Serious games for the job training of persons with developmental disabilities

Jungmin Kwon a, *, Youngsun Lee b

a Dept. of Early Childhood and Special Education, Seoul National University of Education, Seocho Gu, Seocho Dong, Seocho Joongang Ro 96, Eduwell Center #503, South Korea
b Dept. of Education, Inha University, 100 Inharo, Nam-gu, Incheon 402-751, South Korea

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form 1 December 2015
Accepted 5 February 2016
Available online 9 February 2016

Keywords:
Simulation
Serious games
Interactive learning environments
Disability
Job training

ABSTRACT

Serious games use strategies that are often applied in special education such as repetition, immediate feedback, and context-based teaching and learning. Serious games are known to increase performance on the target task and related self-efficacy, using a more cost effective method in a safer environment than traditional training (Michael & Chen, 2006). The main goal of our study was to investigate whether playing serious games can affect hands-on performance of persons with developmental disabilities. For the present study, 47 persons with developmental disabilities were divided into three groups including a control. The two experimental groups performed two hands-on tests in different orders: apple packaging and hydroponics (AH, HA). For the first round of tests, the AH group played the Apple Packaging serious game while the HA group played the Hydroponics game. For the second round of hands-on tests, the AH group played the Hydroponics game, while the HA group played the Apple Packaging game. The control group did not play any games between the hand-on tests. Finally, all three groups performed a third round of hands-on tests. The results show that gameplay of the target task increased speed and accuracy of the target hands-on task performance. The main results of this study show that serious games can be used for training simple job skills in persons with developmental disabilities. Serious games can be implemented as part of an ongoing programme to reduce training time and enhance accuracy in a safe and enjoyable environment.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

People with developmental disabilities generally need a significant amount of time and repetitive practice to master a task because of their difficulty with memory, motivation, and attention (Turnbull, Turnbull, & Wehmeyer, 2012). In addition, due to their deficiency in executive control, people with developmental disabilities struggle to sustain task speed and accuracy (Brewer & Smith, 1982). They tend to be physically slower and have poorer motor functions than non-disabled people do (Rarick, 1973). Such qualities reduce the individual productivity of persons with developmental disabilities (Blackorby & Wagner, 1996), and become the major reasons why industrial employers avoid hiring them (Fuqua, Rathburn, & Gade, 1984; Hanley-Maxwell, Rusch, Chadsey-Rusch, & Renzaglia, 1986; Johnson, Greenwood, & Schriner, 1988).

* Corresponding author.
E-mail addresses: jungminkwon@snue.ac.kr (J. Kwon), leeyoungsun@gmail.com (Y. Lee).

http://dx.doi.org/10.1016/j.compedu.2016.02.001
0360-1315/© 2016 Elsevier Ltd. All rights reserved.
To teach job skills and increase productivity, various strategies have been reported to be effective: the use of visual and auditory prompts (Copeland & Hughes, 2000; Minarovic & Bambara, 2007; Post, Storey, & Karabin, 2002), help of peer mentors (Westerlund, Granucli, Gamache, & Clark, 2006), promotion of social skills (Gear, Bobzien, Judge, & Raver, 2011), self-monitoring (Cavaiulo & Gradel, 1990), self-evaluation (Grossi & Heward, 1998), and use of experimenter surveillance (Belfiore, Mace, & Browder, 1989). One commonly used method is job coaching (Doane & Valente, 1977). A job coach assists the person with disabilities one-on-one at the actual workplace. This approach is considered to be more socially and technologically effective than training at sheltered workshops or training institutions (Hempleman & Longhi, 1996; Inge, Dymond, & Wehman, 1996; Langford & Lawson, 1994; O’Reilly, 2007). However, job coaching still requires long-term training (Olson, Cioffi, Yovanoff, & Mank, 2001), and studies have found that the absence of the job coach sometimes reduces productivity (Rusch, Morgan, Martin, Riva, & Agran, 1985). Training at the workplace may also bear physical and social risks depending on the kind of task, place, and person. For these reasons, in countries with limited welfare systems, the additional hiring or repositioning of employees who can provide job coach services may be costly for the company (Department of Education, Employment, and Workplace Relations, 2011; Domzal, Houtenville, & Sharma, 2008).

To reduce such reliance on human resources and to achieve independence and more self-directed learning, utilization of various electronic technologies is one solution. For example, Riffel et al. (2005) used a palmtop PC-based prompting system to complete tasks without human assistance, while Mitchell, Schuster, Collins, and Gassaway (2000) used an auditory prompting system to provide faded prompts until mastery. Portable DVD players (Mechling, Gast, & Fields, 2008), pocket PCs (Laarhoven, Laarhoven-Myers, & Zurita, 2007), computer programs (Hansen & Morgan, 2008), iPod Touch (Cannella-Malone, Weaton, Wu, Tullis, & Park, 2012), and iPad (Alexander, Ayres, Smith, Shepley, & Matars, 2013; Taber-Doughty, Miller, Shurr, & Wiles, 2013) have been used for similar purposes and methods, mostly for self-prompting. However, some authors suggested that while these technologies may reduce reliance on other humans, they do not actually reduce training time (Riffel et al., 2005); additionally, sometimes these prompting systems cannot be phased out of use—in other words, the workers rely on the prompting, and without it, their performance is reduced (Taber-Doughty et al., 2013).

In this study, we propose a completely different approach to using technology. The main goal was to minimize cost, training time, and reliance on human resources, while maximizing physical and social safety by exploring the possibility of job skills training with serious games. A serious game is designed to train skills, attitudes, and/or knowledge through game playing (Michael & Chen, 2006; Ratan & Ritterfeld, 2009; Ritterfeld, Cody, & Vorderer, 2009). Serious games emphasize functional goals as well as entertainment, whereas traditional simulation games mainly emphasize functional goals (Ritterfeld et al., 2009). Today, serious game task training is used mostly in military, medical, and business education; training in all these areas can be economically, physically, and/or socially costly (Guillem-Nieto and Aleson-Carbonell, 2012; Michael & Chen, 2006). In this study, we suggest that serious games can be a good alternative to traditional training methods not only for these reasons but because they are embedded with special education teaching/learning strategies and other elements that satisfy the special needs of persons with developmental disabilities, as explained in the following sections.

1.1. Special education learning/teaching strategies in serious games

Serious games can be an effective training method for persons with developmental disabilities because they utilize learning and teaching strategies commonly used in special education. First, games can be played repetitively (Aldrich, 2009; Gee, 2007). Repetition is one of the most prominent characteristics of gaming. While repetition of a regular task in the physical world may not always be interesting or even possible for the trainee (Sitzmann, 2011), the repetition inherent to games is often enjoyable and motivating. Repetition is known to be the strongest factor influencing memory retention (Hintzman, 1976), making it one of the most important methods for learning among persons with developmental disabilities, who often have short-term and working memory impairments (McCartney, 1987; Turnbull et al., 2012; Vicari, Carlesimo, & Caltagirone, 2008). Persons with developmental disabilities can achieve mastery of tasks through repetition and reduce the short-term cognitive load by automating the skills (Sweller, 2003).

Serious games are played in a context created through a narrative (Graesser, Chipman, Leeming, & Biedenbach, 2009). Lee, Park, and Jin (2006) stated that having a narrative in game playing vastly reduces the player’s cognitive load. Indeed, memory retention is best when information is embedded in context (Schank & Abelson, 1995; Zumbach & Mohraz, 2008). Without the context, the learner’s cognitive load increases and his/her learning efficiency decreases, as the learner must go through the complex process of constantly remembering the procedure and goal of his/her behaviour. Even in special education, creating a narrative for a task increases efficiency, sustains achievement level (Bottge, 1999; Bottge, Heinrichs, Chan, & Serlin, 2001), and improves the motivation and ability to solve complex problems (Goldman & Hasselbring, 1997).

Games provide immediate and coherent feedback (Ritterfeld et al., 2009). Immediate feedback reduces subsequent errors, helps with independent learning (Dihoff, Brosvic, Epstein, & Cook, 2005), and allows longer retention of the learned material (Epstein, Brosvic, Costner, Dihoff, & Lazarus, 2003). Immediate feedback also assists in controlling one’s own learning, thus increasing motivation (Butler & Winne, 1995). In serious games, feedback on the player’s success or failure is presented visually and/or audibly. This feedback can be effective in promoting a sense of accomplishment and in motivating persons with disabilities to modify their behaviour and try again (Oblinger, 2004).

Serious games simulate the physical world. Sitzmann (2011) conducted a meta-analysis of 65 studies on serious games, and reported that people who had been trained through games made better transitions to performing the actual tasks than did those trained with traditional approaches. These findings were confirmed by Wouters, van Nimwegen, van Oostendorp,
and Spek (2013). Training through game playing can be appropriate for persons with developmental disabilities, who often have difficulty in applying the acquired knowledge to the physical world or in generalizing to different situations in a cost-effective and safe way (Michael & Chen, 2006).

1.2. Advantage of serious games in job training

Serious games allow persons with disabilities to practice task completion skills in a physically safe environment (Aldrich, 2009; Boulos, Hetherington, & Wheeler, 2007; Coles, Strickland, Padgett, & Bellmoff, 2007). Even in a simple task like coffee brewing, there is a risk of injury when handling the hot water or steam, and everyday tasks like making sandwiches or cooking pasta require handling sharp knives and using hot stoves. These physical risks can become actual hazards for beginners, but in the serious game environment, risks are minimal. Similarly, virtual worlds such as Second Life can become a safe social space for persons with autism or developmental disabilities to practice their social skills (Boulos et al., 2007; Kandalaft, Didehbani, Krawczyk, Allen, & Chapman, 2013). Serious games may therefore provide a safe environment where persons with disabilities can practice individually and repetitively with fewer chances of social or physical risks.

Serious games are reported to increase self-efficacy related to the task (Kato, Cole Bradlyn, & Pollock, 2008). The studies by Brown et al. (1997); Beale, Kato, Marin-Bowling, Guthrie, and Cole (2007); and Kato et al. (2008) show that games can increase self-efficacy of children with health disabilities and ultimately lead to behaviour change. For persons with developmental disabilities, who often have lower task self-efficacy, serious games may provide a way to increase self-efficacy in job-related tasks.

Using serious games for job training is cost effective. Although developing a serious game requires financial investment, the resulting costs are much lower than are those of the actual training (Michael & Chen, 2006; Susi, Johannesson, & Backlund, 2007). Orlansky, Taylor, Levine, and Honig (1997) and Cohen et al. (2010) reported that the cost of using a serious game in training is a fraction of the training costs incurred without the use of any type of simulation. As a job coach usually teaches one or a few persons, the costs for training a person with disabilities can be discouraging for some employers. For such employers, serious games could be one way to save costs and time.

Finally, serious games have been reported to improve task completion speed and accuracy. The competition element is known to improve performance speed, and serious games use competition against others (or time) to motivate the player to complete the task in a shorter time (Bryant & Fondren, 2009). According to Brewer and Smith (1982), limitation of time improves the processing speed of persons with developmental disabilities. Therefore, time restriction or competition with others can motivate players with developmental disabilities to improve their speed. Serious games can help players to improve accuracy in task completion by providing explicit, procedural, and implicit knowledge (Aldrich, 2009). For example, auditory cues provide information about the player’s situation; alternatively, the ‘virtual mentor’ in the game can directly instruct or guide the player. Games allow beginners to easily follow step-by-step instructions that are similar to task analysis strategies in special education. Games also utilize alarms and important cues such as hotspot, media controls, sound cues, guide, noise, abstraction, alert, switch, mouse-over, scaffolding, and game tips to help the player perform accurately (Aldrich, 2009; Gee, 2007). In many cases, the player cannot move to the next level before completing the previous level; thus, the player is motivated to accurately complete the given tasks.

1.3. Purpose of the study

Despite the wide use of simulation and serious games in training professionals for job skills, it has rarely been applied for job training of persons with disabilities. Past studies have focused on physical rehabilitation (e.g., Burke et al., 2009), health behaviour (e.g., Beale et al., 2007; Isasi, Basterretxea, Zorrilla, & Zapirain, 2013), sensory perception (Hussaan, Sehaba, & Mille, 2011), and social training (Parsons, Leonard, & Mitchell, 2006). To date, there has been one study of serious games for job-related training for persons with developmental disabilities, conducted by Lanyi, Brown, Standen, Lewis, and Butkute (2012). In their study, the authors provided usability results of a series of serious games they had developed. Their games focused on skills related to work life such as learning fractions and percentages, personal hygiene, money management skills, anger management, and stress at work. In their report, the games were tested for enjoyment, usability, manageability, and graphics; however, the effects of the games on the target skills or knowledge were not tested.

For this study, a team of game designers developed a job training serious game for persons with developmental disabilities to examine whether it could improve the trainees’ skills. The game was based on the job training textbook published by the Korean government, and was distributed to all schools within the country. According to national data (Korea National Institute for Special Education, 2011), over 90% of special education teachers were not using the book because it was unrealistic to learn hands-on skills with written text and pictures. Therefore, it appeared to make sense to convert this well-organized but ‘dead’ text into a hands-on simulation game. The resulting game was titled ‘Adventures on Coolong Island’. The game covered job skills in four different areas: office, farm, grocery, and factory. Among the more than 300 job skills described in the textbook, we chose a total of eight—two from each area—suitable for a low-budget game format, to be developed into mini games. The eight mini games were Apple Packaging, Hydroponics, Cafeteria Food Distribution, Pen Assembly, Wood Work, Mail Delivery, Automated Teller Machine (ATM), and Grocery Assistant.

For the current study, we selected two mini games (Apple Packaging and Hydroponics) to examine the effect of gaming on actual task performance. Both mini games were simple repetitive tasks that persons with developmental disabilities could
easily perform with some practice. We selected two mini games, rather than one, to measure at least two sets of outcomes to reduce the chance of regression and therefore increase internal validity. The goal of this study was to measure the effects of the two mini games on the hands-on performance of persons with developmental disabilities. Specifically, we focused on individual productivity of hands-on performance that was measured by speed and accuracy.

2. Methods

2.1. Participants

The participants were recruited from five high schools in the metropolitan area. Participants who were able to play a computer game and consented to participate were included in the study. Participants were 47 adolescents and young adults, ($M_{\text{age}} = 16.67$, $SD = 0.87$; age range 15–19 years), 10 female (21.3%) and 37 male (78.7%). They were all diagnosed with developmental disabilities based on standardized criteria. Seven participants had autism spectrum disorders and four participants had Down’s syndrome. As standardized test scores were not available for all participants, they were not considered as a grouping factor in the study. Instead, the student ability grouping was based on teachers’ reports on whether the student needed more or less support in daily affairs. Teachers rated each student’s need for support when performing daily functional skills as more support (low ability), moderate support (moderate ability), or less support (high ability). Table 1 summarizes the characteristics of all participants.

2.2. Ethical considerations

Issues of consent were addressed as follows. The full research proposal was sent to and approved by the university’s institutional review board. Written permission from participants, their parents, and teachers were obtained. The permission letter sent to participants, parents, and teachers included a detailed explanation of the study, the benefits and risks of the experiment, and the contact information of the first author. Out of 50 parents contacted, 48 responded to the letter. Among the 48, one participant did not provide consent, resulting in 47 participants. The participants and their teachers were provided with extra details about the study by the experimenter in person. Anonymity was promised to all participants, teachers, and schools. All data gathered were kept secure with double passwords, and only the researchers and a research assistant were given access.

2.3. Research design

Based on Coles’s et al. (2007) study, which tested the effect of two games using a crossover design, this study used randomized, repeated test waves with crossover intervention and a control group design to investigate the effects of gaming. All 47 participants were divided into one of the following groups according to their functioning level reported by the teacher: high, moderate, and low. Within each group, each participant was randomly assigned to one of the three conditions: AH (Apple-Hydro), HA (Hydro-Apple), and control. Throwing dice was used for random assignment. At the most basic level, we were interested in the two game groups: AH and HA, which served as each other’s control group. However, due to the nature of the tests, which involved repetitive actions, it was necessary to also measure the effect of repeated testing itself. Unfortunately, multiple measures analysis was not possible due to the intervention crossover. Therefore, the purpose of the control group was to show the natural gain from repeated testing. Fig. 1 illustrates the sequence of the events in the research design. All participants took a pre-test that was a hands-on performance test (two tasks) using real objects (Test Wave1). The participants in the AH group then played the Apple Packaging game while those in the HA group played the Hydroponics game, followed by a second hands-on test (Test Wave2).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Participant characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control ($n = 16$)</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>Age</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>SD</td>
</tr>
<tr>
<td>Ability</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Low</td>
</tr>
</tbody>
</table>

Note. AH, Apple–Hydro group; HA, Hydro–Apple group; SD, standard error.
Then, the AH and HA groups crossed the interventions over and participated in the Hydroponics game and the Apple Packaging game, respectively. Finally, the groups participated in the third hands-on test (Test Wave3). While the AH and HA groups played games, the control group did not participate in any game or task-related activity but only participated in the hands-on tests. In summary, all three groups performed the same two hands-on tests three times. While the control group did not receive any intervention, the AH group and the HA group played the two different games, each game crossing over the two groups between tests.

2.4. Settings

The experiments were conducted in vacant classrooms at the participating local schools. Two classrooms per school were used. The first room contained the testing stations, where two sets of desks were set up—one for task 1 (apple packaging) and another for task 2 (hydroponics). At the apple packaging station, two boxes (one box holding 14 apples and the other with 10 empty slots), a trash bin, and some packaging materials were set up on the desk, similar to what could be found in the Apple Packaging game. Fig. 2 shows the example of the apple packaging station setting and a screenshot of the Apple Packaging game. At the hydroponics station, there were 20 sets of baby plants, 40 wrapping sponges, a hydroponics box with a lid, and a bucket of enough water to fill the hydroponics box. A testing administrator (a graduate student), who was blind to the experimental condition of each participant, administered both tests. Each time a student entered the room, he/she performed both the apple packaging and hydroponics test.

In the second room, three stations were set up: a game station with a 17-inch laptop and a mouse with the game ready to be played, a video/music station with a laptop with YouTube videos and MP3 music ready to watch or listen to, and a comic book station with six comic books from the library placed on the desk to choose from. An intervention administrator (another graduate student), who administered both game interventions and non-game related activities, was also present in the second room. All stations were equipped with a video camera to record all procedures.

2.5. The games

Two games were used in this study. The goal of the Apple Packaging game was to wrap all apples on the screen within the given time (5 min). The player was to pick up an apple, wrap it with a wrapper, and then place it in an appropriate slot in the box, by clicking once on the right button of the mouse. The goal of the Hydroponics game was to wrap all plants on the screen within the given time (5 min). The player was to pick up a baby plant, wrap it with a small piece of sponge, stick it in the slot of the hydroponics box, and then pour water up to requisite level. Both games were repetitive in nature and utilized time limitations as a motivator. Texts provided the background story for the game, but it was not necessary to read the text to play the game. The game also offered minimal written directions but did not provide corrective feedback or prompts when the player failed to make any moves. The game provided feedback for correct moves (an increase in points) and incorrect moves (a beep and a 3-s pause penalty). When the task was completed in time, a fanfare with fireworks over the game screen appeared for 5 s and asked the player to either quit and go back to the main menu or to play again. When the player failed to complete
2.6. Procedure

Participants scheduled an appointment and visited the experiment rooms during school hours. All of the participants had been randomly assigned to one of the three groups before the visits began. When the student checked into the testing room, the primary investigator introduced herself, asked the student’s name, then briefly explained what the student would be doing during the session. The main purpose of the study was not revealed to the participant. However, we told the participant that we were interested in how well they performed certain job skills. Then, the student was brought to the first testing station and the primary investigator left the room. The test administrator at the station explained to the student what to do. The instructions were written on a piece of paper with graphic examples as a reference for both the administrator and the
participant, and every student received the same set of instructions orally and visually. The instructions were provided in a
direct manner and in detail so an average participant could perform the test without practice or pre-experience. The in-
structions had previously been tested with three non-disabled 7-year-olds and six high school students with developmental
disabilities during a pilot stage to ensure the instructions were understandable and easy to follow even if the participant did
not have any related experience or knowledge.

The participant began at the sound of ‘begin’. Originally, we had planned for the participant to say, ‘done’ to notify that he
or she was finished with the given task. However, we found that during the pilot stage some participants would not say
anything even after they were done. Thus, the experimenter stopped the stopwatch as soon as the participant finished
performing the last step of the task with or without the participant saying ‘done’. The participant performed both the apple
packaging and hydroponics tests during the three testing sessions. The order of the tests during each session was randomized.
During all test sessions, the participants were not given any feedback (correct or incorrect) but received neutral feedback on
their effort (‘Thank you for your hard work’ or ‘I appreciate your effort’). The neutral feedback was given at the end of each test
set and not systematically randomized but given at the administrator’s liberty.

After the test session, participant moved to the intervention room. There, according to the condition group, he or she
participated in one of the three activities: leisure activity, the Apple Packaging game, or the Hydroponics game. The
participant was exposed to the activity for 15 min. Most participants in the game group were able to play the game for two full
rounds during this period. The participant was briefly shown how to play the game by the experimenter (i.e., ‘Take an apple,
wrap it in the wrapper, and place it in the box. Rotten apples go into the trashcan’). All participants were able to play the game
independently after being shown how to play. For the control group, the participants freely chose to read a book, listen to
music, or watch a short video for 15 min in between the test waves.

2.7. Data collection

Two types of data were collected: accuracy (number of correct moves) and speed (time to finish). The main data collectors,
the testing and intervention administrators, collected the data at the time of experimentation. The numbers of errors were
counted using a scoring protocol. The time was measured with a stopwatch. The data collectors silently recorded the result in
writing and did not provide immediate score information to the participant. In addition to the onsite data collection, video
data were collected throughout the study. The video data were observed by a secondary scorekeeper and used for reliability
purposes.

2.7.1. Accuracy

For the apple packaging test, the total number of possible errors was 14. Each apple placed in the packaging box in an
incorrect manner (e.g., apple upside down, without packaging material wrapped around, or outside of the slot) was counted
as one error. Apples that were misplaced (e.g., fresh apple in the trash bin or rotten apple in the packaging box) were each
counted as one error. Since 11 fresh apples needed to be wrapped to go into the packaging box and three rotten apples needed
to go into the trash bin, a total of 14 errors could be made during the test. If the student did not do anything, the number of
errors was counted as 14. If a student correctly packed five apples and then said, ‘finished’, or stopped for longer than 3 s, the
number of errors was nine, since six slots were still empty and three rotten apples were not in the trash bin.

For the hydroponics test, the total number of possible errors was 13. There were 12 young plants that needed to be
wrapped with sponge and put into the hydroponics box. Each plant placed into the hydroponics box in an incorrect manner
(e.g., plant upside down, without sponge around it, or outside of the slot) was counted as one error. The remaining point was
for pouring water into the box. The act of pouring water was counted as one point in the accuracy data, but time taken for
pouring water was excluded from the speed data for two reasons: First, the time required for pouring varied widely between
participants—some pouring all the water at once while others pouring droplets of water for an extended period of time—
therefore seriously distorting the speed data. Second, the pouring method did not match between the game and hands-on
task. In the game, the water runs through a garden hose and the heavy flow thus fills the hydroponics box in a matter of
seconds. In the hands-on task, it was not possible to replicate the game environment because the experiment was conducted
in classrooms. However, knowing that hydroponic plants need water was an important part of the knowledge. Therefore, it
was reasonable to count the act itself but not the time it took for research purposes. For data analysis, the correct percentage
(rather than error percentage) was used for accuracy.

2.7.2. Speed

Measuring time began when the test administrator said, ‘begin’ and ended as soon as the student finished performing the
last step of the task. For participants who did not finish the task, the administrator clicked the stopwatch when the student
did not do anything for more than three seconds. Then, the administrator asked the student whether he or she was finished to
confirm. If the student said ‘yes’, three seconds were subtracted from the stopwatch and marked. When the student said ‘no’
and continued performing the task, the administrator timed the additional time. The iOS clock application on Apple iPhone 5
was used as the stopwatch, as it allowed the user to record several stop times. The same timing procedure was used for both
tasks. For data analysis, speed was calculated as the number of completed moves (e.g., properly wrapped and placed apples)
divided by the total time taken. This excluded all incomplete moves (e.g., no sponge around plant), and if there were no
complete moves, the speed was calculated as zero.
2.8. Reliability

To ensure reliability of the collected data, we performed an inter-observer reliability assessment using the recorded videos from the testing sessions. Because there were 47 participants and each participant performed three rounds of tests, there were 282 5–15 min video clips. This accounted for a total of 2368 min (39.47 h) of standard quality video data. Among the 282 clips, 94 (33.33%) clips were randomly selected for training and reliability measurements. The selected clips were converted to Quicktime format using a MacBook Pro A1286 laptop (hardware) and Final Cut Pro X (software). Among the 94 clips, 10 were used for reliability training purposes. The remaining 84 clips (about 591 min), which accounted for approximately 29.79% of the number of total video clips and 24.96% of total time, were used for assessing inter-observer reliability. A research assistant independently measured the participants’ accuracy and time using this video data. Only when the two observers completely agreed was it counted as agreement in accuracy. For speed, a margin of ±2.0 s was counted as agreement. Inter-observer reliability for accuracy was 97.62% (agreement on 82 trials out of 84); for speed, it was 96.43% (agreement on 81 trials out of 84).

3. Results

3.1. Data analysis framework

In this data analysis, we took the block design structure into account. The crossover study design yielded three groups of data: gain scores of the control group who did not participate in any gaming activities, gain scores of those who played the game matching the task, and gain scores of those who played the other game that did not match the task. For the apple packaging task, we combined Wave2-Wave1 responses from the AH group with the Wave3-Wave2 from the HA group to form the ‘Matching Game Group’. Similarly, we combined Wave3-Wave2 responses from the AH group with the Wave2-Wave1 from the HA group to form the ‘Other Game Group’. Our main interest was to investigate whether the group who played the matching game increased performance of the target task. Therefore, the gain scores of each experimental group were compared to the control group (‘No Game Group’). A significant difference in gain scores in one group but not the other would indicate that playing the target game affected performance of the target task.

3.2. Performance speed

The mean speed, SD, median, and range for the control, AH, and HA groups for all test waves are described in Table 2. The shaded blocks signify the task performance immediately after playing the matching game. Levene’s test of homogeneity of variances was not significant for both tasks (F(2,91) = 1.885, p > 0.05 for apple packaging task and F(2,91) = 1.956, p > 0.05 for hydroponics task). Therefore, a between-subjects one-way ANOVA was conducted to compare the effects of gaming on task performance speed. The ANOVA revealed significant differences between the groups for both apple packaging (F(2,91) = 3.639, p < 0.05, η² = 0.075) and hydroponics (F(2,91) = 4.368, p < .05, η² = 0.087).

A post hoc analysis using the Tukey HSD test was conducted to find treatment effects. For the apple task, significant differences were found between No Game Group and the Matching Game Group at p < 0.05, while no significant differences were found between the No Game Group and the Other Game Group. For the hydroponics task, significant differences were found between the No Game Group and the Matching Game Group at p < 0.05, while no significant differences were found between the No Game Group and the Other Game Group (Table 3).

3.3. Accuracy

The mean accuracy, SD, median, and range for the control, AH, and HA groups for all test waves are described in Table 4. The shaded blocks signify the task performance immediately after playing the matching game. Levene’s test of homogeneity of variances and Welch’s F were significant (p < 0.05) and failed to meet the assumptions for a parametric test. Therefore, a non-parametric equivalent of an ANOVA, the Kruskal-Wallis H test, was conducted to examine the effects of gaming on task performance accuracy. The results indicate that there were statistically significant differences in accuracy scores between the No Game, Matching Game, and Other Game groups: \( \chi^2(2) = 6.517, p = 0.038 \) for the apple task and \( \chi^2(2) = 32.831, p < 0.001 \) for the hydroponics task. Mean ranks for the apple task were 45.86 (No Game), 55.16 (Matching Game), and 41.53 (Other Game). Mean ranks for the hydroponics task were 39.30 (No Game), 63.94 (Matching Game), and 39.53 (Other Game).

To examine the treatment effects, we conducted a Mann-Whitney U test. The post hoc test revealed that for the apple task, playing the target game had no significant effect (\( Mdn = 0.00, U = 394.5, p > 0.05, r = 0.22 \)); while for the hydroponics task, playing the target game yielded strongly significant effects (\( Mdn = 0.00, U = 232.5, p < 0.001, r = -0.59 \)). On the other hand, playing the other game revealed no significant effects on both tasks (\( Mdn = 0.00, U = 447, p > 0.05, r = -0.12 \) for the apple task and \( Mdn = 0.00, U = 495, p > 0.05, r = -0.004 \) for the hydroponics task) (Table 5).
The goal of this study was to investigate whether participants with developmental disabilities would benefit from playing a job-training serious game. To measure effectiveness, we tested whether playing two separate mini games of two individual tasks (apple packaging and hydroponics) affected hands-on task performance in terms of speed and accuracy. The results indicate that playing the matching game affected the target task performance speed. Such effects were witnessed in both apple packaging and hydroponics tasks. However, for performance accuracy, playing the matching game strongly affected the hydroponics task only; while the apple packaging task was not significantly affected.

We suggest several possibilities why the Apple Packaging game yielded little effect on performance accuracy, while the effects of the Hydroponics game were strong. The discrepancy may have been due to the different levels of difficulty of the two tasks. In the national job curriculum for special schools, apple packaging is included in Chapter 2 Section 1, while hydroponics is included in Chapter 3 Section 4 (Korean Ministry of Education, Science, and Technology, 2010). The school-centred vocational curriculum tends to be hierarchical, with easier tasks that require basic skills appearing earlier in the curriculum and the more difficult tasks that require advanced skills appearing later in the curriculum (Korea National Institute of Special Education, 2005). Based on the asserted difficulty of the two tasks, one interpretation of the results may be that games are more effective for difficult tasks than for easy tasks when trying to achieve high accuracy.

The second possibility is the ceiling effect, although this could be regarded a continuum of the first effect. We suspect that there was a ceiling effect for the apple packaging but not for the hydroponics task. Evidence that supports our speculation was found in the distribution of accuracy data at initial testing (Test Wave1). Accuracy scores with the highest frequency (48.9%) for apple packaging consisted of the total available points (14), which resulted in a J-shaped distribution, while for hydroponics, the highest frequency (53.2%) was at a lower score (12.9), which resulted in a bell-shaped distribution. Ceiling effects compromise statistical effects of the experimental data that may lead to the mistaken conclusion that the independent variable has no effect (Cramer & Howitt, 2004, p. 21).

For these reasons, we limit our discussion on the hydroponics results only. The findings from the hydroponics task show that performance speed and accuracy increased significantly when the Hydroponics game was played immediately before and after playing the Hydroponics game did not affect performance in the apple packaging task, and vice versa. Therefore, it is safe to assume that gameplay was directly responsible for the increase in performance of the same task.

The effect of gameplay on task performance has been documented repeatedly in previous studies in the general population. Studies that measured speed and/or accuracy after training with a serious game for medical training found that the game indeed increased task performance (Creutzfeldt, Hedman, & Fellander-Tsai, 2012; Knight et al., 2010; Qin, Chui, Pang, Choi, & Heng, 2010; Seymour et al., 2002; Tuggy, 1998). Connolly, Boyle, MacArthur, Hainey, and Boyle (2012) claimed that this is due to the fact that the main outcomes from playing serious games are knowledge acquisition and content understanding. With increased knowledge and understanding of the target task, performance speed and accuracy are likely to improve as well.

For persons with developmental disabilities, the increase in speed and accuracy after a short period of gaming reveals new opportunities for the training of skills in this population. Previous research in developmental disabilities shows that training for a simple task is time consuming, as it requires a large number of repetitions (e.g., Devlin, 2011; Godsey, Schuster, Lingo, Collins, & Kleinert, 2008; Mechling & Ortega-Hurndon, 2007; Mitchell et al., 2000; Riffel et al., 2005). Extensive repetition is how persons with developmental disabilities achieve task mastery. As we have speculated in the Introduction, the repetition of game playing seemed to have assisted players with developmental disabilities in acquiring knowledge of the task. In addition, various studies have shown that video games can increase visual and attentional skills, resulting in better visual short-term memory (Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Dye & Bavelier, 2004; Hubert-Wallander, Green, & Bevelier, 2011). The visual modality, which is considered to be effective in enabling persons with developmental disabilities to enhance understanding and remember information (Lee, Amos, et al., 2006), and the time constraint as part of the game, which is known to improve processing speed of persons with developmental disabilities (Brewer & Smith, 1982), likely produced the performance increase in the hands-on task in a relatively short time.
The findings from our study imply that games can be a useful tool for training new vocational skills for persons with developmental disabilities. This is relevant because these persons often experience impairments in memory and self-efficacy when learning new tasks, which can result in low performance and learned helplessness (Turnbull et al., 2012). Games present an opportunity to break this vicious cycle by providing learning-by-doing (Aldrich, 2009), feelings of success, and motivation when performing vocational tasks.

There are several limitations to our study. First, the tested tasks were simple repetitive tasks that did not require complex problem-solving behaviour. In any real-world working environment, however, problem solving is required to manage, sustain, and succeed in the job. Thus, for a future study, it will be crucial to investigate the effects of more complex, problem-solving games for persons with developmental disabilities. Second, we did not compare traditional training methods with game-based training in this study. As has been done in studies of pilot and medical simulations, a future study should directly compare traditional and game-based training methods with respect to effectiveness, especially in terms of time and costs. Finally, our study focused on short-term effects. The games were played for only 15 min per task, and we did not measure maintenance effects. It is important for future studies to test whether the skills learned through gaming were maintained after a prolonged period.

The main results of this study show that serious games can be used for training simple job skills in persons with developmental disabilities. However, we do not claim that games or simulations can completely replace traditional training methods, including job coaching. Instead, we suggest that serious games be used as part of an ongoing program to reduce training time and enhance performance within a relatively short time in a safe and enjoyable environment for persons with developmental disabilities.

References


Table 3
Post-hoc analysis of treatment effects for performance speed.

<table>
<thead>
<tr>
<th>Task</th>
<th>Mean difference</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Game vs. Matching Game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>0.028</td>
<td>0.012</td>
<td>0.047*</td>
<td>−0.056 to −0.003</td>
</tr>
<tr>
<td>Hydro</td>
<td>0.016</td>
<td>0.006</td>
<td>0.011*</td>
<td>−0.029 to −0.003</td>
</tr>
<tr>
<td>No Game vs. Other Game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>0.002</td>
<td>0.012</td>
<td>0.009</td>
<td>−0.029 to 0.026</td>
</tr>
<tr>
<td>Hydro</td>
<td>0.007</td>
<td>0.006</td>
<td>0.06</td>
<td>−0.02 to 0.006</td>
</tr>
</tbody>
</table>

*p < 0.05.

Table 4
Summary of accuracy comparisons between control, AH, and HA groups.

<table>
<thead>
<tr>
<th>Wave</th>
<th>Control</th>
<th>AH</th>
<th>HA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wave1</td>
<td>Wave2</td>
<td>Wave3</td>
</tr>
<tr>
<td>Apple</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>11.38 (4.65)</td>
<td>11.75 (4.43)</td>
<td>11.94 (3.94)</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>13.50 (0−14)</td>
<td>14 (0−14)</td>
<td>14 (2−14)</td>
</tr>
<tr>
<td>25% Quartile</td>
<td>11.5</td>
<td>11.5</td>
<td>11.5</td>
</tr>
<tr>
<td>75% Quartile</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Hydro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>10.77 (5.36)</td>
<td>10.70 (5.33)</td>
<td>10.70 (5.33)</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>12.92 (0−14)</td>
<td>12.92 (0−14)</td>
<td>12.92 (0−14)</td>
</tr>
<tr>
<td>75% Quartile</td>
<td>13.19</td>
<td>12.92</td>
<td>12.92</td>
</tr>
</tbody>
</table>

Table 5
Post-hoc analysis of treatment effects for performance accuracy.

<table>
<thead>
<tr>
<th>Task</th>
<th>U</th>
<th>Z</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Game vs. Matching Game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>394.5</td>
<td>−1.752</td>
<td>0.080</td>
<td>−0.221</td>
</tr>
<tr>
<td>Hydro</td>
<td>232.5</td>
<td>−4.643</td>
<td>&lt;0.001*</td>
<td>−0.585</td>
</tr>
<tr>
<td>No Game vs. Other Game</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>447.0</td>
<td>−0.936</td>
<td>0.354</td>
<td>−0.117</td>
</tr>
<tr>
<td>Hydro</td>
<td>465.0</td>
<td>−0.033</td>
<td>0.974</td>
<td>−0.004</td>
</tr>
</tbody>
</table>

*p < 0.05.


