had to pay more attention to vehicles coming from behind.
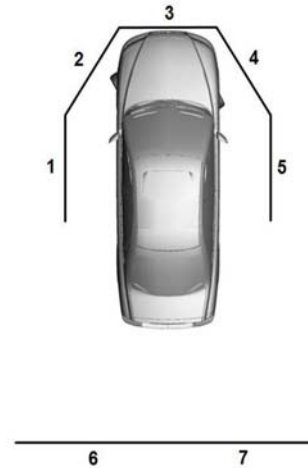
Fig. 1. Views of the simulator gaming environment: the driving simulator uses seven screens to project a nearly 360 degree-field, allowing the use of rear-view mirrors and side-views of surrounding traffic (from Backlund et al, 2007).

Previous research with the simulator has successfully demonstrated its function for studying and training a broad range of driving behaviours relating to traffic safety. The present study focused on the three distinctive driving behaviours for which training in the simulator has shown particularly strong learning effects (Backlund et al, 2007). Those are (1) keeping within speed limits, (2) keeping an appropriate distance to surrounding traffic and (3) correctly performing lane changes, preceded by giving a turn-signal. The operational definitions of these parameters were worked out in collaboration with traffic safety expertise already for the first driving studies by Backlund, Engström and Johannesson (2006). In sum, traffic violations were assessed according to the following three measures:

- *Speed violation* was defined by the average number of seconds per kilometre a participant spent driving 10 or more km/h above the speed limit.

- *Too short headway distance* was defined by the average number of seconds per kilometre a participant spent within a distance of 0–3 seconds behind a vehicle.

- *Improper lane change* was defined by the proportion of lane changes not preceded by the driver using the turn-signal.

Conversely, safe driving was measured as the extent to which the driver kept below 10 km/h too fast at all times, remained in a position with more than three seconds' headway distance to other vehicles, and correctly used the turn-signal before changing lanes.

Aims and scope of the present project

The main aim of the project is pragmatic but theory based, namely to tentatively seek out and detail a procedure for producing a new, self-efficacy promoting, serious game design which would improve results through endorsing the learner's perceived self-efficacy for the learning task (i.e., to exercise certain, safe driving behaviours). This necessarily implied a broad scope of study, however the purpose was as well to provide an example of a design and lines of reasoning which can be applied and increase the learning effectiveness also for other game-based and/or simulation based learning systems.

Specifically, the project was carried out in two phases: first the implementation phase of defining the elements to construct a design for promoting self-efficacy and implement them; second, the performance of an experiment to test the hypotheses which would follow from self-efficacy theory. For the second phase, a traditional pretest/posttest experimental design, with one treatment group and one control group, served for evaluating changes in self-efficacy and learned driving proficiency after having practiced for a certain time in the driving simulator. In contrast to dominant lines of research, the simulator training was not performance-focused (although some performance feedback was provided) but was primarily aimed at increasing the driver's self-efficacy through experiential means (appropriate challenge and feedback), hypothetically leading the driver to greater confidence and a more favourable interpretation of his/her own abilities, which in turn would affect performance positively. A successful design would thus mainly be reflected by an increase of the participant's self-assessed self-efficacy after driving and secondarily by the driving performance. The following main hypotheses were posed:

- The level of driving self-efficacy will increase more for participants in the experimental condition than in the control condition.

- With an increase in self-efficacy, there will be an associated increase in performance after training in the simulator.

- Participants in the experiment condition will become more motivated and engaged in using the simulator, which will reflect in their stated level of enjoyment (fun) when assessing their experience.

In addition, linking to the previous studies on computer gaming and traffic safety at the university, it was considered of interest to explore whether "gamers" (people with much experience of computer games) differ from "nongamers" (people with little gaming experience) with regard to self-efficacy. For instance, would the conventional game (control) condition appeal more to the gamers, which typically are the target group, than the adapted version theoretically designed for promoting self-efficacy? Do gamers experience stronger self-efficacy because they are playing a game?

Method

I. Designing for self-efficacy

Producing a game design to promote self-efficacy essentially concerned two aspects of the existing game: (1) Adapting the *challenge*, as represented by a re-classification of the game's difficulty levels, and (2) Adding appropriate *semantic feedback,* comprising both instructional and performance feedback screens. These two points correspond to two of the principal sources of self-efficacy (Bandura, 1997), namely, (1) enactive mastery experiences and (2) verbal, or semantic, persuasion.

Combining design elements to produce both kinds of experiences in a (serious) gaming context bears the additional advantage of scaffolding different goal orientations, as identified by Dweck (1999): (a) the learning goal orientation (the "serious" aspect of the serious game) which is to develop competence through mastering challenging situations, and (b) the performance goal orientation (the "game" aspect of the serious game), which is to demonstrate and validate one's competence by seeking positive judgments and avoiding negative judgments. The presented design for self-efficacy directs even "performance learners" to learning goals, by making the mastery of challenging tasks a part of the process to obtain favourable judgments.

Challenge

As previously mentioned, self-efficacy theory (Bandura, 1997) suggests that the challenge of a task should be appropriately matched to the person's actual ability. As skill develops, the difficulty needs to increase progressively, to maintain a balanced challenge. Ideally, an individual's level of self-efficacy should even be somewhat higher than the performance would warrant, since this will encourage the individual to keep trying and to take on increasingly more difficult tasks (Stratton & Hayes, 1999).

In practice, however, and especially when it comes to complex tasks such as driving a car, the theoretically optimal design can hardly be achieved, at least not with the technology available in this study, since one cannot know what each individual's "actual ability" is – or which variables in the game would need to be adjusted to affect it accordingly. This problem became tangible when I first presented the idea of a self-efficacy design to the original game designers, since they noted that there was "no obvious answer to what makes one game level more or less difficult than another" (H. Engström, personal communication, February, 2007).

Another way of posing the question, focusing on the artefact, would be "what constitutes a balanced challenge?". In lieu of a systematic analysis, the answer lends itself to heuristic and empirical approaches. Fortunately, in this case, there was a vast amount data from previous test sessions (Backlund et al, 2007), including 1240 driving sessions by 71 drivers on 10 levels. It was agreed that a rough but indicative measure of a single game level's difficulty would be the number of trials, that is, the number of times the participants had needed to repeat the level before completing it successfully. However, in agreement with the game designers, it was decided to disregard the data for the beginning and end levels (1 and 10). The number of trials for level 1 is likely inflated, because it was the drivers' first encounter with the simulator, and they had not had any practice trial to acquaint themselves with the

controls. For level 10, only 15 of the 71 drivers reached this level within the time frame of the driving session. Because of their exceptional skills, these drivers were not deemed representative for the rest of the sample.

Hence, having obtained the raw data from previous experiments, I recompiled the trial statistics according to the number of trials per level (1-10), as depicted in figure 2.[2]
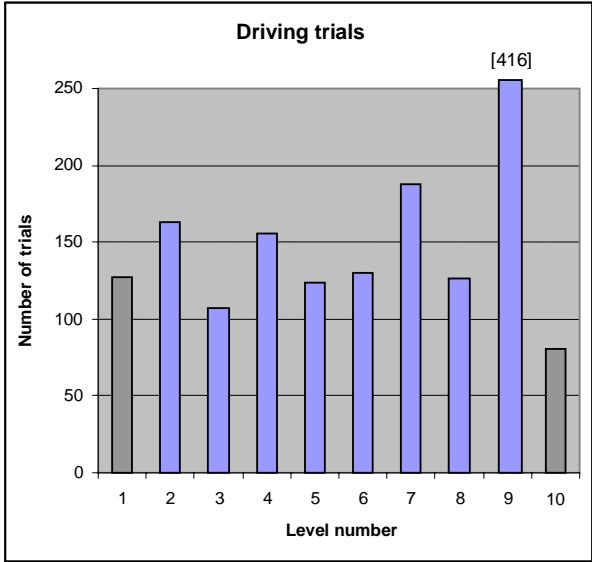

Fig. 2. Number of trials per game level, original distribution (disregarding levels 1 and 10, see the text).
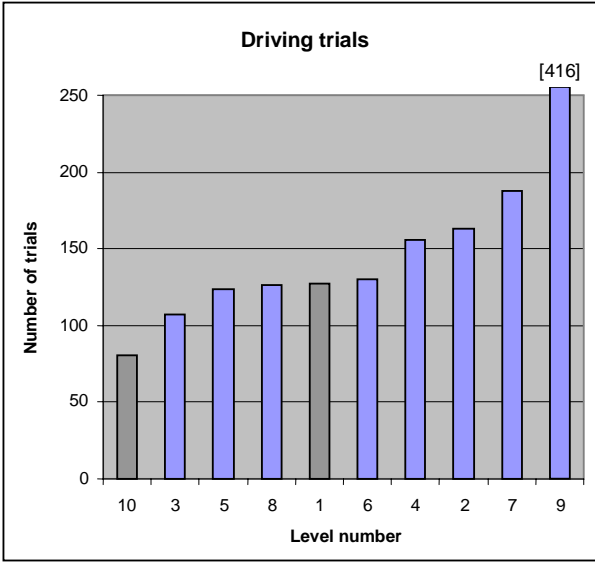
Fig. 3. Game levels re-classified and re-arranged according to "increasing difficulty" (disregarding levels 1 and 10, see the text).

Assuming that the other columns are valid indicators of level difficulty, it stands clear from the diagram in figure 2 that there is no linear progression of difficulty in the game that would correspond to a linear development of skill with practice over time. As confirmed by the original developers (H. Engström, personal communication, February 2007), this pattern is common in game design, which typically varies progressively more difficult levels with occasional "top challenges", or so-called "boss levels", before letting the player proceed to the next level. A closer look at the data reveals that level 9 appears as the undoubtedly most difficult level (having required no less than 416 trials), and level 3 the easiest, whereas the other levels may be categorized in two groups: The levels {2, 4, 7} form the more difficult group, and the levels {5, 6, 8} form the less difficult group.

This said, another reservation to keep in mind when interpreting the data is, that the levels with a lower number are probably more difficult than just the number of trials suggests, due to learning effects when having played a number of levels. Figure 3 shows the rearrangement of levels according to the number of trials only (again, levels 1 and 10 disregarded). This redirects attention to the qualitative features of the individual levels. For this reason, a number of test sessions were conducted with volunteer drivers (employees and friends of the

---

[2] There are some important reservations to these data: first, the participants in the referred study were more proficient than the ones recruited in the present study; second, due to imposed time limits, not all drivers played all levels of the game (and only the most proficient reached the 10th and final level). For this reason, whereas levels 1 and 2 were played by all 71 drivers, the number of trials (t) for levels 3 to 10 have been adjusted upwards by a "missing drivers index", i.e., the adjusted (t) = t * 71/n, where n is the number of drivers having played that level.

experimenter), and the drivers were interviewed about how they experienced the progression of difficulty between levels. Taking both the qualitative and the quantitative data into account allowed for a better approximation of how the levels should be rearranged to more aptly represent a progressive challenge in the game.

In relation to self-efficacy, one must also take into account that, even among inexperienced drivers, some will perceive the driving tasks as easier than others and would therefore require a higher challenge to match their skills. In the end, three blocks of difficulty were constructed, each comprising four levels (to make up an appropriate driving session length of about ten minutes). Each participant was to drive two sessions with a short break in between, repeating the last level after having completed a block. This made use of all ten levels and allowed participants (unknowingly) to start on either an "easy" or "moderate" block, depending on their pre-assessed self-efficacy. The final reclassification of levels to represent a progressive difficulty increase is summarized in order below, having compensated for learning effects by placing higher-number levels last within each respective block. (Note: levels in the adapted design were renamed with letters to avoid confusion with the previous level numbers.)

Block I ("easy"): levels 1 → 3 → 5 → 6 (= a, b, c, d)

Block II ("moderate"): levels 6 → 2 → 4 → 8 (= d, e, f, g)

Block III ("difficult"): levels 8 → 7 → 9 → 10 (= g, h, i, j)

Feedback screens

The importance of encouraging, learning-oriented verbal feedback to promote self-efficacy has been emphasized not only by Bandura (1997), but in a number of studies by other researchers in applied domains, for instance in the research by Dweck (1999), Zimmermann and Schunk (2001) and Maddux (2005). Although most feedback discussed here is verbal, the experiment also included graphical symbols (e.g. a happy face) and some sound signals, which are encompassed under the broader term "semantic feedback".

From a self-efficacy perspective, people need to know that they are developing skill and also attribute their progress to their own skill rather than contextual factors (Dweck, 1999; Zimmerman & Schunk, 2001). Feedback plays a key role in this process, since people may not assess their performance if they receive no information on how they perform (ibid; Bandura, 1997). There was reason to believe that the existing game design could be improved in these respects, since the game contained only very brief feedback after completing a level, and no verbal comments. Specifically, the feedback given was limited to the number of traffic violations made and a conventional grading on a four-point scale, corresponding to the grades given in Swedish schools (IG, G, VG or MVG).[3]

To better meet the suggestions from self-efficacy research, feedback screens tailored for promoting the driver's self-efficacy beliefs were constructed for the experimental game

---

[3] The school grades were obviously intended as a straight and informative feedback on performance that students would be familiar with (one knows that a "Pass" is needed to continue). However, having kept the original between-level feedback screens in the control condition, it became apparent to the author, while conducting the experiments, that several participants reacted negatively to this kind of feedback. Arguably, scholastic grades weakened the association to leisure-time, gaming fun.

condition. Considering that the difference between the experiment and control conditions should pertain to the content of the feedback, and not the amount of feedback given, a corresponding number of feedback screens with neutral information were added to the control group. The screens appeared just before the main driving session, after half the session and at completion of the session. However, Table 1 contains a summary of the standard feedback provided to the respective participant groups, and detailed screenshots are provided in the Appendix.

Table 1. Example extracts from the standardized feedback screens displayed in the experiment and control conditions respectively.

	Welcome screen	**Pause screen**	**Finish screen**
Experiment condition (positive feedback)	E.g. "You affect the difficulty yourself, the better you become … You are only competing against yourself:"	"Congratulations! … You have already improved your driving skills."	"Well done! You have successfully completed the second session."
Control condition (neutral/instructive feedback)	E.g. "The difficulty increases as you proceed in the game … You will get minus points when you make errors."	"Pause. You have finished the first session."	"Game over. You have finished the second session."

Note, that besides the standard feedback screens (same to all participants in the condition), the between-level feedback on individual performance was kept as is in the control condition (grades and number of traffic violations), whereas in the experiment condition it was replaced by a positive number expressed as a percentage, e.g., "You successfully completed 86% of the driving tasks" instead of "Number of violations: 14". In the experiment condition, a final grade was displayed on the finish screen, but it always the highest or next-highest possible (MVG or VG) depending on performance.

In addition, the text on all standard feedback screens were recorded by a professional reader and played aloud through the internal speakers of the car, simultaneously with the graphical presentation. The display of the pause- and finish screens was accompanied by a short fanfare in the experimental condition, and with a "pling" in the control condition.

II. The experiment

Participants

For the first phase of the study, the pre-set goal was to recruit a minimum of 30 participants, with a minimum age of 18 years, who had not yet obtained a driving license and considered themselves to be inexperienced drivers. Participants were recruited through posters on announcement boards at the University of Skövde, in local traffic schools and via contacts in two near-by schools (primarily second- and third-year students on the upper secondary school IT programme).

Due to recruitment difficulties during the time frame of the study and the fact that submission was completely on a volunteer basis, the sample unfortunately became heavily biased. Of the 30 participants who volunteered, the majority was enrolled in education programs relating to computer games development, where the general gaming experience was very high and most pupils were male. As a result, 6 women and 24 men, ages 18-37 years, participated. 15 people were randomized to the control condition and 15 to the experimental condition, regardless of their previous gaming experience. All participants had at least some previous experience of computer games; only four participants reported "some" or "average" gaming experience, whereas the other 26 reported "much" or "very much" gaming experience. None of the participants had a driving license, but several were enrolled in traffic schools.

Questionnaires

1. Self-Efficacy Questionnaire, SEQ (digital). As recommended by Bandura (1997), the participant's self-efficacy was measured by responses on a questionnaire pertaining to the specific tasks at hand. In this case, the main question read (translated from Swedish) "How certain are you that you can perform the following operations in an adequate manner during the driving session?". The three sub-items to be assessed were "Keeping appropriate speed", "Keeping appropriate distance to surrounding vehicles" and "Changing lanes and using the turn-signals". The items were graded on a 10-point scale, from 0, "Very uncertain", to 10, "Very certain". Researchers have recommended the use of ten-point scales for measuring self-efficacy due to their relatively high degree of reliability and discrimination (Bandura, 1997; Sundström, 2006).

2. Post-Driving Experience Evaluation Form (paper). As in the study by Backlund, Engström and Johannesson (2006), a short questionnaire was delivered immediately after the driving session with five questions relating to the participants' experience, regarding how fun, realistic and controllable they found the driving, as well as whether they had experienced any nausea or dizziness. The items were graded on a five-point scale, from "Not at all" to "Very true".

3. Questionnaire on video gaming and previous gaming experience (paper). Attached to the second form were another two pages of questions relating to the participant's computer gaming habits. The main purpose of this questionnaire was to add to previous research on this topic at the University of Skövde and the full results will be published elsewhere. One item of specific interest for the present study was the question asking participants to rate their previous gaming experience on a five-point scale, from "None" to "Very much". Participants

with a score of 1 or 2 were categorized as "nongamers", and those with a score of 4 or 5 were categorized as "gamers" (cf. Backlund et al, 2007).

Procedure

All participants were tested in individual sessions, without any passengers in the car while driving. First, the participant was greeted and introduced to the simulator environment, as depicted in figure 1. The experimenter (in all cases the author) then followed a script to ensure that all participants received the same information and in the same order, as detailed here.

Before the formal experimental session began, the car seat and rear-view mirrors were adjusted to the individual participant. Participants were informed about the general procedure of the experiment – that they were going to drive a number of "levels" under various driving conditions – but that the precise instructions and feedback would be self-explanatory once started. The in-game instructions consisted of information on the small touch-screen inside the car between levels and read out loud from the loudspeakers. The participant progressed from one level or feedback screen to the next by pressing on "Continue" or "Ok" on the touch-screen. No new information was displayed on the screen while driving. After a short introductory level, for the participant to get a first feeling of what driving in the simulator would be like, he/she provided his/her first self-assessment on the self-efficacy questionnaire described above (a second measure was taken after the main driving session, as depicted in figure 2).

Importantly, all participants were asked to drive "as if for real", that is, not deliberately violating traffic rules or causing crashes just because they were playing a game. From an ethical point of view, participants were also informed that the entire session was being recorded, however all data would be anonymized in the final report, and that they were allowed to withdraw from the experiment at any time, in case of dizziness or nausea from driving, or for any other reason.

Following this introduction, the experiment automatically progressed through the following three phases:

Reference tests 1. The reference tests comprised two short driving sessions on different courses, which took 2-3 minutes. On the first course, the task was to drive a certain distance, with very light traffic conditions, while changing lanes (left or right) seven times on command. This type of test allowed for a very accurate scoring of proper lane change behaviour, since the traffic situation was not confounded by other vehicles blocking the lanes or in other ways restraining the task. Drawing on previous experience and in contrast to the studies by Backlund et al (2007), it was decided not to use this test for assessing speed violations or short headway distance, since the situation was also quite artificial and did not permit much interaction with surrounding traffic. To compensate for these limitations, an additional reference test with a game task was included (i.e. "follow the ambulance"), from which speed and distance to other vehicles could be more accurately evaluated.

Main driving session. This was the participants' opportunity to practice their driving skills, by playing the simulator game described earlier. Participants in the experiment condition, who had indicated an average self-efficacy score less than 5,5, were automatically assigned by the computer to start on the "easy" level block (I), whereas those with a self-efficacy score

greater than 5,5 were assigned to start on the "moderate" level block (II). Participants in the control condition all played the existing game, with the level structure used in previous studies. The mid-session break with feedback was displayed when the participant had either completed the four levels in the block, or having played for ten minutes (the upper time limit, unknown to the participants). The participants decided themselves when to continue, by pressing on the touch-screen (typically after a few moments only). They then resumed the last level they had been driving before the break and continued for another four levels or ten minutes. The driving session ended by displaying one of the finish screens described earlier.

Reference tests 2. The reference tests were repeated to provide a second measure of the relevant driving skills, comparable to the test results preceding the practice session. Data for all sessions were automatically recorded on the computer and exported for analysis in MS Excel and SPSS for analysis.

The experimenter did not interfere with the participant at any time during the driving session, except when having to restart the simulator session and telling the participant so, on the few occasions of system crashes. Otherwise, he remained overlooking the system functionality on the computer screens and only returned to the participant when the entire session was over, to administer the final questionnaires. Afterwards, the participants were debriefed and asked not to share any specific information about the design of the study with other, presumptive participants. The complete experimental procedure, including the administration of questionnaires, took 45-50 minutes per participant. Figure 2 shows a complete overview of the experimental design and procedure represented in a flowchart.
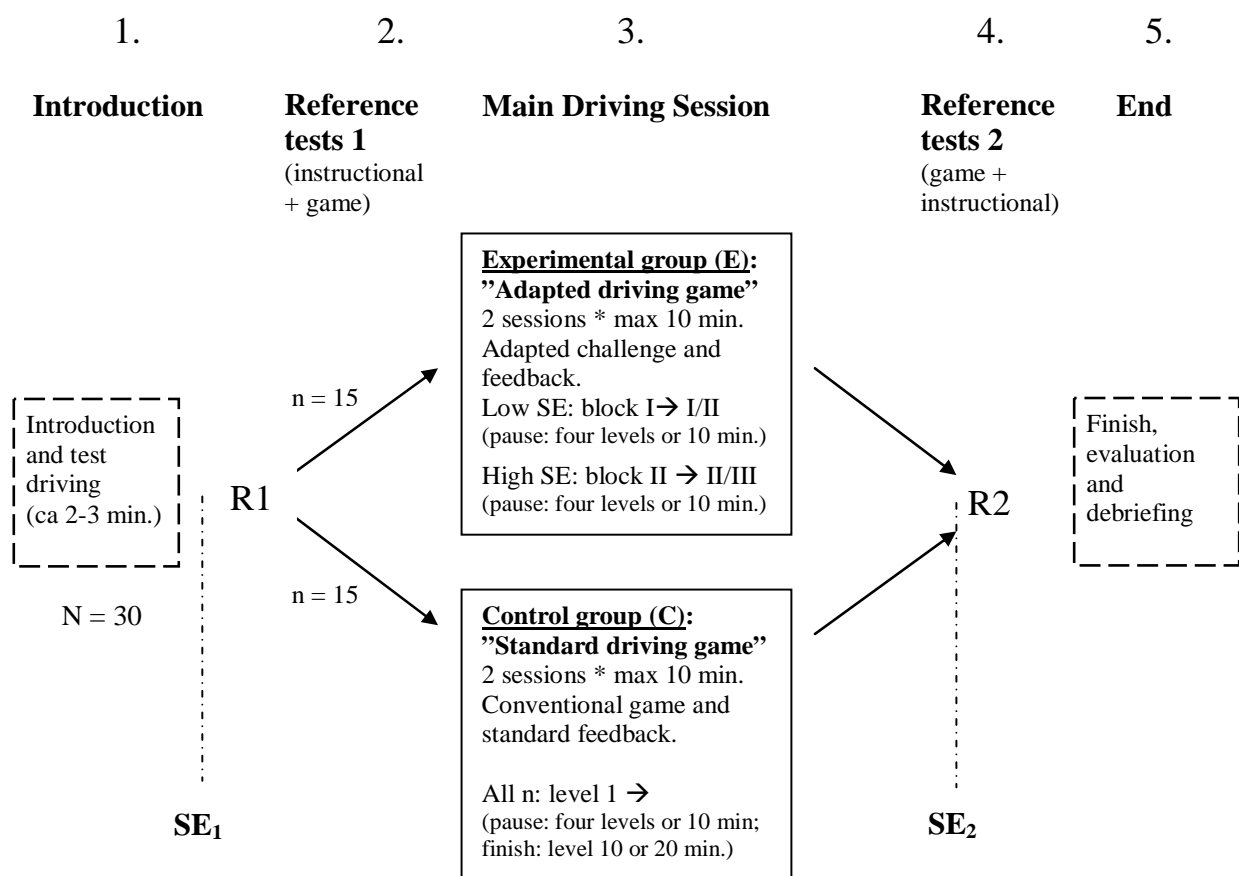

Fig. 2. A schematic overview of the five phases of the experiment.

Results

It is important to note that the results presented here make part of a pilot study, the first of its kind, and are as such not conclusive. The biased sample set some limitations to the analyses, mainly regarding the characteristics of nongamers (only four individuals) as compared to gamers. Also, the procedure of the experiment, especially in the beginning, became disturbed by a number of intermittent computer crashes. The stability of the system has however improved during the course of the experiment, which is encouraging for the data collection to continue in a second phase, in which recruitment efforts will be primarily directed towards female participants, and the goal will be to run a minimum of 40 sessions.

For the statistical tests which were still meaningful, a conventional significance level of $p<0,05$ was chosen (two-tailed). "SD" in the tables is short for "Standard Deviation".

Experience of the simulation game

The responses on the five evaluation items after driving in the simulator were compiled in table 2. A score of five represents the highest degree of agreement, thus indicating that the experience was overall considered fun – somewhat more so in the experimental condition – and did not cause much negative side-effects such as nausea or dizziness. The simulation environment was deemed realistic, but could be more controllable. None of these differences were significant on the $p<0,05$ level.

Table 2. Results on the Post-Driving Experience Evaluation Form (scores 1-5).

Experience: **Condition**:		fun	Realism	nausea	dizziness/ giddiness	sense of control
Control	Mean	**4,40**	**3,27**	**1,20**	**1,20**	**3,07**
(n=15)	SD	0,74	1,10	0,41	0,41	0,96
Experiment	Mean	**4,53**	**3,53**	**1,33**	**1,40**	**3,47**
(n=15)	SD	0,92	0,83	0,62	0,63	0,83
All	Mean	**4,47**	**3,40**	**1,27**	**1,30**	**3,27**
(N=30)	SD	0,82	0,97	0,52	0,53	0,91

Self-efficacy measures

An overview of the mean task-specific self-efficacy scores for participants in the control and experiment conditions respectively is provided in table 3. The self-assessments were generally more positive the second time for the experimental group, whereas the opposite was true for the control group.

Table 3. Self-efficacy self-assessments for specific driving tasks, before and after the main driving session in the simulator (SE_1 and SE_2 respectively).

Task variable:		LANE CHANGE self-efficacy		KEEPING SPEED self-efficacy		KEEPING DISTANCE self-efficacy	
Condition:		SE_1 (lane)	SE_2 (lane)	SE_1 (speed)	SE_2 (speed)	SE_1 (distance)	SE_2 (distance)
Control	Mean	**5,20**	**5,13**	**6,13**	**5,13**	**4,60**	**4,47**
(n=15)	SD	2,62	2,26	2,70	1,55	1,99	1,77
Experiment	Mean	**5,00**	**5,80**	**5,80**	**5,93**	**5,00**	**5,07**
(n=15)	SD	2,24	2,08	2,08	1,91	1,36	1,44
All	Mean	**5,10**	**5,70**	**5,97**	**5,53**	**4,80**	**4,77**
(N=30)	SD	2,40	2,09	2,37	1,76	1,69	1,61

To obtain a somewhat more inclusive measure of the participants' perceived self-efficacy for driving, an average self-efficacy score was calculated as the mean score of the three specific self-efficacy variables, resulting in a mean SE_1 and a mean SE_2. The difference denotes the development (in all cases an increase) of the participant's three-component self-efficacy for safe driving, which is useful for subsequent analysis. Specifically, a question of relevance was whether the experiment condition promoted the driver's average self-efficacy to a higher extent than the control condition. These mean scores are summarized in table 4.

To form an idea of the strength of the apparent effect, the effect size, *Cohen's d*, was calculated according to Cohen (1988). Notably, the effect of the control condition in the analyzed sample turned out *negative* for the development of self-efficacy. These results are in accordance with the specific mean scores observed in table 3 above. The negative effect of the control condition is a small one by Cohen's standards, whereas the experimental condition had a markedly larger *positive* effect, between small and moderate. [4]

Table 4. Mean driving self-efficacy scores, their mean increase (SE_2–SE_1) and the effect size.

Score: Condition:		Mean SE_1	Mean SE_2	SE difference (SE_2–SE_1)	*Cohen's d* (effect size)
Control	Mean	**5,31**	**4,91**	**-0,40**	**(-)0,24**
(n=15)	SD	2,29	1,63	1,67	
Experiment	Mean	**5,27**	**5,75**	**+0,49**	**0,39**
(n=15)	SD	1,69	1,29	1,26	
All	Mean	**5,29**	**5,33**	**+0,04**	**0,03**
(N=30)	SD	1,98	1,51	1,52	

---

[4] Cohen (1988) proposes three categories for judging whether an effect is large or small, where $d = 0,20$ is a small effect, $d = 0,50$ is a moderate effect, and $d = 0,80$ is a large effect.

An independent-samples t-test was carried out to test whether the difference in self-efficacy development between the two groups was significant. The result did not quite reach statistical significance ($p = 0,11$).

Potential interaction effects of gaming experience and experimental treatment on the self-efficacy development were explored in a 2*2 (Gamers vs. Non-gamers * Control vs. Experiment condition) ANOVA with the difference in self-efficacy ($SE_2–SE_1$) as the dependent variable. The results are presented in the cross-table as depicted in Table 5. However there were clear observed differences between the groups, especially between the relatively large number of gamers in the control and experimental conditions, none of the observed differences, nor interaction effects, were statistically significant ($p>0,05$).

Table 5. Interaction effects on increase in self-efficacy (Mean SE_2-SE_1, in unit points) between gamers and nongamers respectively, and the condition exposed to in the experiment.

Type: Condition:		Nongamers' change in self-efficacy	Gamers' change in self-efficacy	All (mean)
Control	Mean	**0,00**	**-0,43**	-0,40
	N	1	14	15
Experimental	Mean	**+0,11**	**+0,58**	+0.49
	N	3	12	15
All (mean)	Mean	**+0,08**	**+0,04**	+0,04
	N	15	15	30

Traffic safety variables and driving behaviour

From the vast amount of data recorded by the driving simulator, the analysis was restricted to the three selected traffic safety variables previously defined. In all cases when measuring the change or improvement in driving skill, the two reference scores were compared, from before and after the main driving session (R2-R1; cf. figure 2). In table 6 below, all scores point to an *improvement* from a traffic safety point of view, that is, the proportion of correctly performed lane changes increased after training, the drivers spent less time at too high speeds and less time positioned too close to other vehicles. (For example, the average score of 7,15 for keeping the speed in the experimental condition means that the drivers improved their speed-keeping scores by spending 7,15 seconds less of their driving time above speed limits in Reference test 2 than in Reference test 1.)

Table 6. Results on traffic safety variables (mean differences R2-R1) for the control and experimental conditions.

Variable: Condition:		Difference R2-R1 LANE CHANGES (proportion correct)	Difference R2-R1 KEEPING SPEED (duration less too fast, s.)	Difference R2-R1 HEADWAY DISTANCE (duration less too close, s.)
Control	Mean	**+0,24**	**5,47**	**9,49**
(n=15)	SD	0,36	7,13	24,71
Experiment	Mean	**+0,18**	**7,15**	**9,48**
(n=15)	SD	0,28	7,20	17,37
All	Mean	**+0,21**	**6,31**	**9,48**
(N=30)	SD	0,32	7,09	20,99

As seen, the results were mixed; whereas there was a notable difference in keeping the speed between the two groups, to the advantage of the experimental condition, the lane changing scores appeared slightly more improved for the control group, and no apparent difference was found with regard to the short headway distance. Independent-samples t-tests were performed to compare the mean differences (R2-R1) on all three traffic safety variables; no significant results were obtained ($p>0,05$).

Discussion

This study stemmed from the interdisciplinary ground of cognitive science to provide one example of how the learner's self-related beliefs – specifically, self-efficacy – quite literally come into play within an advanced learning, game-based simulation environment. It seems a logical consequence that, with greater realism and technological complexity, people will seek to construct – and learn from – their virtual experiences in a similar vein to their everyday experiences.

Nevertheless, a simulation – by definition – remains an imperfect model of reality, in the sense that it will never achieve exactly the same qualities as real-life experience (or else, it would be impossible to distinguish from reality). This fact should, of course, be turned to the advantage for the learner: Since the simulation is *different* from reality, it may also be *better* than reality with regard to specific aspects for facilitating learning. After all, the real word does not only offer positive opportunities for learning, but also obstacles to learn effectively. The objective of the virtual environment, then, needs to be on complementing and scaffolding, rather than substituting or governing, the learner's abilities. It is in this context self-efficacy becomes imperative; the power of control needs to remain with the learner, not with the external system.

Relevant to self-efficacy theory, one conclusion from the present experiment seems to be that performance and self-efficacy do not always go hand-in-hand. This supports the notion of self-efficacy as not simply another measure of ability, but worthy of attention in its own right. In the line of reasoning by Sundström (2006), in a traffic safety context, neither too "over-efficacious" nor "under-efficacious" drivers are wished for, since misjudging one's own ability might make one more accident prone, in the interaction with other road-users.

Although there was only a small difference in improvement of overall driving performance between the control and experimental groups in this study, the results, importantly, pointed in opposite directions: Only participants in the experimental group indicated *higher* self-efficacy, as hypothesized, as their skill developed, whereas participants in the control group actually indicated *lower* self-efficacy, both on the specific parameters assessed and on average. If this tendency persists for a greater sample (and yields statistical significance), serious attention needs to given to the game's feedback and scoring patterns, since a positive learning outcome in general should never be associated with a negative self-related outcome. In any case, taking nothing but the present sample into account, it is notable that the decrease in self-efficacy for the participants in the control group was also linked to an overall more negative experience of the driving session, including feeling less "in control" over the car.

The potentially negative effects of the game in the control condition make a crucial issue for future studies, especially considering that this version has already been used in hundreds of previous driving sessions from which learning outcomes, but not self-efficacy, have been assessed (Backlund et al, 2007). This observation is reinforced by the fact that, judging from the relatively large sample of "gamers" in this study, having much gaming experience had little – if any – influence on how participants were affected by the control and experimental conditions. In fact, the difference between lowered self-efficacy in the control group and increased self-efficacy in the experimental group was even greater among gamers than for the overall sample. That is, gamers *did not* express any greater self-efficacy for the driving task(s) just because they were playing a game, but quite the opposite.

On the positive side, it can be concluded that the driving simulator again proved useful for practicing and improving a range of safe driving behaviours, since not only relatively simple tasks (such as learning to keep the speed by not pressing too hard on the accelerator) but also more complex manoeuvres (such as performing a lane change preceded by a turn-signal while keeping an eye on surrounding traffic) were improved. There was a positive change in performance after the main driving session for all traffic variables analyzed, whether in the control or in the experiment condition.

For the experimental group, there was an associated increase in self-efficacy, which corresponds to the hypotheses of the experiment. Considering this, it is of secondary importance that the differences in driving performance were less clear between conditions, and it remains to be seen whether these discrepancies persevere with more participants and/or in future studies. That the number of correctly performed lane changes, in particular, did not increase more in the experiment condition than in the control condition, may be explained by the fact that the main practice session was completely focused on the game tasks and did not involve any "on command" instructions, as in the reference tests. In addition, for inexperienced drivers, the task for both groups of having to follow the ambulance, might have conflicted with the attention-demanding task of performing proper lane changes.

Another remarkable result was that the participants' evaluation scores on *all* experience items were greater for the experimental than for the control participants. A possible explanation lies in previous findings that positive emotion intensifies experience and facilitates associative thinking through spreading activation of nearby structures (e.g. Olsson, 2003). Thus, having experienced strong emotions might have been expressed in the questionnaire through generally higher scores. If so, this would truly warrant the description of self-efficacy as a "global affective state" (ibid), which has general effects on the appreciation of one's environment. It has also been noted that high perceived self-efficacy make people respond more productively to feedback (Stratton & Hayes, 1999), which possibly made the participants more attentive to the evaluation parameters and thereby rendered higher scorings.

In sum, since none of the differences between the groups measured in the present sample reached conventional statistical significance ($p<0,05$), one must be careful not to draw too grand or general conclusions about the effects of the experimental game design. However, one must neither ascribe too grand importance to significance tests, since they say nothing about the *practical* significance of the findings. As argued by Armstrong (2007), researchers should in general give more attention to effect sizes, replications and meta-analyses. Reflecting the benefits and detriments of a pilot study, the present results are promising for both a continuation and further development of the experimental design.

Future studies

Taking a broad approach entails several issues to be examined and clarified in future studies. Besides a continuation of the pilot study, assembling a greater and more varied sample of participants, and improving the technical details, some examples of follow-up questions would be:

- On a specific scale, *which specific elements and aspects* of the experimental design yield the greatest effects on self-efficacy: e.g., the verbal/visual feedback, the adaptive challenge or the gaming factor?

- On a large scale, how can other sources of self-efficacy be employed in the design, such as learning through modelling and observation, from real or virtual driving teachers? What possible contributions can be made through physiological feedback?

- From a traffic safety perspective, how can the knowledge and empirical support obtained from this type of studies help the development and/or evaluation of self-efficacy based instruments for driving school students?

Finally, it must be pointed out that there is a prominent advantage of employing a design for self-efficacy even if it does not result in improved performance at the time of testing. Self-efficacy is primarily a social-cognitive construct, that is, it tells something not only about how one thinks or performs on a specific task, but how one approaches similar situations and responds within a social context. Therefore, a person's level of self-efficacy has consequences not only for the performance of the particular task, but also for what activities he/she chooses to engage in and the persistence invested in them. In other words, a reasonably high level of self-efficacy also encourages the person to keep on learning and undertake the same tasks again, which were not accomplished the first time. In the case of inexperienced drivers, this might mean that a person, who has never before dared to sit behind the steering wheel, may well be encouraged to take on formal driver's education later on, relying upon the experience with the vehicle which had not been achieved otherwise.

Considering that much more work is needed to understand the complex learning mechanisms when interacting with computers, one may as well design the learning environments to simultaneously teach us more about the learning process. The sense of novelty often accompanied with new technology is a good starting point for studying the effects of a wide range of positive cognitive and emotional states for learning – from promoting the belief that one can, to demonstrating what can ever be done.

References

Armstrong, J. S. (2007). Statistical Significance Tests are Unnecessary Even When Properly Done. *International Journal of Forecasting*, *23*, 335-336.

Backlund, P., Engström, H., Johannesson, M. (2006). *Computer Gaming and Driving Education*. In the Proceedings of the workshop Pedagogical Design of Educational Games affiliated to the 14th International Conference on Computers in Education (ICCE 2006). Beijing, China, December 1.

Backlund, P., Engström, H., Johannesson, M., & Lebram, M. (2007). *Games and Traffic Safety – an Experimental Study in a Game-Based Simulation Environment.* 11th International Conference Information Visualisation (iV07), July 4-6, 2007, Zurich, Switzerland.

Bandura, A. (1977). Self-Efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84,* 191-215.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman.

Britain, S. & Liber, O. (1999). *A Framework for Pedagogical Evaluation of Virtual Learning Environments*. Retrieved 2007-08-13 from http://www.jisc.ac.uk/uploaded_documents/jtap-041.doc

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

Dweck, C. S. (1999). *Self-theories: Their Role in Motivation, Personality and Development*. Philadelphia: Psychology Press.

Gee, P.J. (unpublished manuscript). *Why are video games good for learning?*
Retrieved 2007-08-01 from: http://www.academiccolab.org/initiatives/papers.html

Gee, P.J. (2004). *Situated language and learning: A critique of traditional schooling*. New York: Routledge.

Kay, J. (2001). Learner Control. *User Modelling and User-Adapted Interaction*, *11*, 111-127.

Lebram, M., Engström, H., & Gustavsson, H. (2006). A driving simulator based on video game technology. In *Proceedings of SIGRAD 2006*, November 2006.

Maddux, J. (2005). Self-Efficacy: The Power of Believing You Can. In Snyder, C. R. & S. J. Lopez, (Eds). *Handbook of Positive Psychology*, 277-287.

Middup, C. P., & Johnson, P. (2006). Towards Using Technological Support of Group Memory in Problem-Solving Situations to Improve Self- and Collective Efficacy. *Proceedings of the 39th Hawaii International Conference on System Sciences.*

Mitchell, A. & Savill-Smith, C. (2004). *The use of computer and video games for learning: A*

review of the literature. Learning and Skills Development Agency. Retrieved 2007-08-01 from: http://www.LSDA.org.uk

Olsson, A. (2003). Emotion and Motivation in Learning. *Lund University Cognitive Studies, 112.*.

Quinn, C. N. (2005). *Engaging Learning: Designing e-Learning Simulation Games*. San Fransisco: Pfeiffer.

Reeves, C., & Nass, B. (1996). *The Media Equation*. NY: Cambridge University Press.

Stratton, P., & Hayes, N. (1999). *A Student's Dictionary of Psychology*. London: Arnold.

Sundström, A. (2006). *Utveckling av instrument för att mäta begreppet upplevd förarkompetens.* (BVM No. 22). Umeå universitet: Institutionen för beteendevetenskapliga mätningar. Retrieved 2007-05-02 from: http://www.umu.se/edmeas/publikationer/pdf/BVM%20nr%2022.pdf

Susi, T., Johanneson, M, & Backlund, P. (2007). *Serious Games – an Overview*. Technical Report HS- IKI -TR-07-001, School of Humanities and Informatics, University of Skövde, Sweden.

Van Merriënboer, J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training.* Englewood Cliffs, NJ: Educational technology publications.

Vogel, J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*, 229-243.

Zimmerman, B. J. & Schunk, D. H. (2001). *Self-regulated Learning and Academic Achievement*. Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer,* 38(9), 25-32.

Appendix

Examples of the feedback screens presented to the participants in the experimental (E) and control (C) conditions respectively (in Swedish; the number 0 is substituted in the real game).

Experiment group ("self-efficacious") screenshot: Corresponding control condition screenshot:

Välkommen till simulatorspelet!

Syftet med spelet är att ge dig tillfälle att träna de moment som du tyckte var svåra under den första körningen.
Du påverkar själv svårigheten, genom att utmaningen ökar ju bättre du blir för varje bana.

Kom ihåg att du inte tävlar mot någon annan än dig själv. Låt spelet börja!

Fortsätt

E1: Welcome screen

Välkommen till simulatorspelet!

Syftet med spelet är att förbättra och utvärdera din förmåga att köra bil.
Svårigheten ökar i takt med att du kommer vidare i spelet och du får poäng efter varje bana.

Kom ihåg att du får högre poäng om du gör rätt, men minuspoäng om du gör fel. Låt spelet börja!

Fortsätt

C1: Welcome screen

GRATTIS

Du har nu klarat av första sessionen.

Din körförmåga har redan förbättrats med 7% Sedan första körningen

Ju mer du tränar, desto bättre blir du. Fortsätt så!

Fortsätt

E2: Displayed after half the main driving session.

PAUS

Du har nu spelat första sessionen.

Du har kört 0 banor av spelet.

I nästa session kommer du att få köra fyra banor till. Lycka till!

Fortsätt

C2: Displayed after half the main driving session.

BRA GJORT

Du har nu klarat av andra sessionen.

Din körförmåga har redan förbättrats med 52% Sedan första körningen

Ditt inlärningsbetyg är: VG

Fortsätt

E3: Displayed upon completion of the game.

SPELET SLUT

Du har nu avslutat andra sessionen.

Du har totalt kört 0 banor av spelet.

Fortsätt

C3: Displayed upon completion of the game.