Running head: Engagement enhances skill learning

Engaging environments enhance motor skill learning in a computer gaming task.

Lohse, K. R., Boyd, L. A., & Hodges, N. J. (2016). Engaging environments enhance motor skill learning in a computer gaming task. *Journal of motor behavior*, *48*(2), 172-182.
https://www.tandfonline.com/doi/abs/10.1080/00222895.2015.1068158

https://doi.org/10.1080/00222895.2015.1068158

Keith R. Lohse[1], Lara A. Boyd[2], & Nicola J. Hodges[3]

[1]School of Kinesiology, Auburn University

[2]Department of Physical Therapy, University of British Columbia

[3]School of Kinesiology, University of British Columbia

Key Terms: motor learning; video games; engagement; motivation
Word Count: 5,026
Contents:

Corresponding Author:
Keith Lohse, PhD
School of Kinesiology
Auburn University
301 Wire Road
Auburn, AL, 36830
USA
p: 1-334-844-1982
e: lohse@auburn.edu

Abstract

Engagement during practice can motivate a learner to practice more, hence having indirect effects on learning through increased practice. However, it is not known whether engagement can also have a direct effect on learning when the amount of practice is held constant. To address this question, 40 participants played a video-game that contained an embedded repeated sequence component, under either highly engaging conditions (the “Game” group) or mechanically identical but less engaging conditions (the “Sterile” group). The game environment facilitated retention over a 1 week interval. Specifically, the Game group improved in both speed and accuracy for random and repeated trials, suggesting a general motor-related improvement, rather than a specific influence of engagement on implicit sequence learning. These data provide initial evidence that increased engagement during practice has a direct effect on generalized learning, improving retention and transfer of a complex motor skill.

Engaging environments enhance motor skill learning in a computer gaming task.

Goal-directed, task-specific, and repetitive physical practice is an important determinant of motor learning in animal studies (Kleim, Barbay, & Nudo, 1998; Nudo & Milliken, 1996), human motor skill learning (Williams & Hodges, 2012; Wulf, Shea, & Lewthwaite, 2010), and also in the reacquisition of motor skills during neurorehabilitation (Birkenmeier, Prager, & Lang, 2010; Waddell, Birkenmeier, Moore, Hornby, & Lang, 2014). Beyond the *quantity* of practice required for motor learning, there are important considerations about the *quality* of practice. Not all practice is equally efficacious for long term retention, or what has been referred to as "learning" (Schmidt & Lee, 2011; Soderstrom & Bjork, 2015). For example, distributed or spaced practice of one skill is better for retention than equal amounts of massed practice (e.g., Shea, Lai, Black, & Park, 2000) and random or interleaved practice of multiple skills is better for retention than equal amounts of blocked or repetitive practice (e.g., Kantak & Winstein, 2012). What is especially interesting about these examples above is that short term performance in practice does not positively predict which conditions will best aid long term learning. Hence, distinctions between performance (during practice) and learning (during a delayed testing phase) are needed to enable conclusions about latent variables in practice which might later impact retention and hence motor learning (for the most recent discussion of this distinction see Soderstrom & Bjork, 2015).

In recent years, practice quality and its implications for motor learning have been studied with respect to the affective experiences of the learner (Lewthwaite & Wulf, 2010; Sanli, Patterson, Bray, & Lee, 2013; Gabriele Wulf, Chiviacowsky, & Cardozo, 2014) and important theoretical questions have been raised about how a learner's *motivation* to practice as well as their *engagement* during practice might influence the motor learning process (Lohse, Shirzad, Verster, Hodges, & Van der Loos, 2013; Zimmerli, Jacky, Lünenburger, Riener, & Bolliger, 2013).

To clarify terms, we operationally define *motivation* as a psychological property that encourages action towards a goal by eliciting and/or sustaining goal directed behavior (see Mogenson, Jones, & Yim, 1980; Wise, 2004). *Engagement*, however, refers to the affective quality or experience of a person. Interactivity, choice, exploration, and reward are all environmental properties that are thought to contribute to the experience of engagement (Hunicke, Leblanc, & Zubek, 2004; O'Brien & Toms, 2008, 2010). Thus, engaging environments are likely to be motivating (i.e., you will return to an activity that was engaging), but motivation is not a guarantee of engagement (i.e., you may be motivated to play a game only to find it is no longer challenging, reducing engagement and potentially reducing future motivation).

Increased motivation and engagement during practice have the potential to increase the amount of practice in which a learner chooses to participate (Hunicke et al., 2004; O'Brien & Toms, 2008). This might be referred to as an indirect effect of engagement on learning through motivation that manifests as practice amount. However, there is also neurophysiological evidence to suggest that motivation and engagement can have direct effects on learning (not mediated by increased practice). From research in rodents, for instance, "enriched" environments (those containing complex inanimate and social stimulation) can increase the retention of new neurons (Kempermann, Kuhn, & Gage, 1997), the number of synapses per neuron (Anderson et al., 1994), and the expression of brain-derived neurotrophic factor (Klintsova, Dickson, Yoshida, & Greenough, 2004) compared to "sterile" environments in which similar amounts of repetitive exercise occur. Enriched environments in animal research show a strong correspondence to engaging environments as defined in human psychological literature: providing choices, novelty, complex physical interactions, and opportunities for exploration (Hunicke et al., 2004; Lohse et al., 2013; Zimmerli et al., 2013).

To our knowledge, there is no research on engagement-mediated learning effects in humans, but considerable research has been conducted on the role of motivation in learning. Monetary rewards, for instance, but not equivalent punishments, given during practice, improve retention of explicitly learned motor sequences (Abe et al., 2011) and memory for visually presented stimuli (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Wittmann et al., 2005; Wolosin, Zeithamova, & Preston, 2012). Similarly, participants' endogenous curiosity for certain topics (e.g., facts about dinosaurs) has been shown to modulate the strength of explicit memories (Gruber, Gelman, & Ranganath, 2014; Kang et al., 2009). Neuroimaging data suggest that memory benefits are attributable to interactions between dopaminergic midbrain structures (viz., substantia nigra and ventral tegmental area) and the hippocampus during the anticipation of reward (Adcock et al., 2006; Wolosin et al., 2012) or during periods of increased curiosity (Gruber et al., 2014). To date then, both rewards, which are extrinsic motivators, and curiosity, which is an intrinsic motivator, have been tested in human participants and found to enhance learning. For both types of motivation, neuroimaging data suggest that learning is modulated by the interaction of dopaminergic pathways and the hippocampus during encoding (i.e., during practice).

Combining ideas about engagement with neurophysiological data showing that activity in dopaminergic pathways during encoding facilitates subsequent retrieval, leads to the hypothesis that engaging environments can positively impact motor learning. Much has been written recently about the potential for video game technology to enhance motor learning and more specifically rehabilitation following disease or injury (e.g., Lohse, Hilderman, Cheung, Tatla, & Van der Loos, 2014; Lohse et al., 2013; Zimmerli et al., 2013). Theoretically, these technologies should help to enhance motivation to practice more by making the practice environment more engaging (i.e., indirect effects), but it is also possible that engaging, gaming environments could have immediate, direct effects on motor learning. Previous work in humans has led to suggestions that motivation also directly affects processes involved in long-term skill retention (Abe et al., 2011; Adcock et al., 2006; Wolosin et al., 2012), but human research on engagement and motor learning is currently lacking.

The goal of the present study was to explore the role of engagement in motor learning. However, it is important to consider that motor learning and memory are not singular processes, but are composed of separate abilities. The broad categories of learning and memory can be subdivided into 2 main types—explicit and implicit (Squire, 1987). Explicit knowledge is represented as memory for facts, events and episodes, and may be formed very quickly (even following one exposure to explicit information). Explicit knowledge is directly accessible to conscious recollection and is used to guide high-level cognition when decisions are based on complex rules and information. By contrast, the functions of the implicit system are highly distributed, supporting multiple behaviors, including skills and habits (e.g., sequence learning) (Squire, 1987). Our current study was designed to test whether game features, thought to promote engagement, would impact motor learning, and if so, whether implicit and explicit processes would be differentially affected.

Thus, we conducted a behavioral experiment comparing across two different groups that varied in the environmental conditions of their practice. Our aim was to see if practicing a motor skill in an engaging computer game (the "Game" condition) would facilitate motor learning compared to a mechanically identical version of the same game with aesthetic features removed (the "Sterile" condition). In both conditions, the game required participants to "catch" objects that flew onto the screen as quickly as possible and then "throw" these objects at targets on the screen in order to score points while we recorded accuracy and timing metrics of their performance. In the game, we also embedded a specific sequence of locations from which the objects would originate; at other times individuals practiced random sequences of objects. By comparing performance on random trials with repeating sequence trials we tested whether any motor learning that occurred was sequence-specific (correct anticipation of predictable objects) or generalized (equally improved for repeated and random sequences). Post-test surveys were also used to check if potential sequence learning was explicit or implicit.

We predicted that the game environment would provide a more engaging experience than a sterile condition, promoting better retention of skilled movements. Although rewards have been shown to affect implicitly acquired sequence learning (Gong & Li, 2012), we do not know whether engagement during practice will similarly affect implicit learning. To test these predictions we ran two successive studies where we varied the dose of practice (200 trials vs 400 trials). Because implicit sequence learning can take longer to develop than more explicitly acquired processes, we doubled the amount of practice in Experiment 2 to ensure that any potential effects were not hidden because of the potential lack of practice. We did not expect differences in the overall pattern of results across the two studies. Measures of engagement as well as intrinsic motivation were collected in order to inform about potential psychological mechanisms underpinning our practice manipulation. The game environment was expected to be rated as more engaging than the sterile environment and we anticipated that this would translate to increased enjoyment and potentially higher intrinsic motivation in general.

Methods

Participants

Forty participants were recruited through an online advertisement at the University of British Columbia (23F and 17M). The average age of the participants was 23.68 years (SD = 3.41). None of the participants regularly or currently played games using the Kinect. However, 37 participants endorsed that they played video games in other media (e.g., tablets, phones, or console systems; Game, n = 10; Sterile, n=12) and 22 reported that they had played a motion controlled game, with n=17 reporting experience with the Kinect (at least once) in their lifetime (Game, n=9; Sterile, n=8). All participants were blind to the hypotheses of the experiment.

Participants were successively assigned to a low dose of practice (200 trials on 1 day; Experiment 1) or a high dose of practice (400 trials over 2 days; Experiment 2), with a delayed retention/transfer test that occurred 5-9 days after the first day of practice (depending on the participant's availability). Although all the low-dose groups were run before the high-dose groups, participants across both studies were drawn from the same population, responded to the same advertisement and were run at the same time of year (summer term). Within each study, participants were pseudo-randomly assigned to either the Game group or the Sterile group, using blocked-randomization within sex to balance the groups. One participant in Experiment 1 (23 year old female in the Game group) did not return for retention testing although her data were included in acquisition analyses. An additional participant in this group (22 year old female) was a significant outlier, with in-game scores 2.82 SDs below the sample mean. Data for this latter participant were removed before analysis. Thus there were four groups: A Game-200 trials (6F, 3M during acquisition; 5F, 3M during retention/transfer) and Sterile-200 trials (6F, 4M) in Experiment 1 and a Game-400 trials (5F, 5M), and Sterile-400 trials (5F, 5M) in Experiment 2.

Apparatus and Measures

Participants played a custom built computer game using the Microsoft Kinect® (Microsoft, Redmond, WA) which was written in Visual Studio 2010 using XNA Game Studio 4.0 and the Kinect SDK v1.8 (Microsoft, Redmond, WA). The game was played on a 60 inch LG flat panel television with the Kinect sensor centered in front of the television approximately 1 m off the ground, 7.6 cm below the television, and 1.4-1.6 m away from the participant (the exact distance changed from participant to participant as they stepped forward/backward to improve motion tracking).

Procedures

Two experiments were run that differed with respect to the dose of practice and the details of the repeating sequence. In Experiment 1, each participant completed 200 practice trials on one practice day. Experiment 2 was almost identical except that participants completed two days of practice (200 trials/day). Within each Experiment, participants were randomly assigned to either the Game group or the Sterile group and completed all acquisition trials within the assigned condition. Delayed retention and transfer tests (each 20 trials long) were conducted 5-9 days after the last acquisition day. Participants completed the retention test in the same condition as practice first, but then switched to the opposite condition for the transfer test.

Two 5-location sequences were pseudo-randomly generated with the restriction that object locations did not repeat (e.g., AB but not AA) or trill (e.g., ABCB but not ABAB). The sequence referred to one of 8 locations around the edge of the screen where the object would originate (top-left, TL, –middle, T or –right, TR; bottom-left, BL, -middle, B, or –right, BR; centre-left, L, or –right, R). Sequence A (L, TR, R, TR, BR) was used for the 200-trial group and Sequence B (R, TL, BL, B, TR) was used for the 400-trial group. Different sequences were used to reduce potential item effects that could confound sequence learning effects. There was no evidence that the sequences were acquired differently across the two experiments, suggesting the sequences had comparable difficulty.

In the game/sterile conditions, participants used their non-dominant arm (in order to reduce potential transfer from previous gaming experiences) to control the motion of a spaceship/cursor on the screen in order to “catch” asteroids/circles that appeared on the screen from various pre-determined locations, and then “throw” the objects at yellow targets (see Figure 1). In both conditions, the asteroid/circle flashed red when it was caught to indicate a successful catch. The game condition included sound, both background music and action-specific sounds, whereas there was no sound in the sterile condition. In both conditions, a participants’ in-game score was shown in the top-left corner of the screen. As explained to participants before the experiment, their score was based on how quickly they could catch the asteroid and how many targets they could hit. Each trial was worth a maximum of 100 points.

Participants lost a single point per 10 frames (~167 ms) until they caught the object, but they scored 100 points for every target that they hit.

In retention and transfer, the scores remained on the screen and hence the conditions in retention (and transfer) replicated those experienced during practice. These sessions were divided into alternating blocks of 5 random trials and 5 sequence trials.

Dependent measures included a participants' in-game score, whether or not an object was caught, the time-to-catch (or total-time if no catch was made), whether or not the target was hit, and the total-time for the trial (time from appearance of the object to the object hitting the target or leaving the screen). These were calculated for each trial. Across a block of 5 trials, data were condensed to: points scored per block (500 points maximum), proportion of objects caught (out of five), the average time-to-catch for successful catches, the proportion of targets hit for successful catches, and the average total-time of the trial on successful catches.

Survey Measures

Before Acquisition 1, all participants completed a pre-training demographic survey that assessed age, sex, previous experience with video-games, previous experience with motion controlled games and any neurological/visual/musculoskeletal impairments that may have influenced their ability to perform the task. Following practice in Acquisition 1, all participants completed a post-training questionnaire that contained a language-modified version of a user engagement scale that was specific to our task (O'Brien & Toms, 2010). We also included a language-modified version of the intrinsic motivation inventory, IMI (McAuley, Duncan, & Tammen, 1989), which again was modified to be specific to our task and that contained interest/enjoyment, perceived competence, effort, and pressure/tension subscales. We expected the interest/enjoyment subscale to be most related to our game-based manipulation, with enhanced interest being a consequence of an aesthetically engaging environment, but we maintained all components of the IMI to explore any general intrinsic motivation benefits from game-based practice.

At the end of the delayed testing session, participants also completed an exit-survey to assess explicit awareness of the embedded sequence. The survey consisted of four questions asking participants if they noticed "anything unusual" while playing the game to specifically asking participants if they noticed a sequence. If participants indicated that they noticed a sequence, they also completed a recognition test in which three sequences were shown on the screen. Participants then had to indicate which, if any, sequence was the one they practiced.

Statistical Analyses

This study was designed to detect two primary effects with a priori statistical power of $(1-\beta) = 0.80$. First, we were interested in an overall learning effect, operationally defined as the main effect of training group (Game versus Sterile) on the delayed test. Assuming $\alpha = 0.05$, and Cohen's $d = 0.8$ (a large effect), a total sample size of N = 40 was needed to achieve 80% power (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). The primary outcome for the overall learning effect was participants' in-game score.

Second, we wanted adequate statistical power to detect an implicit learning effect from the repeating asteroid sequence. Specifically, we were interested in potential differences between groups, thus we operationally defined the implicit motor learning effect as the interaction between training group (Game versus Sterile) and Trial-type (Random versus Sequence) on the delayed test. Assuming $\alpha = 0.05$, $\eta_p^2 = 0.05$ (a small effect), and a strong correlation between Random and Sequence Trials (r = 0.50), a total sample size of N = 40 was again needed to achieve 80% power.

This study was designed to detect hypothesized differences on the delayed retention/transfer tests, thus any differences during acquisition or on survey measures should be treated as exploratory. To test for differences on the delayed tests, a 4-way, mixed design ANOVA, with between-subjects factors of group (Game versus Sterile) and dose (200 trials versus 400 trials) and repeated-measures on the last two factors of trial-type (Random versus Sequence) and test-type (Retention versus Transfer) was used. These analyses were conducted on the primary outcome of in-game score, and the secondary outcomes of % asteroids/circles caught, time-to-catch, % target hits and total time/trial.

For acquisition data, a 4-way, mixed design ANOVA with between-subjects factors of group (Game versus Sterile) and dose (200 trials versus 400 trials) and repeated-measures on the last two factors of trial-type (Random versus sequence) and block (Blocks 1-10) was conducted for all outcome measures on Day 1. For participants in the high dose group, an additional analysis was conducted comparing Day 1 to Day 2 and the interaction of Day with Block.

Survey data from the engagement scale and the IMI were analyzed using independent samples t-tests comparing the Game training group to the Sterile training group. These tests were conducted for the composite score of each survey and the sub-scales of each survey.

Maulchy's test of sphericity was conducted prior to all analyses. The Greenhouse-Geisser correction was applied to any tests where sphericity was violated. Analyses were conducted using IBM SPSS v22.0.

Results

Retention and Transfer Tests

Training in an engaging environment improves learning.

For in-game scores on the delayed retention (same game context as practice) and transfer tests (opposite context), participants showed similar levels of performance on both tests and there was no statistically significant effect of test-type ($F < 1$). Importantly, there were significant main effects of group, $F(1,34) = 8.19$, $p = .007$, $\eta_p^2 = 0.19$, and dose, $F(1,34) = 13.10$, $p = .001$, $\eta_p^2 = .28$, but they did not interact ($F<1$). Data for all dependent measures are shown in Table 1. For points per block, shown in Figure 2, participants in the Game group scored more points on retention and transfer tests than participants in the Sterile group. Similarly, participants who had 400 practice trials scored more points on retention and transfer tests than participants who had 200 practice trials. The only interaction was for Dose by Test-type, $F(1,34) = 6.64$, $p = .01$, $\eta_p^2 = .16$. The low dose group scored more points on the retention test (397.79 points) than the transfer test (371.65 points), pointing to a specificity of practice effect for low practice amounts. Although the high dose group scored more points in general, more points were scored on the transfer test (429.37 points) than on the retention test (410.95 points). As detailed in Table 1, the group effects were replicated across all dependent measures with the exception of time-to-catch. The dose effect was also seen across most measures, except for % caught (p=.07) and total time.

Weak evidence of implicit sequence learning, independent of game group.

For in-game scores on the delayed retention and transfer tests, there was a significant main effect of trial-type, $F(1,34) = 6.89$, $p = .01$, $\eta_p^2 = .17$, such that participants scored more points on Sequence trials (411.86 points) than on Random trials (393.02 points). Importantly, Trial-type did not significantly interact with Training group (F<1). A priori, the study had 80% statistical power to detect an interaction with an expected correlation of r = .50 between random and sequential trials. The actual correlation was r(36) = .48. As such, although we had reasonable power to detect an effect, there was no evidence that the game environment augmented implicit learning of the sequence as opposed to a general learning effect.

The implicit learning effect itself was not particularly strong. Although there was an effect of trial-type for in-game score, this effect was not significant for any of the other outcome variables, shown in Table 2. The effect on in-game score appeared to be driven by percentage of targets hit.

Exit survey data suggested that participants had no explicit awareness of the embedded, repeated sequence. Although five participants indicated they thought there was a sequence present, none of these participants were able to accurately recall all or part of the sequence and only one of these participants correctly identified their sequence in the recognition test.

Practice data

Self-reported engagement was higher for the Game group, but not motivation.

As shown in Figure 3a, for the Engagement Scale, there was a significant difference between groups for engagement overall, $t(38) = 2.15$, $p = .04$. Although the mean ratings for the game group were numerically higher on all of the subscales, the only subscale on which ratings were significantly different was the aesthetics subscale, $t(38) = 4.47$, $p < .001$.

For the IMI (Figure 3b), there was no difference between groups for overall intrinsic motivation, $t(38) = 0.31$, $p = .76$. The only significant difference between groups was on the effort subscale, $t(38) = -2.43$, $p = .02$, with the Game group reporting less effort than the Sterile group.

There was also a significant positive correlation between overall engagement scale scores and overall motivation scale scores $r(38) = .66$, suggesting that engagement and motivation are different, but related constructs. Neither engagement, $r(36) = .12$, nor intrinsic motivation, $r(36) = .18$, correlated significantly with points scored per block on the delayed test.

All groups improved during practice

For in-game scores during Day 1, improvements were evidenced by a significant main effect of block, $F(9,306) = 4.89$, $p < .001$, $\eta_p^2 = .12$ (see Figure 4). There was also a main effect of trial-type, $F(1,34) = 4.12$, $p = .05$, $\eta_p^2 = .11$, but this was superseded by a Trial-type by Block interaction, $F(9,306) = 2.34$, $p = .02$, $\eta_p^2 = .06$. From inspection of Figure 4, this interaction appeared to be driven by the fact that participants scored more points on random trials early in practice (Blocks 2 and 3), but as expected, scored more points on sequence trials later in practice (Blocks 6, 7, 8, and 10). The group effect was not statistically significant, $F(1,34) = 3.30$, $p = .08$, $\eta_p^2 = .08$, but there was a main-effect of dose, $F(1,34) = 6.45$, $p = .02$, $\eta_p^2 = 0.15$, with participants in the high-dose group scoring more points (386 points) than participants in the low-dose group (359 points). This effect appears to be attributable to more proficient participants being assigned to the high-dose group and not a difference in the difficulty of the two sequences, as the difference was found for both random and sequential trials and there was no Dose by Trial-type interaction, $F(1,34) = 1.62$, $p = .21$, $\eta_p^2 = .04$.

For the high dose participants only, we conducted an additional analysis comparing Day 1 to Day 2. As above, there were no group related effects. There were significant effects of day,

F(1,17) = 26.75, $p < .001$, $\eta_p^2 = .61$, block, F(9,153) = 3.62, $p < 0.001$, $\eta_p^2 = .18$, and a Day by Block interaction, F(9,153) = 2.04, $p = .03$, $\eta_p^2 = .11$. Although participants improved from Day 1 (mean = 386.79 points) to Day 2 (416.97 points), within-day improvements were larger on Day 1 than on Day 2.

As in-game score was a composite of % caught, time-to-catch, and targets hit, we also ran secondary analyses on these data, which were again based only on the high-dose participants to make comparisons across days, see Figure 5a-d. For % caught (Figure 5a), there were no significant group-related effects. Participants improved across days, F(1,17) = 49.71, $p < .001$, $\eta_p^2 = 0.75$, and blocks, F(9,153) = 5.46, $p = .001$, $\eta_p^2 = 0.24$, and there was a Day by Block interaction, F(4.96,84.38) = 2.77, $p = .02$, $\eta_p^2 = 0.14$. As predicted, participants caught a higher percentage on Sequence trials (96%) than Random trials (94%), shown by a main effect of trial type, F(1,17) = 6.56, $p = .02$, $\eta_p^2 = .28$.

For time-to-catch (Figure 5b), only significant main-effects of day, F(1,17) = 48.98, $p < .001$, $\eta_p^2 = .75$, and block, F(9,153) = 5.60, $p = .001$, $\eta_p^2 = .25$ were observed. There were no group or trial-related effects, nor any statistically significant interactions.

For % of targets hit (Figure 5c), although there looked to be group differences, there were no significant effects involving group. Participants did improve in % of target hits across days, F(1,17) = 9.02, $p < .01$, $\eta_p^2 = .35$, but not significantly across blocks, F(9,153) = 1.48, $p = .16$, $\eta_p^2 = .08$. There was no Day by Block interaction and no main-effect of trial type (both Fs<1).

For total-time (Figure 5d), there were significant main-effects of day, F(1,17) = 15.22, $p < .001$, $\eta_p^2 = .47$, block, F(4.54,77.10) = 5.45, $p < .001$, $\eta_p^2 = .24$ and trial-type, F(1,17) = 7.61, $p =$

.01, η_p^2 = .31, but no group-related effects. Participants got faster across days and blocks, but they were slower on random trials (3,406 ms) than sequence trials (3,333 ms).

Discussion

Our data provide evidence that training in an engaging, game environment improves the learning of a novel motor skill compared to an equal amount of mechanically similar training in a less engaging environment. Significant effects of the training condition on the retention and transfer tests suggest that the more engaging “Game” conditions improved the retention of motor skill learning. Retention, specifically, is implicated because the Game and Sterile groups showed similar performance curves during acquisition, but differences between groups emerged after the delay between acquisition and retention testing.

There is considerable evidence that physiological processes enhance learning after physical practice is finished, what is known as memory consolidation (Krakauer & Shadmehr, 2006; McGaugh, 2000; Robertson, Pascual-Leonne, & Miall, 2004). Our data suggest that psychological states during practice may alter this process of consolidation. Specifically, while keeping rewards and mechanics constant, the more engaging aesthetics of the Game condition appeared to facilitate the consolidation of this complex motor skill as evidenced by improved retention. A lack of difference in the IMI data suggest that this was not a result of a general enhancement in intrinsic motivation. Although enjoyment was generally rated as higher in the Game group than in the Sterile group, the groups only differed significantly on the effort subscale of the IMI, with the Game group rating their practice as less effortful (i.e., performed with more ease) than that of the Sterile group. However, there was a statistically significant difference in overall levels of engagement between the two groups, with the largest difference being participants' preference for the aesthetics of the game condition. Moreover, engagement was positively correlated with overall IMI ($r = .66$). While aesthetics contribute to engagement, other factors such as choice/interactivity, clear goals and mechanics, and optimal levels of difficulty are known to contribute to engagement as well (Lohse et al., 2013). Some of these factors have been explored in previous motor learning studies (e.g., Guadagnoli & Lee, 2004; Wulf & Adams, 2014), but to our knowledge the current experiment is the first demonstration that increased engagement through aesthetic, sensory features of the task can facilitate learning.

We are currently researching the neurophysiological differences between Game and Sterile conditions during practice, and how these proximal differences in responses to stimuli might affect retention. Candidate mechanisms would be dopaminergic systems that have been implicated in both rodent and human research. For instance, post-training injections of dopamine receptor agonists enhance learning for stimuli specific to the area of injection (Hitchcott & Phillips, 1998; Packard, Cahill, & McGaugh, 1994). Neuroimaging studies in humans similarly implicate dopaminergic midbrain structures (viz., substantia nigra and ventral tegmental area) (Adcock et al., 2006; Gruber et al., 2014; Wolosin et al., 2012) in motivation-mediated learning effects. Thus, these substrates could also explain engagement-mediated learning effects.

The learning advantage of the game group was found for both random and sequential trials. This result suggests that the engaging environment benefited both learning of specific sequences of action as well as generalized motor control. The lack of a sequence specific motor learning result could be attributable to the fact that there was very little evidence of a sequence learning effect overall (see Table 2), or it could be reflective of the differential impact of engagement on motor learning memory processes. Without randomly assigning participants to conditions of explicit or implicit knowledge of the sequence, it is not possible to make any definite conclusions with respect to engagement and implicit learning.

There was also a benefit to receiving more practice in the task, shown by the effect of dose on the delayed retention and transfer test. This effect must be interpreted with caution, however, because the 400-trial group started to show a difference from the 200-trial group during acquisition Day 1. Importantly, this difference did not appear to be driven by a difference in the difficulty of Sequence A (given to the 200-trial group) and Sequence B (given to the 400-trial group), as the 400-trial group showed an advantage for both random and sequential trials.

One limitation of the present study is that the Sterile condition, although less engaging than the Game condition, was still quite engaging according to the survey data (both groups rated the experience above the median score of 4). Although the aesthetic features of the game were removed in the Sterile condition, participants were still playing a game that provided considerable challenge and interactivity. These results, and the sequence learning results in particular, might therefore be stronger with a less engaging control condition.

Furthermore, although the sequence learning effect was quite weak, it may have been affected by the amount of practice that participants received and the implicit nature of the sequence. Within our own data, there was little evidence that additional practice trials made the sequence learning effect stronger (beyond within day improvements for the sequence trials). It is possible that 400 trials of practice (30 exposures to the 5-item sequence) were not enough to elicit a sequence learning effect, given that increased exposures have been the norm in other serial response paradigms, although admittedly with longer sequences (e.g., 80 exposures to a six item sequence, Curran & Keele, 1993; 80 exposures to a 12 item sequence, Boyd, Vidoni & Siengsukon, 2008). Making the sequence explicit at the beginning of the experiment may make the sequence easier to learn (Boyd & Winstein, 2001) and perhaps more susceptible to modulation through engagement. However, as above, it is possible that such a manipulation would change the pattern of results because of the change to explicit rather than implicit learning conditions. Previous research demonstrating effects of motivation on learning was based on the acquisition of an explicit motor sequence (Abe et al., 2011) or explicit knowledge (Adcock et al., 2006; Wittmann et al., 2005). These studies, however, manipulated rewards between groups to influence motivation, in contrast to our experiment in which rewards were kept constant.

In conclusion, there is evidence that extrinsic motivators related to the learning environment directly impact motor learning (Abe et al., 2011), but in the present study we sought to manipulate engagement with the task by creating a more stimulating environment (Lohse et al., 2013; Zimmerli et al., 2013). Intrinsic motivation did not differ between the game and sterile group, when assessed using the Intrinsic Motivation Inventory. Only the engagement scale reflected training related differences across the groups. Moreover, differences in performance between the groups were only seen after a 1-week retention interval and not during practice, suggesting that the impact of this variable happened sometime between the end of practice and retention, reflective of enhanced consolidation (similar to Abe et al., 2011). Improvements for the Game group were seen for both random and predictable (repeating sequence) trials, suggesting a general motor-related improvement, rather than an implicit learning effect. This is the first study to evaluate whether game-like contexts, when learning a novel motor task, impacts retention of the skill. While previous data have shown that engagement can have an indirect effect on learning by increasing the amount of self-selected practice, the current data show that engagement can also have a direct effect on learning, improving the consolidation of a complex motor skill.

Funding sources: This work was supported by the Natural Sciences and Engineering Research Council (Discovery grants awarded to Boyd and Hodges).

Boyd receives salary support from the Canada Research Chairs and the Michael Smith Foundation for Health Research. Hodges receives salary support from the Canadian Institute for Health Research

Disclosures: None.

Acknowledgements: The authors would like to thank Kris De Asis and Bryan Luu for their assistance in programming the game, and Beverly Larssen for help with participant testing.

References

Abe, M., Schambra, H., Wassermann, E. M., Luckenbaugh, D. A., Schweighofer, N., & Cohen, L. G. (2011). Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current Biology*, *21*, 557–562.

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*, *50*(3), 507–17. doi:10.1016/j.neuron.2006.03.036

Anderson, B. J., Li, X., Alcantara, A. A., Isaacs, K. R., Black, J. E., & Greenough, W. T. (1994). Glial hypertrophy is associated with synaptogenesis following motor-skill learning, but not with angiogenesis following exercise. *Glia*, *11*(1), 73–80. doi:10.1002/glia.440110110

Birkenmeier, R. L., Prager, E. M., & Lang, C. E. (2010). Translating animal doses of task-specific training to people with chronic stroke in 1-hour therapy sessions: a proof-of-concept study. *Neurorehabilitation and Neural Repair*, *24*(7), 620–35. doi:10.1177/1545968310361957

Boyd, L. A., Vidoni, E. D., & Siengsukon, C. F. (2008). Multidimensional motor sequence learning is impaired in older but not younger or middle-aged adults. *Physical Therapy*, *88*(3), 351–62. doi:10.2522/ptj.20070131

Boyd, L. A., & Winstein, C. J. (2001). Implicit motor-sequence learning in humans following unilateral stroke: the impact of practice and explicit knowledge. *Neuroscience Letters*, *298*(1), 65–69. doi:10.1016/S0304-3940(00)01734-1

Curran, T., & Keele, S. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*(1), 189–202.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60. doi:10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–91.

Gong, M., & Li, S. (2012). Reward Prompts Visual Short-Term Memory Consolidation. *Journal of Vision*, *12*(9), 171.

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. *Neuron*, 1–11. doi:10.1016/j.neuron.2014.08.060

Guadagnoli, M. a, & Lee, T. D. (2004). Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, *36*(2), 212–24. doi:10.3200/JMBR.36.2.212-224

Hitchcott, P., & Phillips, G. (1998). Double dissociation of the behavioural effects of R (+) 7–OH–DPAT infusions in the central and basolateral amygdala nuclei upon Pavlovian and instrumental. *Psychopharmacology*, *140*(4), 458–469.

Hunicke, R., Leblanc, M., & Zubek, R. (2004). MDA : A Formal Approach to Game Design and Game Research. *Challenges in Game Artificial Intelligence: Papers from the 2004 AAAI Workshop*, 1–5.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963–73. doi:10.1111/j.1467-9280.2009.02402.x

Kantak, S. S., & Winstein, C. J. (2012). Learning – performance distinction and memory processes for motor skills : A focused review and perspective. *Behavioural Brain Research*, *228*(1), 219–231. doi:10.1016/j.bbr.2011.11.028

Kempermann, G., Kuhn, H. G., & Gage, F. H. (1997). More hippocampal neurons in adult mice living in an enriched environment. *Nature*, *368*, 493–495.

Kleim, J. A., Barbay, S., & Nudo, R. J. (1998). Functional Reorganization of the Rat Motor Cortex Following Motor Skill Learning. *J Neurophysiol*, *80*(6), 3321–3325.

Klintsova, A. Y., Dickson, E., Yoshida, R., & Greenough, W. T. (2004). Altered expression of BDNF and its high-affinity receptor TrkB in response to complex motor learning and moderate exercise. *Brain Research*, *1028*(1), 92–104. doi:10.1016/j.brainres.2004.09.003

Krakauer, J. W., & Shadmehr, R. (2006). Consolidation of motor memory. *Trends in Neurosciences*, *29*(1), 58–64. doi:10.1016/j.tins.2005.10.003

Lewthwaite, R., & Wulf, G. (2010). Grand challenge for movement science and sport psychology: embracing the social-cognitive-affective-motor nature of motor behavior. *Frontiers in Psychology*, *1*(August), 42. doi:10.3389/fpsyg.2010.00042

Lohse, K. R., Hilderman, C. G. E., Cheung, K. L., Tatla, S., & Van der Loos, H. F. M. (2014). Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PloS One*, *9*(3), e93318. doi:10.1371/journal.pone.0093318

Lohse, K. R., Shirzad, N., Verster, A., Hodges, N. J., & Van der Loos, H. F. M. (2013). Videogames and rehabilitation: Using design principles to enhance patient engagement. *Journal of Neurologic Physical Therapy*, *37*(4), 166–175. doi:doi: 10.1097/NPT.0000000000000017

McAuley, E., Duncan, T., & Tammen, V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58.

McGaugh, J. (2000). Memory--a century of consolidation. *Science*, *287*, 248–251. doi:DOI: 10.1126/science.287.5451.248

Mogenson, G. J., Jones, D. L., & Yim, C. H. I. Y. I. U. (1980). From motivation to action: functional interface between the limbc and motor system. *Progress in Neurobiology*, *14*, 69–97.

Nudo, R., & Milliken, G. (1996). Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *Journal of Neuroscience*, *16*(2), 785–807.

O'Brien, H. L., & Toms, E. G. (2008). What is User Engagement ? A Conceptual Framework for Defining User Engagement with Technology. *Journal of the American Society for Information Science & Technology*, *59*(6), 938–955.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement in e-commerce environments. *Journal of the American Society for Information Science & Technology*, *61*(20 10), 50–69. doi:10.1002/asi.21229.1

Packard, M., Cahill, L., & McGaugh, J. L. (1994). Amygdala modulation of hippocampal-dependent and caudate nucleus-dependent memory processes. *Proceedings of the National Academy of Science*, *91*(18), 8477–8481.

Robertson, E., Pascual-Leonne, A., & Miall, R. C. (2004). Current concepts in procedural consolidation. *Nature Reviews Neuroscience…*, *5*, 576–582.

Sanli, E., Patterson, J., Bray, S., & Lee, T. (2013). Understanding self-controlled motor learning protocols through the self-determination theory. *Frontiers in Psychology*, *3*, 611. doi:doi: 10.3389/fpsyg.2012.00611

Schmidt, R., & Lee, T. (2011). *Motor control and learning* (5th ed.). Champaign, IL: Human Kinetics.

Shea, C. H., Lai, Q., Black, C., & Park, J.-H. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human Movement Science*, *19*(5), 737–760. doi:10.1016/S0167-9457(00)00021-X

Soderstrom, N., & Bjork, R. (2015). Learning Versus Performance An Integrative Review. *Perspectives on Psychological Science*, *10*, 176–199. doi:10.1177/1745691615569000

Squire, L. (1987). *Memory and brain.* New York, New York, U.S.A.: Oxford University Press.

Waddell, K. J., Birkenmeier, R. L., Moore, J. L., Hornby, T. G., & Lang, C. E. (2014). Feasibility of high-repetition, task-specific training for individuals with upper-extremity paresis. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association*, *68*(4), 444–53. doi:10.5014/ajot.2014.011619

Williams, A. M., & Hodges, N. J. (Eds.). (2012). *Skill Acquisition in Sport: Research, Theory and Practice* (2nd ed.). London: Routledge.

Wise, R. a. (2004). Dopamine, learning and motivation. *Nature Reviews. Neuroscience*, *5*(6), 483–94. doi:10.1038/nrn1406

Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, *45*(3), 459–67. doi:10.1016/j.neuron.2005.01.010

Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of Cognitive Neuroscience*, *24*(7), 1532–47. doi:10.1162/jocn_a_00237

Wulf, G., & Adams, N. (2014). Small choices can enhance balance learning. *Human Movement Science*, *38*, 235–40. doi:10.1016/j.humov.2014.10.007

Wulf, G., Chiviacowsky, S., & Cardozo, P. L. (2014). Additive benefits of autonomy support and enhanced expectancies for motor learning. *Human Movement Science*, *37*, 12–20. doi:10.1016/j.humov.2014.06.004

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: a review of influential factors. *Medical Education*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2923.2009.03421.x/full

Zimmerli, L., Jacky, M., Lünenburger, L., Riener, R., & Bolliger, M. (2013). Increasing patient engagement during virtual reality-based motor rehabilitation. *Archives of Physical Medicine and Rehabilitation*, *94*(9), 1737–46. doi:10.1016/j.apmr.2013.01.029

Table 1. Effects of training group and dose on all outcome measures at test.

	Game Group		Sterile Group		ME of Group		ME of Dose	
Measure	200 Trials (n = 8)	400 Trials (n = 10)	200 Trials (n = 10)	400 Trials (n = 10)	*F*	*p*	*F*	*P*
In-game Score	400.31 (27.02)	432.58 (22.44)	369.13 (31.32)	407.73 (36.83)	8.19	0.007	13.10	0.001
% Caught	97.20 (2.09)	98.50 (2.11)	94.0 (4.89)	96.75 (3.34)	5.05	0.031	3.42	0.073
Time-to-catch (ms)	1,447.93 (95.83)	1,289.03 (101.64)	1,445.84 (148.90)	1,367.22 (128.22)	0.92	0.345	8.94	0.005
% Hit	91.30 (3.89)	95.60 (3.44)	87.90 (6.43)	92.60 (5.61)	3.89	0.057	7.41	0.010
Total Time (ms)	3,210.34 (327.44)	3,000.26 (186.56)	3,360.92 (283.42)	3,327.36 (338.54)	6.48	0.016	1.68	0.203

Note. Cells show the mean (and SD) between participants. ME = main effect. The *F*- and *p*-values reported are from a mixed-factorial ANOVA that included test-type and trial-type as repeated measures, and training group and dose as between-subjects factors. All *F*-values had (1,34) degrees of freedom.

Table 2. Implicit learning effects (Random – Sequence trials) as a function of training group and dose.

	Game Group		Sterile Group		ME of Trial-Type		Trial-Type X Group Interaction	
Difference (RAND – SEQ)	200 Trials (n = 8)	400 Trials (n = 10)	200 Trials (n = 10)	400 Trials (n = 10)	*F*	*p*	*F*	*P*
In-game Score	-25.44 (42.66)	-6.43 (36.75)	-36.74 (57.60)	-6.76 (35.28)	6.89	0.013	0.16	0.688
% Caught	0.60 (7.30)	-1.00 (2.10)	-2.00 (4.20)	-1.50 (5.80)	1.37	0.249	0.89	0.351
Time-to-catch (ms)	14.77 (196.93)	-35.72 (103.76)	58.00 (102.75)	-24.78 (84.05)	0.23	0.881	0.45	0.509
% Hit	-5.78 (8.56)	-0.50 (8.25)	-4.54 (12.82)	-0.38 (5.07)	3.54	0.069	0.05	0.82
Total Time (ms)	160.58 (254.00)	145.29 (347.13)	93.09 (277.97)	-100.81 (345.28)	2.15	0.152	2.38	0.132

Note. Cells show the mean difference between Random and Sequence trials (and between subjects SD of the difference). Negative values for accuracy measures, but positive values for time measures, are suggestive of an implicit learning effect. ME = main effect. The *F*- and *p*-values reported are from a mixed factorial ANOVA that included test-type and trial-type as repeated measures, and training group and dose as between-subjects factors. All *F*-values had (1,34) degrees of freedom.

Figure Captions

Figure 1. Screenshots from separate trials in the Game condition (top-left) and the Sterile condition (top-right), and a schematic showing the timeline of the experiment (bottom). Participants in Experiment 1 (200-trials) completed Acquisition 1. Participants in Experiment 2 (400-trials) completed Acquisition 1 and 2. Survey 1 = pre-training demographic survey. Survey 2 = post-training engagement survey and intrinsic motivation inventory. Survey 3 = post-test sequence recognition survey.

Figure 2. Points scored per block on the delayed retention/transfer tests as a function of training group and dose of practice. As there was no main effect of test-type, data are presented averaging across retention and transfer tests. Error-bars show the between-subjects standard-error.

Figure 3a,b. Likert scale ratings for the O'Brien and Toms Engagement scale (a) and the Intrinsic Motivation Inventory (b) as a function of the composite (overall) score on each survey and the contributing subscales. Error-bars show the between-subjects standard-error.

Figure 4. Acquisition data for participants' in-game scores (shown as points per block, 500 max). Data for the Game training group are shown by white circles, the Sterile training group by black squares. Means shown for Acquisition 1 include data from the low- and high-dose groups. Means shown for Acquisition 2 include data only from the high-dose group (as the low-dose group did not have a second acquisition).

Figure 5. Acquisition data for the percentage of asteroids caught (a), time-to-catch (b), percentage of targets hit (c), and total trial time (d). Data for the Game training group are shown by white circles, the Sterile training group by black squares. Means shown for Acquisition 1 include data from the low- and high-dose groups. Means shown for Acquisition 2 include data only from the high-dose group (as the low-dose group did not have a second acquisition).

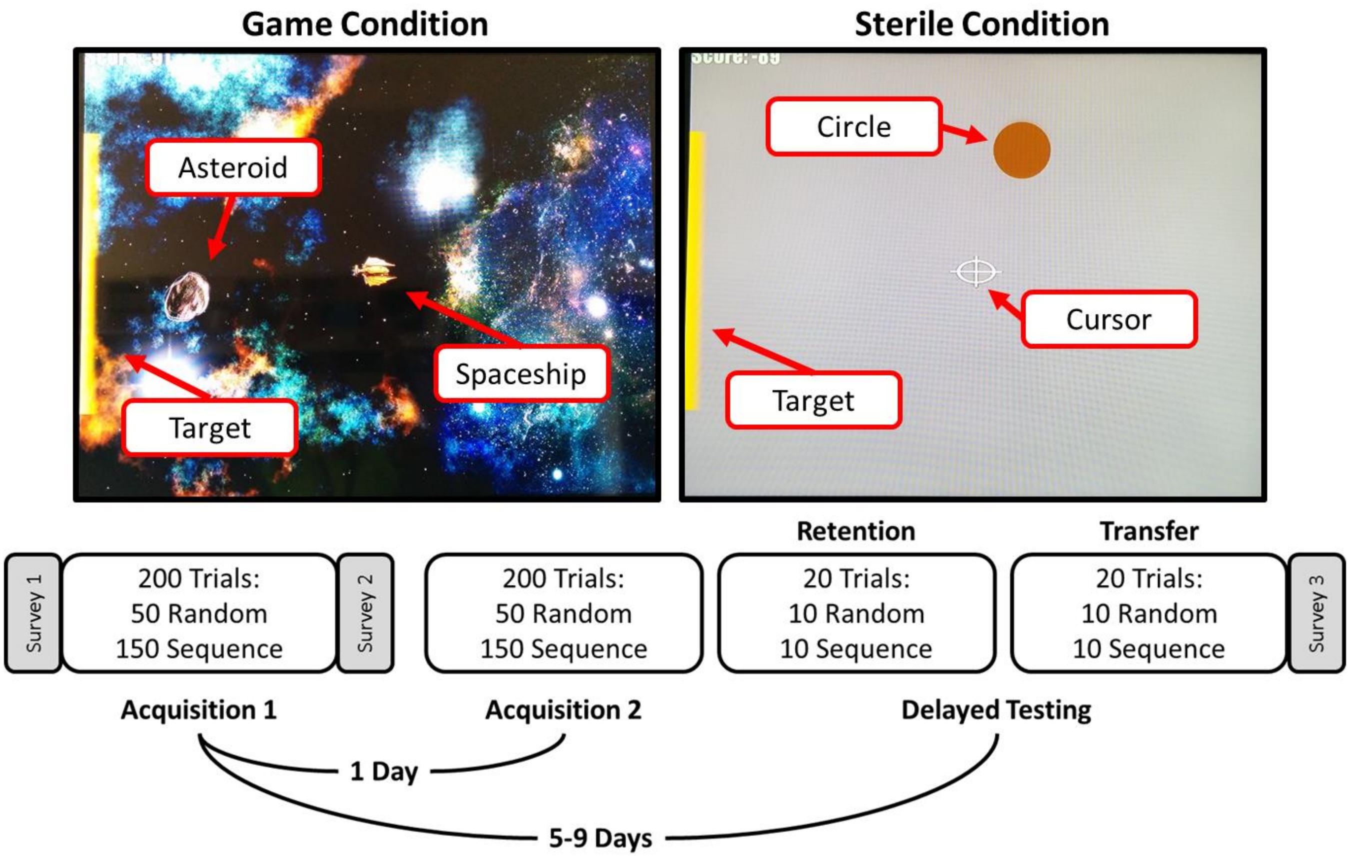
Game Condition
Sterile Condition
Asteroid
Spaceship
Target
Circle
Cursor
Target
Retention
Transfer
Survey 1
200 Trials:
50 Random
150 Sequence
Survey 2
200 Trials:
50 Random
150 Sequence
20 Trials:
10 Random
10 Sequence
20 Trials:
10 Random
10 Sequence
Survey 3
Acquisition 1
Acquisition 2
Delayed Testing
1 Day
5-9 Days

Figure 1.

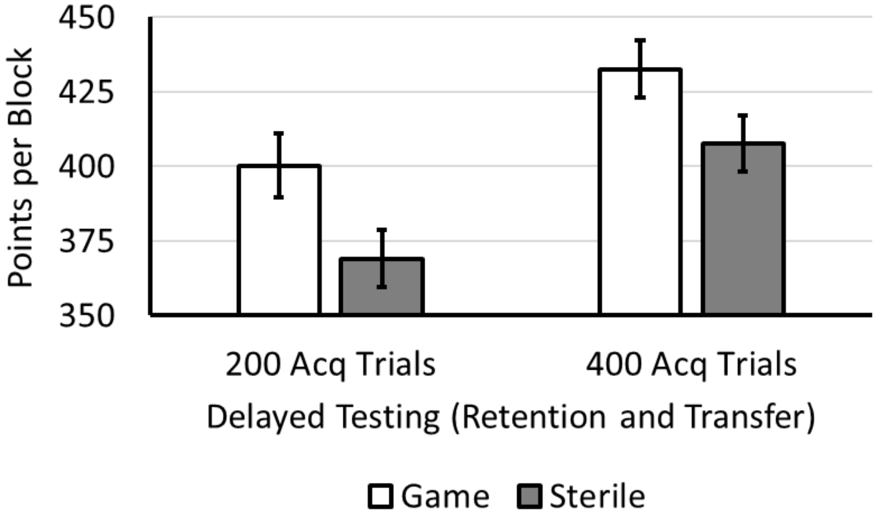
450
425
400
375
350
Points per Block
200 Acq Trials
400 Acq Trials
Delayed Testing (Retention and Transfer)
Game
Sterile

Figure 2.

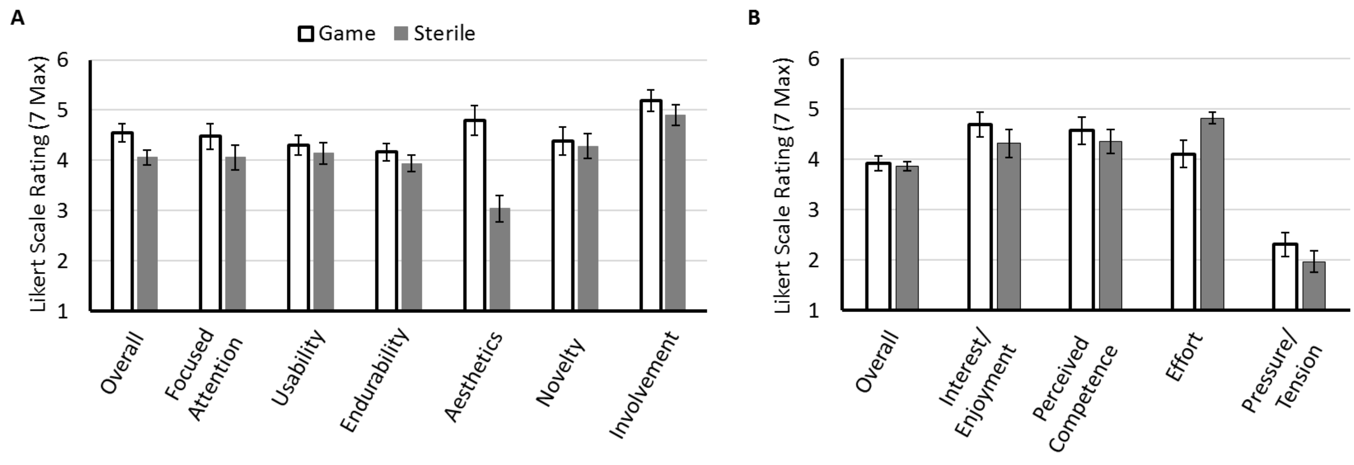
A
Game
Sterile
Likert Scale Rating (7 Max)
6
5
4
3
2
1
Overall
Focused Attention
Usability
Endurability
Aesthetics
Novelty
Involvement
B
Likert Scale Rating (7 Max)
6
5
4
3
2
1
Overall
Interest/ Enjoyment
Perceived Competence
Effort
Pressure/ Tension

Figure 3a,b.

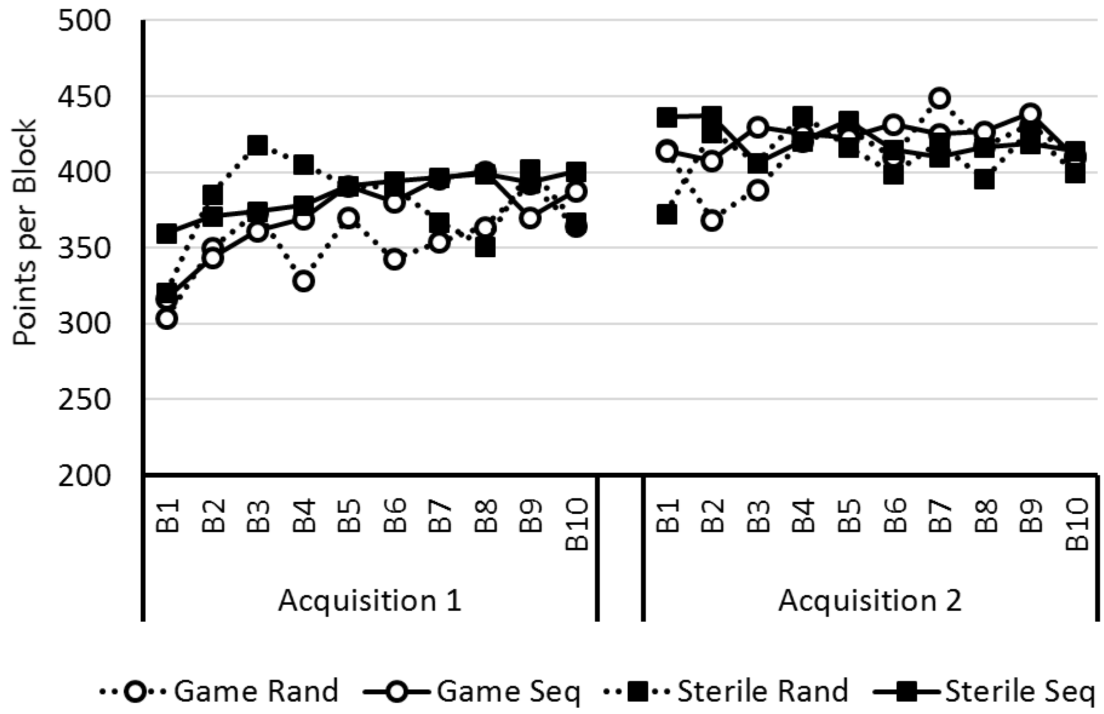
500
450
400
350
300
250
200
Points per Block
B1
B2
B3
B4
B5
B6
B7
B8
B9
B10
Acquisition 1
Acquisition 2
Game Rand
Game Seq
Sterile Rand
Sterile Seq

Figure 4.

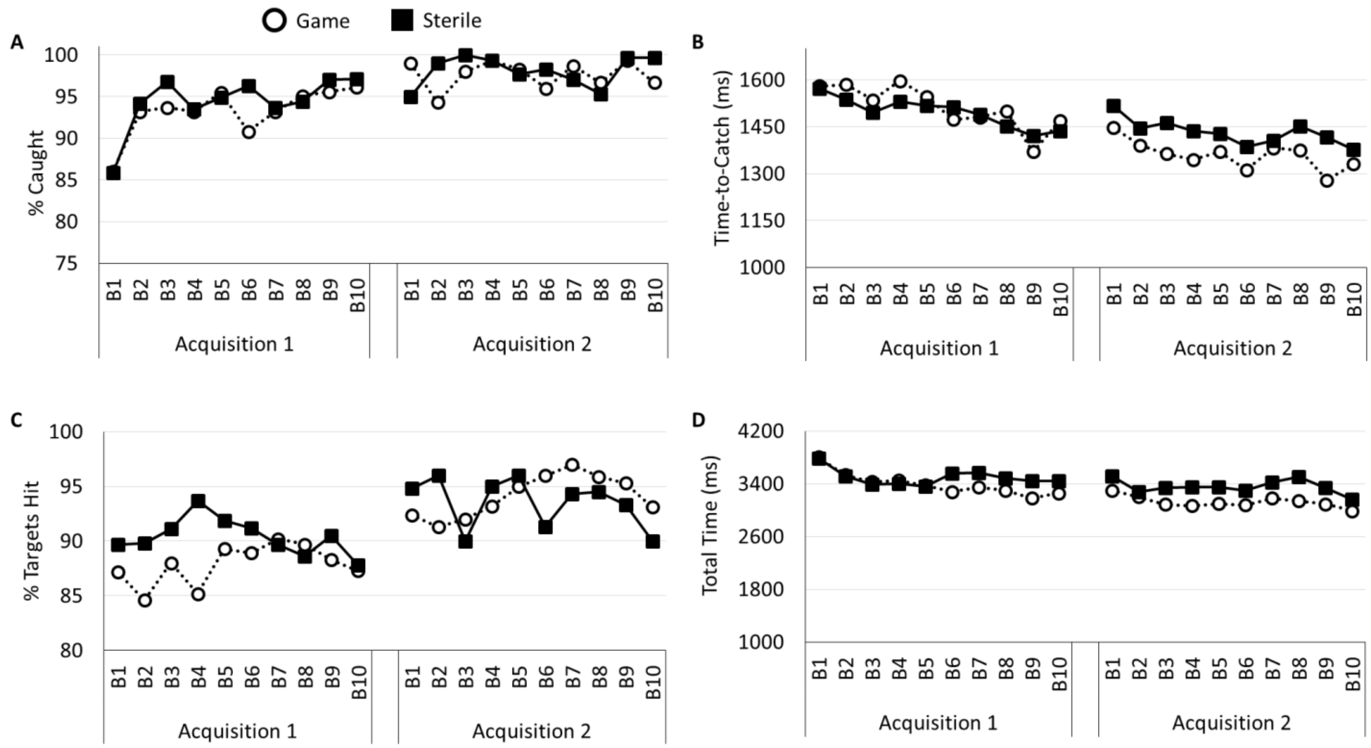
Game
Sterile
A
% Caught
100
95
90
85
80
75
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10
Acquisition 1
Acquisition 2
B
Time-to-Catch (ms)
1600
1450
1300
1150
1000
C
% Targets Hit
100
95
90
85
80
D
Total Time (ms)
4200
3400
2600
1800
1000

Figure 5.