Aangepaste versie hiervan verschijnt als hoofdstuk 15 van de springer versie van de CASA conferentie
Oprins, E.A.P.B. & Korteling, J.E. (2012). Transfer of Training of An Educational Serious Game: The Effectiveness of the CASHIER TRAINER. In CASA conferentie Signapore 2012 (pp. 247-285).

Chapter 13: Transfer of gaming: effectiveness of a Cashier Trainer

Esther Oprins, Hans Korteling
TNO, The Netherlands

Abstract

In this study transfer of gaming is investigated for a particular case, the Cashier Trainer. This is an instructional simulation game with an intelligent tutoring system that can be done individually. Its purpose is to train new cashier employees. These types of games should save time and costs of supervisors during learning at the workplace. Training effectiveness is measured by comparing an experimental Cashier Trainer group with a control group who received conventional on-the-job training (OJT). The candidates from both groups were administrated to the same performance test after a comparable duration of training. Performance and competence were measured by observation and by self-assessment. The results showed differences between the Cashier Trainer condition and OJT: performance and competence were assessed higher for the Cashier Trainer. Especially tasks that occur less frequently were performed better. These results point at a positive transfer of gaming. In addition, the candidates' opinion on the Cashier Trainer was related to performance and competence. Specific didactical features of these types of games such as the automated instruction and feedback, well-structured scenarios and the virtual environment influence learning processes positively. Self-directed learning, experiential learning, and active learning are encouraged and learners are more motivated and self-confident to transfer the acquired competences to the workplace.

13.1. Introduction

Serious games are increasingly gaining acceptance by the education and training community as a potentially valuable, efficient, and effective alternative for conventional training at the workplace. These so called instructional games are here defined as artificial and interactive games that represent (aspects of) the real world that can be used in order to obtain training goals (Sitzmann, 2011). This is especially useful when conventional training, e.g. on-the-job training (OJT) does not provide suitable opportunities because of risks caused by errors, costs of using the 'real' system, legislation, sustainability, training scenario generation, or the limited availability of the operational environment (Korteling et al., 2012). In OJT a supervisor usually coaches the learner while performing tasks in the real-life set-

ting (Oprins et al, 2011). Serious games may thus be a cost-effective, alternative for OJT.

However, this result can only be achieved if there is sufficient transfer of training to the real-world environment. After definitions provided by others for transfer of training (Baldwin & Ford, 1988; Gielen, 1995), we define transfer of gaming as 'the degree to which knowledge, skills and attitudes that are acquired by playing a game can be used effectively in real-life situations'. This definition refers to the notion of competence that we define as: 'applying the whole of acquired knowledge, skills and attitudes effectively in the job' (Oprins, 2008). Evidence-based studies in which transfer of gaming has been verified are hardly available (Peck, 2012). A few quantitative reviews of effectiveness studies for instructional games have been done (Lee, 1999; Randel et al., 1992; Vogel et al.,2006; Ack et al, 2010; Sitzmann, 2011). This research has shown a large diversity in the quality of research and many different results. The effectiveness of instructional games appears to be also moderated by the didactical and technical characteristics of the game design as well as the instructional context in which the game is embedded. Especially transfer is difficult to measure because it requires research at the workplace that cannot easily be controlled (Veldhuis & Theunissen, 2009). Getting evidence on transfer of gaming is important to prove the surplus value of serious gaming and to get insight into the most effective didactical, technical and organizational gaming design properties (Korteling & Oprins, 2012). This research is not only relevant for the game itself but the results could be generalized to other related games.

In this study transfer of gaming is investigated based on a case from the retail industry. Training with the serious game 'Cashier Trainer' developed by Jutten Simulation is compared with conventional on-the-job training in an empirical effectiveness study at the workplace: the supermarket.

13.2. Transfer of gaming

13.2.1 Learning approach

Serious gaming may have many potential didactic and motivational advantages in line with modern learning approaches. In the traditional 'instructor-centered' situation the instructor is dominant. He stands in the center of the attention whereas the learner more or less passively absorbs ('lean back') the information that is provided. Contemporary theories of training and instruction, however, conjecture that learners should participate more actively during classroom lessons in a more 'lean-forward' style. Learners should have an active, central role while the instructors should get a more supportive rather than directive role (Johnston & McCormack, 1996; Petraglia, 1998; Van Merrienboer & Kirschner, 2007). Many contemporary 'constructivist' visions of learning and instruction promote these active forms of learning through experience such as discovery learning (Steffe & Gale, 1995), action learning (Smith & O'Neill, 2003), and experiential learning (Jiusto

& DiBiaso, 2006). Although constructivism may take many forms (Philips, 1998; Petraglia, 1998), an underlying premise is that learning is an active process in which learners are active sense makers who seek to build coherent and organized knowledge. The merit in this premise of learning can be considered as knowledge construction since it is fully consistent with the basics of neuronal development and the functioning of the brain (Hebb, 1949; Meltzoff et al., 2009; Korteling et al., 2012).

13.2.2 Instructional games

Many different types of serious gaming exist and the literature does not show consensus on its definition. Several popular definitions agree that they are entertaining, interactive, rule-governed, goal-focused, and competitive (Vogel et al., 2006; Tobias & Fletcher, 2007; Bell et al, 2008). They serve other goals than entertainment, for instance, learning. The distinguishing feature of serious games is that they are reality based, but they can also incorporate common game features such as rules and competition (Sitzmann, 2011).

In this paper serious games for learning purposes are referred to as interactive 'instructional games' representing (aspects) of the real world. If the games are well-designed, they may be brought to fit very well to the aforementioned constructivistic notions. They can be used to train many competences of workers in a realistic, attractive and challenging manner, using authentic (realistic, practical, job-related) and flexible learning environments in an active and self-directed way (Korteling et al, 2012). Ideally, these games utilize a combination of entertainment and active learning principles to immerse learners in learning. They are assumed to be intrinsically motivating and engaging (Malone, 1981; Csikseszentmihaly, 1990), and enhance the learner's self-efficacy (Bandura, 1997). The learners usually have relatively high own responsibility for learning in accordance with the principles of self-directed learning (Stehouwer et al, 2005). In addition, unlike learning at the workplace, the simulated environment offers the opportunity, to systematically train and practice tasks that do not occur very frequently (Farmer et al, 1999; Van Merrienboer & Kirschner, 2007; Oprins et al, 2008).

Instructional games can have many forms. There exist stand-alone or internet-based games, individual or collaborative games, low-fidelity or high-fidelity games, games with artificial tutoring or with human supervision, etc. This paper focuses on a stand-alone instructional simulation game that must be done individually supported by an intelligent tutoring system (Polson & Richardson, 1988).

13.2.3 Measurement of transfer of gaming

An important question concerns the degree to which instructional games are beneficial and cost-effective since they are mostly developed and used to enhance the effectiveness of training and to reduce training costs. In addition, games itself

may require substantial investments, not only in the design and development of the device itself, but also in the implementation, management, supervision, and education and training of personnel.

A critical assumption underlying the utility of these games is transfer to the real-world environment: transfer of gaming. The degree to which training results in behavioral change on the job is the gold standard of training. This also refers to Kirkpatrick's third level of evaluation, i.e., concerning the enhancement of actual behavior on-the-job based on the training activities (Kirkpatrick, 1998). However, many studies fall short in measuring training effects beyond Kirkpatrick level 1 (reactions) and 2 (declarative learning) Cohn et al., 2009). Also, many studies that have been executed for instructional games have limitations in their design (Ack et al, 2010; Sitzmann, 2011), for instance, focusing specifically on fidelity of the simulation (Allan et al, 1986) or relying on subjective reactions and opinions of trainees, users, or experts (Stehouwer et al, 2005). In addition, studies are often confounded by differences between experimental and control groups in educational context, characteristics of the groups, test- and selection-effects, and preconceived opinions of personnel (Korteling & Oprins, 2012). Finally, it should also be noted that not all instructional games seem to live up to their potential. Hays (2005) has reviewed 48 empirical research articles on the effectiveness of instructional games. This extensive report also includes summaries of 26 other review articles and 31 theoretical articles on instructional gaming. One major conclusion is that instructional games are more effective if they are embedded in adequate instructional programs that include debriefing and feedback. In addition, instructional support during play increases the effectiveness of instructional games. Therefore instructor-less approaches (e.g., web-based applications) must include all instructor functions.

On the basis of a meta-analysis of the literature on instructional effectiveness of computer-based games, Sitzmann (2011) draws similar conclusions concerning the role of instructional support. She also deduces that trainees learn more, relative to comparison groups, when instructional games convey content actively rather than passively and when trainees could access the game as many times as desired. Referring to these recommendations, one important issue for knowledge development concerns the beneficial effects of stand-alone instructional games provided with intelligent tutoring functions (Polson & Richardson, 1988), which are needed to compensate for the lack of human instructional support. Studying this issue will provide more insight into the potential of stand-alone instructional games and (required) supporting components.

These insights from the literature are crucial for all professionals working with serious games or trying to enhance the cost-effectiveness of training by saving on instructional personnel. The present study was aimed at collecting evidence on the transfer of gaming of an instructional simulation game with artificial instructional tutoring and feedback functions. We expect that this type of game will be more cost- effective than conventional on-the-job training.

13.3. Method

13.3.1 Apparatus: Cashier Trainer

For the purpose of this study we used the so-called Cashier Trainer that has been developed by a leading Dutch training enterprise in the retail sector, Jutten Simulation. This Cashier Trainer is being used by various retail organizations in The Netherlands. The game is representative of instructional games with a simulated environment that must be done individually, supported by a virtual tutor. The Cashier Trainer is a low-cost virtual training program combining a simulation of a complete task and a simplified 3D representation of the job environment with didactical principles in an attractive and accessible way, see Figure 13.1.

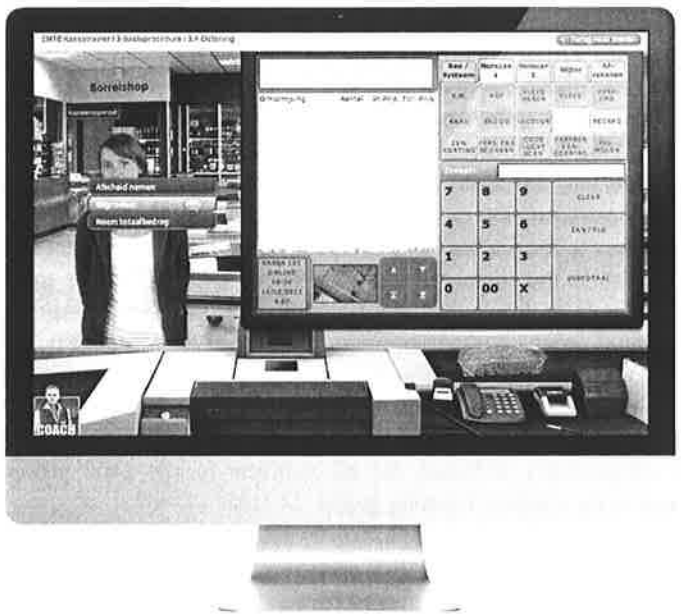

Figure 13.1: The Cashier Trainer (Jutten Simulation)

This Cashier Trainer is a typical instructional simulation game with an intelligent tutoring system. The Cashier Trainer does not only represents the cash desk, as many other related training systems, but it simulates the whole cashier environment including virtual customers. In this way, the game is intended to simulate the entire cashier task and to immerse the learners in the virtual environment of the supermarket. The competences to be acquired refer to operating the cash desk,

communication with customers, dealing with special payments, applying control and safety procedures etcetera. A sophisticated intelligent tutoring system supports the learners with support and feedback independently of a human coach. A scoring system is built in as a special game element for motivating and engaging the learners. The scenarios are sequenced from part-task to whole task training, from only regular tasks to also special tasks. This Cashier Trainer is in use by many retail organizations in The Netherlands to train new cashier employees as a replacement of on-the-job training. The game is internet-based, 3D and runs on each home computer that is connected to the internet.

13.3.2 Experimental design

A total of 45 subjects (7 male, 38 female), who were trained as new cashier employees for a large supermarket chain in The Netherlands, were recruited to participate in the experiment. Subjects were randomly assigned to the conventional OJT group (N=23) and the experimental group trained with the Cashier Trainer (N=22). Average age of all subjects was 20 years. An independent T-test did not show significant differences in average age between the two groups.

The OJT group was trained according to the usual training procedures in the retail company. After introduction of the general procedures of the cashier task they were trained on the real cash desk under supervision of an experienced cashier employee. The experimental group was trained with only the Cashier Trainer. This group gets access to the Cashier Trainer from home circa three weeks before their first working day in the supermarket. Only for the experiment, we tried to fix the training time for both groups at three hours in total in order to make the control and experimental groups comparable with each other for research purposes. Then it could be compared how much was learned in the Cashier Trainer versus OJT within the same time brackets. However, since the study was executed in the real-life practice of supermarket policy, this training time expressed in minutes turned out to be substantially different for all subjects in the OJT group (M=213; SD=129) and in the Cashier Training group (M=188; SD=61).

Both groups were administered to exactly the same performance test. This test consists of handling a customer with a shopping cart of preselected groceries that must be paid. The customer was played by a role-player and the candidates must use the cash desk and apply the correct procedures to handle this customer in the real environment of the supermarket. In this way, an experimentally controlled setting on location was created. The content of the test was based on a task analysis that was carried out in cooperation with the retail company. The assignment for the candidates included regular articles to be scanned, special articles such as price reduction, showing ID for alcohol, vouchers, different payment methods, safety procedures, etcetera.

The task analysis resulted in a set of training outcomes that were measured in various ways in order to make the measurements more reliable. We made a distinction between 'performance' during the test: technical errors directly related to

the task, and 'competence' in a more generic way: the ability of new cashier employees to do their job. Two types of measurements could be distinguished. First, two well-trained observers noted the type of errors during the test and counted them, called 'observed performance', and they assessed general competences at a 5-point rating scale, called 'observed competence'. These observers were educated into the requirements, boundary conditions, and performance criteria of the cashier task. The involvement of two observes makes it possible to calculate inter-rater reliability. Second, the subjects did a self-assessment on the performance variables that are directly related to possible errors during the test, called 'self-assessed performance', and they rated the same set of competences as rated by the observers, also on a 5-points rating scale, called 'self-assessed competence'. The subjects in the Cashier Trainer condition also filled out a questionnaire with items rated on the same 5-points rating scale concerning the Cashier Trainer. This questionnaire includes the training technology (cf. serious gaming), the quality and amount of practicing, the didactical characteristics (e.g. adequacy of instructions), and feeling of reality (cf. immersion).

Finally, structured interviews with retail managers, who had worked with both groups of candidates, were carried out. In these interviews we inventoried subjective observations about task performance and types of errors. We also asked for expected savings in training time with the Cashier Trainer, both for new cashier employees and for the supervisors needed in on-the-job training.

13.4. Results

13.4.1 Observer assessments

The observers assessed both task performance during the test and general competence. They also measured the duration of the task 'scanning 15 regular articles' to get insight into the speed of working. This sub task was chosen to make this measurement maximally objective. Too many factors might influence the total task, such as actions of the customers, to include the whole task. The average speed hardly varies for the OJT group (34,7 seconds) and Cashier Trainer group (35,4 seconds). Figure 13.2 presents the scores of the two groups for the number and type of errors (averaged per sub task; usually only one error could be made per task) made during the performance test, the so-called 'observed performance'.

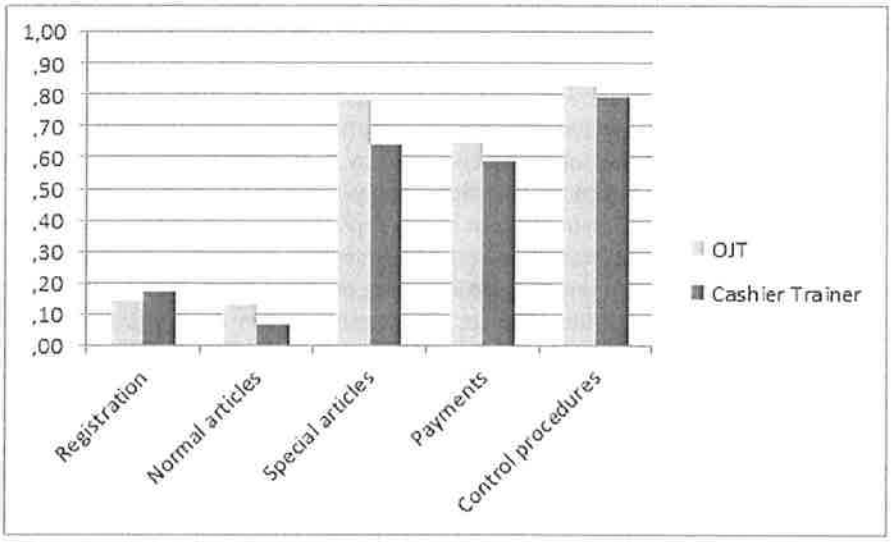

Figure 13.2: Results 'observed performance'

The scores in Figure 13.2 are the averaged error-scores of the two observers. The reliability was assumed to be sufficiently high: the correlation (Pearson, two-tailed) between the scores were .83 significant at $p < ,01$. An independent T-test did not show significant differences between the two groups. In Figure 2 we see that the number of errors is much lower for regular tasks ('registration', 'normal articles') in comparison with tasks that occur less frequently ('special articles', 'control procedures'). The task in 'payments' was also special because it included a combination cash and pin payment.

Figure 13.3 presents the observed scores for the two groups on general competence, called 'observed competence'.

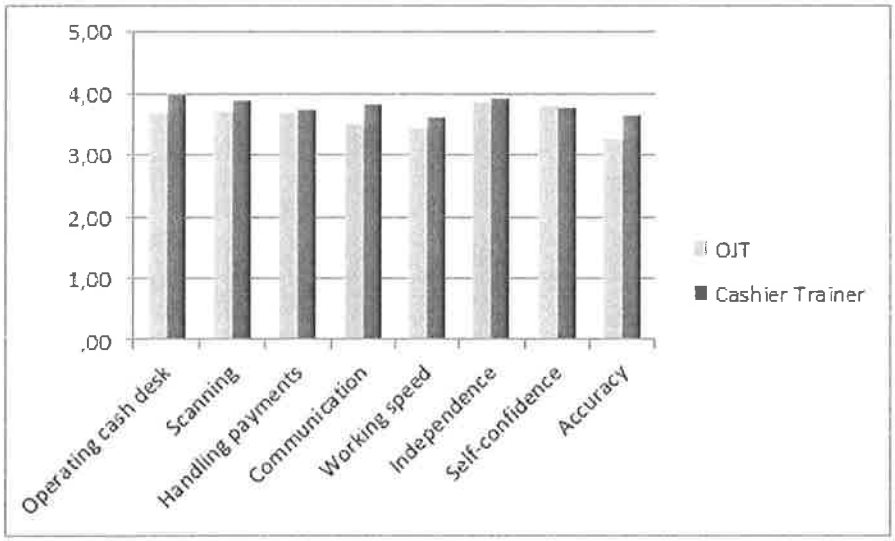

Figure 13.3: Results 'observed competence'

These are also the averaged scores of the two observers. The reliability was assumed to be sufficiently high here as well: the correlation (Pearson, two-tailed) between the scores was .70 significant at p<,01. Again, an independent T-test did not show significant differences between the two groups.

13.4.2 Self-assessments

Figure 4 presents the self-assessments of task performance of the two groups, OJT and Cashier Trainer, called 'self-assessed performance'. The variables are not similar to the variables for the observers (see Figure 13.2) due to the measurement method: the observers counted the errors while the candidates rated their task performance on a rating scale in a more global way. An independent T-test showed that three of the five variables are significantly higher rated in the Cashier Trainer group than in the OJT group: 'cashier station' (T=--2,068, df=43, p=.045), 'payments' (T=--2,015 df=43, p=.050), and 'control procedures' (T=--2,195, df=43, p=.034). Figure 13.4 also shows that the candidates rate 'corrections' relatively low in comparison with the other variables.

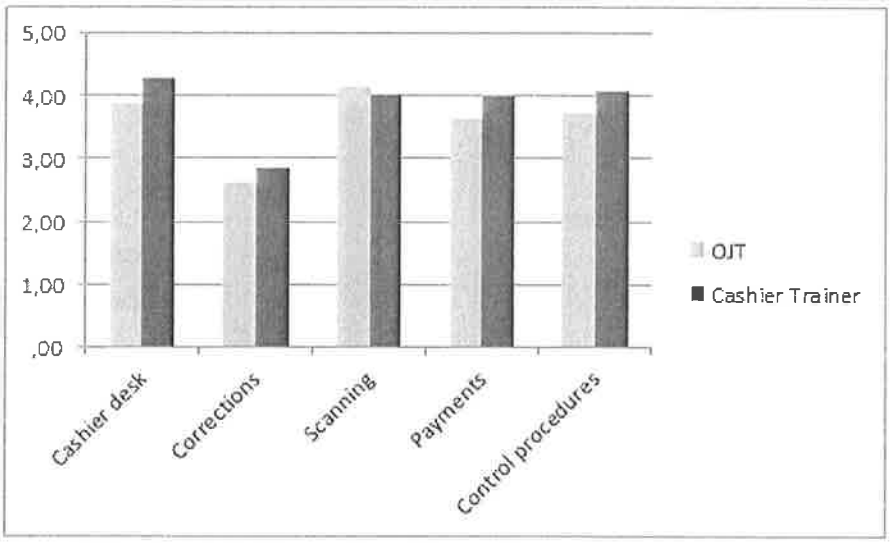

Figure 13.4: Results 'self-assessed performance'

Figure 13.5 presents the self-assessments of general competence of both training groups, called 'self-assessed competence'. The set of rated competences is similar to the 'observed competence' (see Figure 13.3) to keep the results comparable. There were no significant differences found with an independent T-test. It should be noticed that the differences between the rated competences are higher for the self-assessments (see Figure 13.5) than for the observers (see Figure 13.3).

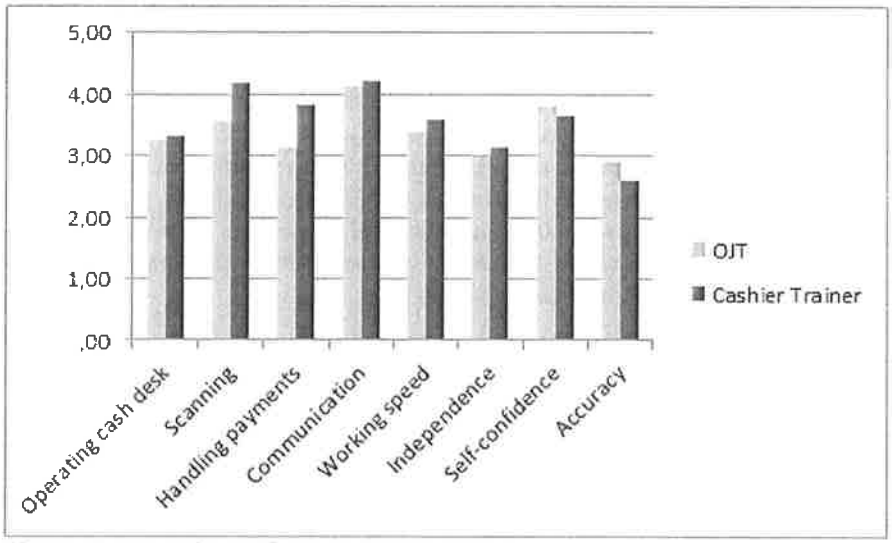

Figure 13.5: Results 'self-assessed competence'

13.4.3 Intercorrelations

Table 13.1 presents the intercorrelations (Pearson, two-tailed) between the various types of measurements (N=45):

Table 13.1: Intercorrelations

	1	2	3	4
1 Observed performance	1	-,55**	-,26	-,18
2 Observed competence	-,55**	1	,26	,31*
3 Self-assessed performance	-,26	,26	1	,80**
4 Self-assessed competence	-,18	,31*	,80**	1

Note. * $p < .05$, ** $p < .01$

Both measurements of performance and competence for the observers are correlated and both measurements for the candidates are correlated as well. We only found a significant correlation between observers and candidates for performance but not for competence.

The candidates' performance was controlled for the influencing factors age, education level, gender and type of job (fulltime, part-time), but no significant correlations were found for any of the four types of measurements, performance or competence. In addition, the amount of time practiced in OJT or in the Cashier Trainer was investigated but this time factor did not influence performance as well.

13.4.4 Didactical features of the Cashier Trainer

The candidates in the Cashier Trainer condition (N=22) have a positive opinion on the Cashier Trainer very high in general. All average ratings, rated at a 5-points scale, are higher than 3.4. Table 13.2 shows the correlations with the various types of measurements. Table 13.2 shows significant correlations for 'quality of practice' and 'realism' but not for 'training technology' and 'didactics'.

Table 13.2: Correlations with opinions

	Training technology	Quality of practice	Didactics	Realism
1 Observed performance	-.08	-.30	-.23	-.23
2 Observed competence	.10	.52*	.22	.32
3 Self-assessed performance	.24	.39	.21	.44*
4 Self-assessed competence	.19	.42*	.21	.38

Note. * $p < .05$

13.5. Discussion and conclusions

13.5.1 Transfer of gaming

The objective of the present study was to investigate the transfer of gaming of the Cashier Trainer. The Cashier Trainer simulates the entire procedural cashier task, immerses the learners in the virtual environment of a supermarket, and is fitted out with an intelligent tutoring system. Training outcomes of the experimental group trained with the Cashier Trainer were compared to a control group trained conventionally on-the-job. The two groups did the same performance test at the workplace. Both observed and self-assessed performance (errors during the test) as well as observed and self-assessed competence were measured. Transfer of gaming refers to the (positive) training outcomes of the Cashier Trainer group in comparison to the OJT group. The results show that the Cashier Trainer candidates obtained generally higher scores on almost all training outcomes than the OJT candidates. The new employees trained with the Cashier Trainer made less errors, received higher observed competence rates, and they assess their own task performance and competence as substantially higher. From these results we may conclude that the new employees in the Cashier Trainer condition generally have learned more than in the OJT condition. This result is even more solid since the mean duration of training was shorter for the Cashier Trainer group than for the OJT group. The Cashier Trainer is also supposed to save on time and costs of supervisor capacity because the new cashier employees can be trained adequately and sufficiently with this stand-alone game as possible replacement of OJT time. In the interviews the supermarket managers estimated these savings on training costs at about 50-70%. This estimation has not been quantified.

In sum, we may conclude that transfer of gaming of the Cashier Trainer is quite positive. Since the Cashier Trainer is representative for instructional simulation games with an intelligent tutoring system, this result suggests that these types of games are effective if well-designed. Because at least the same training outcomes can be achieved, these games can replace expensive and time-consuming on-the-job training.

13.5.2 Type of tasks

To get more insight into the reasons why training effectiveness appeared to be higher in the Cashier Trainer condition, we looked for some trends in the data presented in Figure 13.2-13.4. The relatively easy tasks, such as scanning normal articles and registration at the cash desk, were rated rather high with slight differences between OJT and Cashier Trainer. Apparently, this task is so clear that it can be trained in any form. The most remarkable differences between the two training conditions were found in the tasks that occur less frequently, such as

scanning special articles (e.g. price reduction, alcohol), complex payment procedures (e.g. combination pin and cash) and control procedures (e.g. safety). In the interviews, the supermarket managers also pointed at the difficulties with particular payment and control procedures. These tasks cannot easily be trained systematically in real-life environments because they occur too infrequently. Instructional games offer the possibility to put them in a well-structured sequence of scenarios: one of the reasons to use gaming with simulation instead of OJT (Farmer et al, 1999; Korteling et al, 2012; Oprins, 2008). The intelligent tutoring system (Polson & Richardson, 1988) of the Cashier Trainer offers systematic instruction and feedback. This helps the candidates to train regular tasks and also to practice the less frequently occurring tasks and procedures in part-task training (Van Merrienboer & Kirschner, 2007;). In domains in which safety plays a role and training might cause risks, this is extra important (Oprins et al., 2011). Thus, these results provide evidence for the surplus value of instructional simulation games to train special tasks systematically in a safe and controlled environment with automated instruction and feedback.

13.5.3 Future learning at the workplace

Although transfer of gaming to the workplace is positive, a certain amount of OJT is still needed after having finished the Cashier Trainer. Especially the various self-assessed competences were rated differently. They suggest that the candidates must acquire more routine in operating the cash desk and that working speed, independence and accuracy must be developed further. This is an expectable outcome since only three hours of training is rather low for building routine, in any training system. In addition, the retention time must be taken into account. It should be noticed that the time between doing the Cashier Trainer and the first working day (at which the performance test took place) was circa two to three weeks. This increases the risk of skill loss during this retention period and thus may lead to lower transfer. This risk counts especially for procedural tasks that must be practiced often to acquire routine (Van Merrienboer & Kirschner, 2007). If the Cashier Trainer was used only a few days before the first working day, the differences with the OJT conditions might have been even higher. Procedural tasks must be frequently practiced and repeated just before entering OJT, in general and in serious gaming, to achieve maximal transfer.

In addition to routine tasks, also more general competences must be acquired further during prolonged working at the workplace. The stress component, i.e. rows of waiting customers, is missing in the Cashier Trainer. Specific communication procedures can be trained in the Cashier Trainer but the game does not have natural speech. Automating natural speech for serious gaming is still in development. The customers in the Cashier Trainer are virtual with a limited arsenal of sentences that can be answered by the candidates in a multiple choice menu. Nevertheless, both the observers and the candidates rated communication as rather high both in the Cashier Trainer and OJT condition. Apparently, the virtual cus-

tomers can replace real-life customers for the acquisition of a specific subset of communication procedures despite of a lack of natural speech. This is particularly relevant in domains where practicing with real customers or clients is difficult to realize, for instance, virtual patients may be used within the medical domain (Akl et al., 2010).

The combination of the acquisition of routine tasks and general competences such as communication supports the additional value of using both part-task and whole-task training in serious gaming (Van Merrienboer & Kirschner, 2007), as applied in the Cashier Trainer, which not only represents the technical system (cf. cash desk) but simulated the complete cashier environment.

13.5.4 Observation versus self-assessment

We also investigated the interrelations between the different types of measurements, especially the correlations between the observers' ratings and the self-assessments. Within observers and within candidates the correlations were very high as expected since they refer to their own judgments: both the performance and competence ratings refer to the same task execution. We would also expect a significant correlation between 'observed performance' and 'self-assessed performance' for the same reason but this was not found. The main reason is probably that both measurement methods were different: counting the occurred errors or self-assessing general performance. As expected as well, 'observed competence' and 'self-assessed competence' are strongly related due to way in which the competences were measured (Oprins et al., 2006). These results suggest that the measurements were sufficiently reliable and that it makes sense to combine observed and self-assessed measurements in these types of effectiveness studies.

It is remarkable that the differences between the various self-assessed competences are much higher than for the observed competences. These results suggest that candidates were very good capable to assess their own competences, sometimes even more clear or reliable than external observers. This might be explained by the fact that insight into own performance of the trainees was lacking for the observers. For instance, only the candidates themselves can have a precise feeling about their self-confidence in operating the cash desk independently of the supervisor. The observers can only rely on what they see. This increase the chance on the halo error: the tendency to think of a person in general as rather good or inferior and to color the ratings on specific dimensions by this general feeling (Thorndike, 1920; Oprins, 2008). Probably the observers' ratings are influenced by this rating error that often occurs.

13.5.5 Didactical features

Finally, the candidates' opinion on the didactical features of the Cashier Trainer was related to observed and self-assessed performance and competence. In gen-

eral, the candidates were quite positive on these didactical features. These results were supported by the interviews with the supermarket managers. The new cashier employees were motivated to do the Cashier Trainer, they liked the tutoring and feedback system, and they felt immersed in the simulated environment of the supermarket. In general, they had more self-efficacy and self-confidence at their first working day because they were able to practice both regular and special tasks before in their own tempo. These didactical features are applied in many instructional simulation games with an intelligent tutoring system as argued (e.g. Hays, 2005; Sitzmann, 2011). Thus, the results provide evidence that games are motivating (Malone, 1981; Csikseszentmihaly, 1990), and enhance the learner's self-efficacy (Bandura, 1997). They offer the opportunity for more active and self-directive learning in an authentic learning environment which has been proved to enhance learning processes (Petraglia, 1988; Van Merrienboer & Kirschner, 2007). We would therefore expect that comparable results will be found for other games fitted out with an adequate intelligent tutoring system (Polson & Richardson, 1988) although this study is restricted to only one case.

In this study we correlated the candidates' opinions on the didactical features with the measurements of performance and competence. The results suggest that the applied technology and didactics (cf. tutoring) are less directly influencing performance than the type and amount of practice and feeling of reality or immersion. This supports the idea that fidelity in particular is a very important feature of simulation (Allan et al, 1986; Farmer et al, 1999; Korteling et al, 2012). However, only the candidates' opinion was used without really measuring the underlying learning processes such as self-directed learning, way of practicing, repetition of exercises, etcetera. Therefore, the relationship between the didactical features and learning processes should be investigated in more detail in the future. In order to get more insight into the most effective design characteristics of instructional simulation games, this kind of research should be based on a larger set of comparable instructional games with a variety of didactical features This conclusion confirms similar kinds of suggestions found in previous meta reviews (Lee, 1999; Randel et al., 1992; Vogel et al.,2006; Ack et al, 2010; Sitzmann, 2011).

13.5.6 General conclusion

This is one of the first empirical studies in which real evidence was collected for transfer of gaming in a stand-alone instructional game with an intelligent tutoring system. We may conclude that transfer of gaming of this kind of instructional games, i.e., the Cashier Trainer, can be considerably high. Therefore these games may substantially reduce the amount of required on-the-job training, which may substantially save on time and costs. We suppose that this finding may be promising for future development of similar types of instructional games. In the present Cashier Trainer, the simulated environment offered the opportunity to practice both regular and special tasks in well-sequenced immersive and engaging scenarios. In addition, the well-designed intelligent tutoring system replaced human su-

pervisors, which lead to additional training benefit and cost-savings. Finally, playing these games individually at own tempo may enhance self-efficacy and self-confidence of learners. This, in turn, may contribute to transfer, i.e. application of the acquired competences at work.

Acknowledgements

This study was a Knowledge Transfer Project of the research program 'Games for Training and Entertainment' (GATE) which was supported by the Dutch ICT-Regie and NWO. We would like to thank Jutten Simulation for using the game Cashier Trainer as a case in this study. We also would like to thank the retail company Sligro Foodgroup n.v. in The Netherlands for making it possible to collect data in their EMTÉ supermarkets.

References

Akl EA, Pretorius, RW, Sackett K, Scott Erdley W, Bhoopath P, Alfarah Z, Nemann HJ (2010) The effect of educational games on medical students' learning outcomes: a systematic review. BEME guide no 14

Allan JA, Hays JT, Buffardi LC (1986) Maintenance Training Simulator Fidelity and Individual Differences in Transfer of Training. Human Factors 28: 297-509

Baldwin TT, Ford JK (1988) Transfer of training: a review and directions for future research. Personnel Psychology 41: 63-105

Bandura A (1997) Self-efficacy: the exercise of control. Freeman, New York.
Bell BS, Kanar AM, Kozlowski SWJ (2008) Current issues and future directions in simulation-based training in North America. The international journal of human research management 19: 1416-1436

Cohn J, Kay S, Milham L, Bell Carroll M, Jones D, Sullivan J, & Darken R. (2009) Training effectiveness evaluation: from theory to practice. In: Schmorrow D, Cohn J, Nicholson D (Eds). The PSI Handbook of Virtual Environments for Training and Education

Csikseszentmihaly M (1990) Flow: the psychology of optimal experience. Harper & Row, New York

Farmer E, Rooij J van, Riemersma J, Jorna P, & Moraal J. (1999) Handbook of simulator-based training. Ashgate, Aldershot

Gielen EWM (1995) Transfer of training in a corporate setting. Dissertation, University of Twent

Hays RT (2005) The effectiveness of instructional games: a literature review and discussion. Technical Report 2005-004. Naval Air Warfare Training Systems Division. Orlando, U.S.A.

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Wiley, New York

Jiusto S, DiBiasio D (2006) Experiential learning environments: Do they prepare our students to be self-directed, life-long learners? Journal of Engineering Education 95: 195-204

Johnston S, McCormack C (1996) Integrating information technology into university teaching: Identifying the needs and providing the support. International Journal of Educational Management, 10/5: 36-42

Kirkpatrick DI (1998) Evaluating Training Programs: The Four Levels. 2nd ed, Ashgate, San Francisco

Korteling JE, Helsdingen AS, Theunissen NCM (2012) Serious Games @ Work: Learning job-related competences using serious gaming. In: Bakker A, Derks D (Eds) The Psychology of Digital Media at Work. Psychology Press LTD / Taylor & Francis Group.

Korteling K, Oprins E (2012) True evidence for transfer of training of serious gaming. GATE magazine, 2012

Lee J (1999) Effectiveness of a computer-based instructional simulation: a meta-analysis. International journal of instructional media 26: 71-85

Malone TW (1981) Towards a theory of intrinsically motivating instruction. Cognitive science 4: 333-369

Meltzoff AN, Kuhl PK, Movellan J, Sejnovski TJ (2009) Foundations for a new science of learning. Science 325: 284-288

Merrienboer JJG van, Kirschner P (2007) Ten steps to complex learning: a systematic approach to four-component instructional design. Lawrence Erlbaum Associates, Mahwah NJ

Oprins E (2008) Design of a competence-based assessment system for air traffic control training. Dissertation, Maastricht University

Oprins E, Burggraaff E, Roe R (2011) Analysis of learning curves in the on-the-job training of air traffic controllers. In: D'Oliviera T (Ed.) Mechanisms in the chain of safety. Ashgate Publishing Company, Aldershot

Oprins E, Burggraaff E, Van Weerdenburg H (2006) Design of a competence-based assessment system for air traffic control training. The international journal of aviation psychology 16(3): 297-320

Peck M (2012) Tools or toys? Training games are popular, but no one knows how well they work. Training and simulation journal, dec2011/jan2012.

Petraglia J (1998) Reality by design: the rhetoric and technology of authenticity in education. Erlbaum, Mahwah NJ

Phillips DC (1998) How, why, what, when, and where: Perspectives on constructivism in psychology and education. Issues in Education 3: 151–194

Polson M, Richardson J (1988) Foundations of Intelligent Tutoring Systems. Lawrence Erlbaum Associates, Hillsdale NJ

Randal JM, Morris BA, Wetzel CD, Whitehill BV (1992) The effectiveness of games for educational purposes: a review of recent research. Simulation & gaming 23: 261-276

Sitzmann T (2011) A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology 64: 489-528

Smith PAC, O'Neil J (2003) A review of action learning literature 1994-2000. Part 1: Bibliography and comments. Journal of Workplace Learning 15: 63-69

Steffe L, Gale J (1995) Constructivism in education. Lawrence Erlbaum Associates, Hillsdale NJ

Stehouwer M, Serné M, Niekel C (2005) A tactical trainer for air defence platoon commanders. In: Proceedings of the Interservice/Industry, Training, Simulation, and Education Conference. Orlando I/ITSEC 2005, Paper no. 206

Tobias S, Fletcher JD (2007) What research has to say about designing computer games for learning. Educational Technology 47: 20-29

Thorndike EL (1920) A constant error in psychological ratings. Journal of applied psychology 4: 25-29

Veldhuis GJ, Theunissen NCM (2009) Transfer of Training. Onderzoek naar een maximaal leereffect [Transfer of training. Research on a maximal learning effect]. Opleiding en Training 11: 20-22

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