Quantifying Individual Differences, Skill Development, and Fatigue Effects in Small-Scale Exertion Interfaces

Mike Sheinin and Carl Gutwin
Department of Computer Science, University of Saskatchewan
110 Science Place, Saskatoon, Saskatchewan, Canada
mike.sheinin@usask.ca gutwin@cs.usask.ca

ABSTRACT
Game mechanics in sports video games for skills like running and throwing are nothing like those skills in real sports. Adding small-scale exertion to the control scheme – using small muscle groups such as hands and fingers – can re-introduce some degree of physicality into sports video games. However, there is little quantitative knowledge about how small-scale exertion affects individual variability, skill development, or fatigue – and how it compares to traditional game mechanics. We carried out two studies to provide this quantitative information. Our studies showed that controlling movement with small-scale exertion was significantly and substantially different from rate-based control, and that both movement and passing skills showed significant increases with practice. Our work provides valuable information that can help designers decide when and how to use small-scale exertion, and provides an empirical basis for the design of new game interaction techniques.

Author Keywords
Exertion UIs; Skill development; Individual differences.

INTRODUCTION
Sports video games – such as *FIFA Soccer* or *Madden NFL* – let people play their favorite team sports in a computer simulation. These games provide highly realistic graphics and movement, using motion capture techniques and actual visuals of real players and stadiums. However, although the appearance of on-screen characters and environments in these games is very similar to the real world, other aspects are not like sports at all. In particular, the gameplay of sports video games involves simulations of expert physical actions such as running, throwing, and kicking – but these actions in the game are performed using techniques that are very much *un*like the ways that athletes operate in the real world.

Most sports video games use a standard game controller – with right and left joysticks (thumbsticks), a directional pad, and a series of buttons – to control characters in the game. (We will use *player* to refer to the human player, and *character* to refer to the on-screen avatar). The actions that players perform to run, pass, throw, or shoot are all carried out with this controller – and are much the same regardless of which human player is at the controls. This is in stark contrast to real-world sports, where the ways that different athletes perform these actions is the basis of their skill level. For example, there are major differences in the way that a professional soccer player shoots compared to a junior athlete, but very little difference between two players of a sports video game (since both players shoot by pressing a button on the game controller). Although there are other game areas where players can improve (e.g., selecting line-ups, deciding when to change characters, learning new button combinations), important skills like movement speed and throwing are essentially undifferentiated.

Sports video games do have expertise differences – but these are mostly built into the game characters, rather than originating in the physical abilities of the human players. This means that if a human player wants to perform better in the game, they need to choose a character that has better statistics – but this is very unlike real-world sports, where expertise and skill are developed through time, effort, and practice.

Researchers have proposed that this discrepancy can be reduced by adding physical activity back into sports video games [16]. Many video games now involve some form of exertion (e.g., *Wii Sports*, or several research systems [4, 10, 12]). Because sports video games are closely associated with game controllers, however, many traditional forms of exertion interface do not apply. Instead, taking inspiration from early button-intensive games such as *Mattel Football* or Konami's *Track and Field*, researchers have proposed using *small-scale exertion* that makes use of small muscle groups such as the hands and fingers [16]: for example, using repeated thumb movements on the thumbstick to move an on-screen character; or requiring precise movements of the hand for activities such as throwing.

Previous qualitative studies have shown that small-scale exertion increases variability in movement and precision passing, allows players to develop expertise, and causes fatigue over the course of a game [16]. These studies show that small-scale exertion has potential for improving richness and expressiveness in sports video games. However, there is still little quantitative information about this approach. In particular, three questions remain unanswered: how much variability is introduced by small-scale exertion; what is the rate and magnitude of improvement over time; and how does small-scale exertion compare to traditional input schemes.

CHI PLAY 2015, October 03 - 07, 2015, London, United Kingdom

DOI: http://dx.doi.org/10.1145/2793107.2793129

To answer these questions, we carried out two empirical studies. The first study asked people to perform three skills over several sessions: a "running" race, an obstacle course, and a passing drill. Movement was done with both impulse-based controls (requiring repeated movements of the thumb) and traditional rate-based controls (requiring the thumbstick to be held down in the direction of movement); the passing drill required precise movements of the thumb to control the ball's speed and direction. The second study involved twelve people playing two versions of the game over 6 weeks, alternating between rate-based and impulse-based controls.

We gathered performance data from the first study to investigate player variability, fatigue, expertise development, and differences between impulse and rate-based controls; we used the second study to gather illustrative examples of phenomena that were identified in the first study, and to test subjective player engagement with the two control schemes.

The two studies provide five main findings.

- There was substantial individual variability in movement skills with the impulse-based controls (12% variation from the mean), but not with rate-based controls (1.5%)
- Small-scale exertion led to significantly larger skill development – people were 11% faster in the final race with impulse control, but only 0.8% faster with rate control.
- In the passing drill (which used the same mechanic for both versions), there was also a significant improvement over time – error was reduced by 23% by the final session and the number of targets hit improved by 44%.
- There were substantial effects of fatigue in the race test: players had far more speed drops with impulse control (5.4 vs. 0.24 for rate) and were much less able to run at near-maximum speed (14% for impulse vs. 95% for rate).
- Player ratings of engagement in the second study were significantly higher for the small-scale exertion approach; other subjective measures (such as overall preference) were split between the two versions of the game.

Overall, our studies demonstrate that small-scale exertion provides significantly more variability than standard control schemes, and show that both small-scale exertion and high-precision control allow people to significantly improve their game skills. The studies provide new data about the differences that can be expected from adopting these approaches, and provide additional evidence that small-scale exertion can increase the complexity and richness of sports video games.

BACKGROUND

Traditional Sports Video Games and Controls

Sports video games are a large part of the video game industry – 12% of console sales [5]. Little research has been done on this genre, although studies have considered the audience for these games [17], and the link between playing a sports game and success in the real-world version of the sport [3].

Most sports video games provide similar play experiences for their users – they typically use game controllers as input devices, and provide a realistic visual simulation of a professional sports league. Games can involve both on-field play and team management, but here we focus only on the on-field activities. For this part of the game, there are two main game mechanics: moving and passing/shooting.

Traditional sports video games use *rate-based movement*: the player manipulates the thumbstick of the controller, and the direction and displacement of the stick from its center position determines the direction and percentage of maximum speed the character will move. Movement speed is dependent on the statistics given to the character by the developer: a character with a higher speed rating will always run faster.

Sports games control passing and shooting using *managed throwing*, where the computer controls most of the aspects of the action. Passing in a sports game usually only requires a single button press – the trajectory and distance of the pass is entirely controlled by the game. In most games, indicating a general direction and pressing the pass button will usually make a perfect pass to another character on the same team.

These control schemes mean that fatigue is not a factor for players in sports video games. The game might implement an artificial fatigue system (e.g., an energy gauge), but this is never a measure of the human player's physical fatigue.

Exergames and Small-Scale Exertion

An exergame is any game that uses physical exertion [13]. There are many examples of exergames in research [4, 12] and in commercial video games such as *Wii Sports*. The goal of exergames is mostly to show health benefits by getting players to be more active [4]; researchers have also looked at fatigue as a game design element [10]. Recent work has pointed out some drawbacks of these systems, including their full-body nature, their need for specialty equipment, and their departure from traditional gaming controls [16].

Small-scale exertion has been investigated as a way to alleviate the drawbacks of full-body exertion interfaces. Small-scale exertion uses small muscle groups – fingers, hands or feet – for repeated movements [16]. Several early games provide inspiration for the small-scale approach. One of the earliest examples of this is *Mattel Football*, a handheld video game made in 1977. This game required quick presses of arrow buttons to try and dodge defenders on a 9x3 grid [9]. Other early examples included the Track and Field and Olympic video games released in the 1980's (e.g., Konami's *Track and Field* [20] and *Daley Thompson's Decathlon* [18]). These games included events such as the 100m dash and the long jump. Movement and jumping mechanics in these games were limited to button-pressing – for example, running was carried out by alternately pressing two buttons as fast as possible. Although seen in several titles, players cannot sustain this mechanic for long – as a result, most events in these games were very short (e.g., less than 10 seconds). Rapid button-pressing does not work well in traditional sports video games where players play for longer periods (e.g., 10-40 minutes). In the 1990s, this game mechanic was used less often – although there are a small number of recent commercial games that mimic these early methods (e.g., *Mario and Sonic at the Olympic Games* [19]).

Small-scale exertion is not limited to these button-pressing mechanics, however, and can be used in modern video game environments with some adaptation. Instead of alternating buttons, games have used rapid thumbstick flicking to control movement and throwing [16]. This type of movement cannot be sustained at maximum for long, but small-scale exertion can also have interesting effects on other gameplay elements. For example, players need to be strategic about their energy use – a full-speed burst when on a breakaway, playing goalie or defense for a quick rest, or move slowly into position for a pass. Recent small-scale exertion mechanics allow players to play for extended periods of time without breaks. A recent study also showed that small-scale exertion led to greater differentiation between players, enabled increased opportunities for expertise development, and made fatigue an important gameplay element [16].

Motor Skills, Practice, and Fatigue

A motor skill is a voluntary pattern of movement acquired through practice that is used to complete a task [8]. Motor skills can be divided into two groups: gross motor skills involve coordinated movement of large muscle groups, such as when jumping, and fine motor skills involve moving smaller muscle groups such as the fingers. Early research suggested that each person had a certain capability to move and that this capability generally worked the same for all motor skills. Recent work, however, suggests that there are many different motor abilities (possibly more than 100) that are independent from each other and that certain skills may involve many of these abilities at once [15].

A critical aspect of motor skill is that it can be improved through learning. Schmidt and Lee list four important characteristics for motor learning: learning is a process of gaining capabilities to become more skilled; learning directly comes from practice and experience; we are not able to observe learning directly, we can only infer that learning has occurred by observed behavior change; learning seems to produce permanent changes in the capability for skilled behavior [15]. Most motor-skill learning is achieved through practice. There is a large body of research dedicated to practice in many areas (e.g., sports, cognitive skills, motor skills). Most of this research shows that learning follows the *power law of practice*: people improve quickly, and can continue to improve (although at a reduced rate) for a long time [15].

Fatigue is the impairment of performance which increases the perceived effort needed to perform a force and, in time, degrades the ability to produce the force [6]. Physiological research has also shown that different muscle types (e.g., "fast-twitch" vs. "slow-twitch") fatigue at different rates [2], that there is high variability (across muscle groups and individuals) in the ability to recover from fatigue [10], and that fatigue can occur peripherally (i.e., due to physiological processes in the muscles themselves) or centrally (usually involving exhaustion more generally).

Fatigue must be considered carefully when used in game design because of the risk of over-exertion or injury. However, fatigue as a game design element does not imply over-exertion. In games that are more complex than repetitive button presses, fatigue can have effects without over-exertion (e.g., forcing strategy changes). Fatigue may also not be the only mechanic of the game. Designers could vary exertion to avoid over-exertion of one muscle group, or they could mix exertion with non-exertion mechanics.

STUDY SYSTEM AND GAME MECHANICS

We modified a small-scale exertion system used in previous research [16]. The game, called Jelly Polo, is a simple three-on-three team game similar to handball. The onscreen character has an arm that can hold and throw the ball; the body and arm are controlled with the left and right thumbsticks of a standard game controller. Jelly Polo uses small-scale exertion in both movement and throwing.

Impulse-based movement. The left thumbstick controls the movement of the character, and uses an impulse mechanism rather than rate-based control. Impulse-based movement requires the player to repeatedly flick the left thumbstick to move; the direction and speed of the flicks control the direction and speed of the impulse given to the character.

Precision throwing control. The right thumbstick controls passing and shooting. A threshold-based release mechanism makes the ball shoot out from the character in the exact direction the thumbstick is flicked and at a velocity that is based on the speed of the flick. This means the accuracy of shots and passes is solely based on the player's ability to flick in the desired direction at the desired speed.

In order to compare small-scale exertion with the standard controls available in sports video games, we created a new version of the Jelly Polo game that used rate-based movement control. In this version, holding the left thumbstick moved the character at a rate determined by the displacement of the stick – for example, players could simply push the thumbstick over to move at maximum speed. We did not add the "managed throwing" scheme to this version because this scheme essentially leads to error-free performance, and so both versions of the game used precision passing control.

STUDY 1: MOVEMENT AND PASSING SKILLS

Our first study examined basic skills in the Jelly Polo game over several sessions, and with both impulse and rate movement control. Our goals were to quantify individual variability, skill development over time, and differences between the two control schemes.

Experimental tasks (game skills)

Speed test: "running" race

Participants used their controller to move an on-screen character across the game screen and back again, as quickly as possible (see Figure 1). Participants started at the left side of

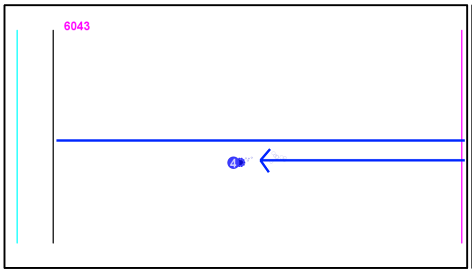

Figure 1. Race in progress (timer at top left; arrow shows motion, not shown in trials).

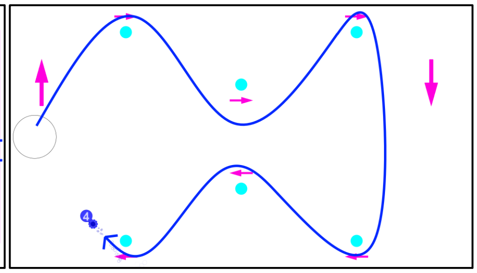

Figure 2. Obstacle course in progress (curve shows path, not shown in trials).

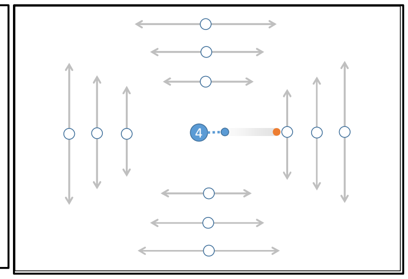

Figure 3. Passing task, showing all target locations.

the screen, and after a 5-second countdown, "ran" to touch the right edge of the screen, then turned around and returned to the finish line at the left side. The right wall turned green once they had touched the wall successfully. In order to give the participants a goal to beat for the next race, the total time was displayed during and after the race.

Maneuverability test: obstacle course
In this task, participants moved their character in a particular path around a series of obstacles, as quickly as possible. Participants began in the "start" circle (Figure 2, left), and after a 5-second countdown, began moving around the obstacles in the path shown in Figure 2. They finished by returning to the start circle. Once a marker arrow was passed, it turned green. If the participant hit an obstacle, they would bounce off, causing people to lose time and disrupting their rhythm. Participants ran 10 laps of the obstacle course in a row. Each lap time was displayed during the next lap's countdown, again to provide a goal and spur performance.

Accuracy test: precision passing
In this task, participants controlled their character to throw a ball at a series of moving targets. Figure 3 shows a diagram of the task setup and all targets. The participant was put in the middle of the screen without the ability to move. A series of round targets appeared, one at a time, as shown in Figure 3; the targets moved back and forth in a predictable linear pattern. There were three difficulty levels in the targets: easy (close+slow), medium (middle distance and medium speed), and difficult (far+fast) and four directions: right, up, left, down. There were 12 targets (3 difficulty x 4 directions).

The participant's goal was to throw the ball to the target. Participants were clearly told that speed did not matter in the passing section of the study; we were only asking them to be as accurate as possible. Once the participant attempted the pass, we kept track of the shortest distance between the ball and the target throughout the entire motion of the ball. If the ball made contact with the target, the target would turn green.

Once the ball either hit a wall or stopped its motion, it was moved back to the participant for the next throw. Participants were given five attempts at each target one by one (e.g., five attempts at right+easy then move to the next target). Once the five attempts were taken, the next counter-clockwise target appeared; in addition, all targets of one difficulty level were completed before moving to the next level.

Study Methods

Participants and Procedure
Twelve participants were recruited from a local university (7 male, 5 female; mean age 25.2 years). 9 of the participants played games regularly (> 3-7 hours/week), and 7 people were familiar with sports video games (> 1 hr/wk).

Each participant performed the skills tasks described above on three different days, with gaps of 2 days and 7 days between sessions. In each session, participants used the impulse-based movement controls first, and then the rate-based controls (this fixed ordering was done as our main goal was to explore the impulse-based controls, and we did not want behavior in this condition to receive benefit of practice). With each control type, participants first completed a running race, then did 10 laps of the obstacle course, then a second running race. After these six tasks (3 tasks x 2 control types), participants performed the passing test.

The system logged all performance data; participants also filled out questionnaires after each control type, and at the end of the study. Overall, each participant completed 6 running races for each control scheme, 30 laps of the obstacle course for each control scheme, and 3 passing drills.

Design and Analysis
Running race. To look for differences in individual variability across control types, we re-coded our dependent measures (average speed and max speed) as the percent difference between the participant's score and the overall mean (we call this measure *variability*). We analyzed variability across control type with a one-way RM-ANOVA with a single within-subjects factor, ControlType (impulse or rate). To look for expertise development, fatigue effects, and differences across control types, we used a 2x2x3 RM-ANOVA with three within-subjects factors: ControlType (impulse, rate); Race (1, 2); and Session (1-3). Dependent variables were average speed and maximum speed. We also looked for fatigue within races with two additional measures: number of speed drops during the race, and the amount of time participants were able to move at nearly their maximum speed.

Obstacle Course. We measured participant variability with the same method used for the running race. We analyzed variability across control type with a one-way RM-ANOVA with the factor ControlType (impulse, rate). To look for expertise development, fatigue effects, and differences across control types, we used a 2x10x3 RM-ANOVA with three within-subjects factors: ControlType (impulse, rate); Lap (1-10), and Session (1-3). Dependent variables were average speed and number of collisions.

Passing. Both control types used the same passing controls, and fatigue was not a factor in the passing test, so here we analyze individual variability and expertise development. Variability is again calculated as the percent difference between a participant's accuracy and the mean accuracy, and a participant's precision and the mean precision. There is no comparison of control types, because the real-world technique used in sports games is managed throwing, which has a near-100% success rate for passing. Expertise development is analyzed using a 4x3x3 RM-ANOVA, with three within-subjects factors: Direction (up, down, left, right); Difficulty (close+slow, medium+medium, far+fast); and Session (1-3). Dependent variables were error and number of targets hit.

Results – Skills Study

We organize our results below by the three skills we tested: running, maneuvering (obstacle course), and passing. For each skill, we explored four main issues:

- *Individual variability* – were players different in their performance, and by how much;
- *Expertise development* – how did player performance change over time, and by how much;
- *Fatigue* – did performance change over an individual session, and by how much;
- *Differences between control schemes* – are there differences between impulse-based and rate-based controls.

It is important to note that we do not compare values between the control schemes, as these are arbitrary depending on how the developers determine the speed variable in their system.

Running race

For the running race, we collected maximum speed and average speed, and also calculated variability (mean percent difference from the group average, as described above). Figures 4-6 show the results for the two control schemes.

Running – individual variability

RM-ANOVA showed a main effect of Control Type on Max Speed Variability ($F_{1,11}$=18.51, p<0.001). The variability with impulse-based control was more than 6%, and less than 1% for rate-based control.

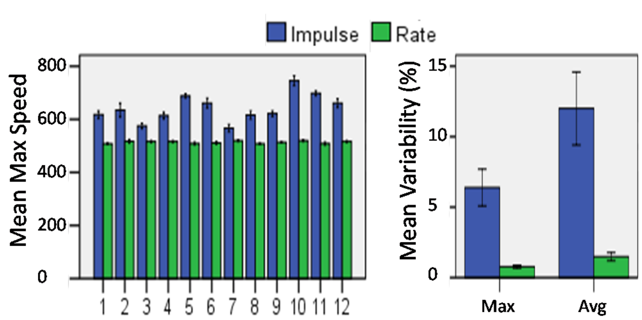

Figure 4: (Left) participant max speed (px/sec) (6 races, ±s.e.); (Right) variability in max and average speed

We carried out similar analyses on average speed. RM-ANOVA showed a main effect of Control Type on Average Speed Variability ($F_{1,11}$=16.20, p=0.001). There was much higher variability in performance with impulse-based control (12%) than with rate-based (1.5%) as seen in Figure 4.

Running – expertise development

RM-ANOVA showed a main effect of Session on Average Speed ($F_{2,22}$=325.32, p<0.001). There was also an interaction between Control Type and Session ($F_{2,22}$=228.80, p<0.001); Figure 5 shows that the increase in speed for impulse-based control was much larger than for rate-based control (impulse: 11.3% faster by final session; rate: 0.8% faster).

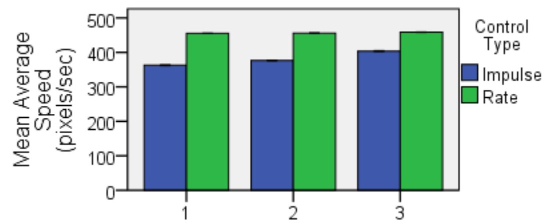

Figure 5. Mean avg. speed (± s.e), by session and control type.

For Max Speed, an RM-ANOVA showed no main effect of Session ($F_{2,22}$=1.27, p=0.28) and there was no interaction between Control Type and Session ($F_{2,22}$=1.07, p=0.35).

Running – fatigue

We looked for speed changes during a race (short-term fatigue), and changes across races (longer-term fatigue). To examine short-term fatigue, we used two performance measures: first, the number of speed drops of more than 100 pixels/sec, indicating the number of times people slowed down substantially (we used a 15-sample rolling average for this measure, to smooth the effects of the impulse-based mechanism); and second, the fraction of the total race where speed was at or above 90% of the max speed for the race.

Figure 6 shows representative data from impulse and rate-based races for one participant; these charts clearly show the higher variability in the impulse version. RM-ANOVA on both measures of short-term fatigue showed significant differences between impulse and rate-based movement: for number of speed drops, $F_{1,11}$=110.01, p<0.001 (mean of 5.4 drops for impulse vs. 0.24 for rate); for time near maximum speed, $F_{1,11}$=2298, p<0.0001 (13.9% for impulse, 95.5% for rate).

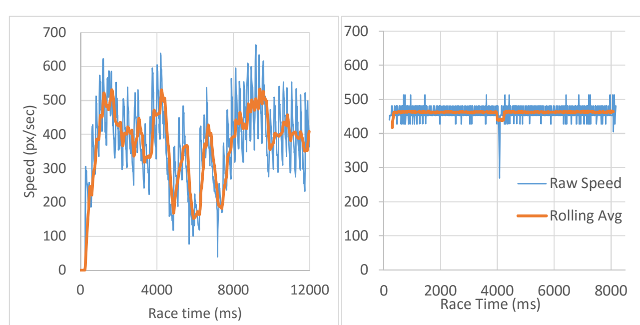

Figure 6. Example data for impulse (left) and rate races.

These results are in line with prior research on muscle fatigue. The muscles of the thumb and hand (used for repeated flicking motions) contain fast-twitch fibres that are susceptible to short-term fatigue, but can recover quickly [2]. This suggests that people cannot continuously maintain their maximum movement frequency, but can increase speed again after a short recovery period of slower movement or rest – which mirrors the bursty speed profiles seen in the data.

We also investigated longer-term fatigue by looking for declining average or maximum speeds in the second race of each session. However, there was no significant decrease for maximum speed ($F_{1,11}$=2.74,p=0.099), and for average speed, there was actually a small increase ($F_{1,11}$=223.45, p<0.001). These results suggest that practice effects overshadowed longer-term fatigue. Finally, subjective responses indicated that players felt the effects of fatigue (see below).

Obstacle course

Maneuvering – individual variability

RM-ANOVA showed that the effect of Control Type on Average Speed Variability was significant ($F_{1,11}$=29.95, p<0.001). Variability with impulse control was more than 8% on average, and 1% for rate-based. RM-ANOVA also showed a main effect of Control Type on Collisions ($F_{1,11}$=5.22, p=0.032). Overall, there were more collisions with impulse control than with rate-based control (possibly due to the higher difficulty of this method).

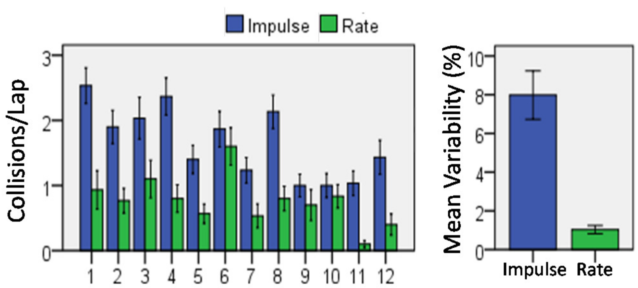

Figure 7. Left: mean collisions/lap (± s.e.) Right: mean variability in average speed.

Maneuvering – expertise development

RM-ANOVA showed a significant effect of Session on Average Speed ($F_{2,22}$=3012.35, p<0.001). There was also an interaction between Control Type and Session ($F_{2,22}$=1774.99, p<0.001); Figure 8 indicates that the increase in speed for impulse control was much larger than for rate (impulse: 13.3% faster by the final session; rate: 1.2% faster).

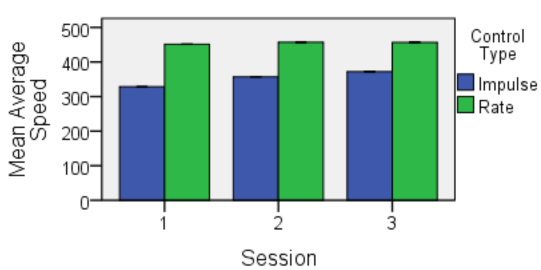

Figure 8. Mean avg. speed (± s.e), by session and control type.

Number of collisions indicates the amount of error in the obstacle course. RM-ANOVA showed no significant effect of Session on Collisions ($F_{2,22}$=2.53, p=0.08) and no interaction between Control Type and Session ($F_{2,22}$=1.78, p=0.170).

Maneuvering – fatigue

We used the same measures for short-term fatigue as described above (i.e., number of speed drops, and fraction of time spent above 90% of max). In the obstacle course, there are other factors that contribute to these measures (i.e., needing to slow down to go around obstacles) – but because the courses were equal for the two control conditions, the measures are an accurate reflection of the difference. As in the running race, there were significant differences for both measures. Impulse control had a mean of 13.2 speed drops, and rate control had a mean of 0.26 ($F_{1,11}$=436.4, p<0.0001. Percent of time near maximum speed was 5.6% for impulse, and 97.0% for rate, $F_{1,11}$=39045, p<0.0001.

To test for longer-term fatigue, we looked at average speed through the laps of the course. We found significant differences, but in opposite directions: average speed actually increased for impulse control (21.47 pixels/sec faster by the final lap; $F_{9,99}$=82.33, p<0.001), and decreased for rate control (7.69 pixels/sec slower by final; $F_{9,99}$=81.13, p<0.001). It is possible that boredom might be a factor in this result.

Passing

We gathered data about both *error* (the minimum distance to the target for each throw), and *accuracy* (the number of targets successfully hit with the ball). Note there was no difference in the control schemes for passing, because "managed throwing" would lead to nearly 100% accuracy and 0% error.

Passing – individual variability

An RM-ANOVA showed a main effect of Participant on both Error ($F_{2,22}$=4.909, p<0.001) and Accuracy ($F_{2,22}$=2.696, p=0.020). Figure 9 shows the mean of all participants' difference from the average minimum distance and the average percentage of targets hit.

An RM-ANOVA showed a main effect of Difficulty on Error ($F_{2,22}$=33.32, p<0.001) but no main effect of Direction on Error ($F_{2,22}$=0.98, p<0.402). There was no interaction between Direction and Difficulty on Error ($F_{2,22}$=0.41, p<0.872).

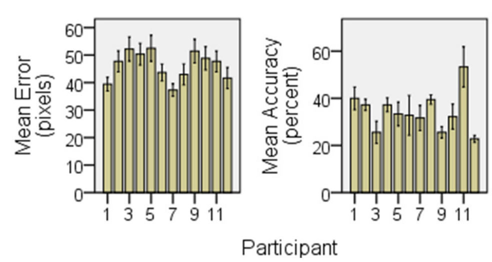

Figure 9. Mean error (left) and accuracy, by participant.

Passing – expertise development

RM-ANOVA showed significant effects of Session on Error ($F_{2,22}$=14.66, p<0.001) and Accuracy ($F_{2,22}$=4.82, p=0.015).

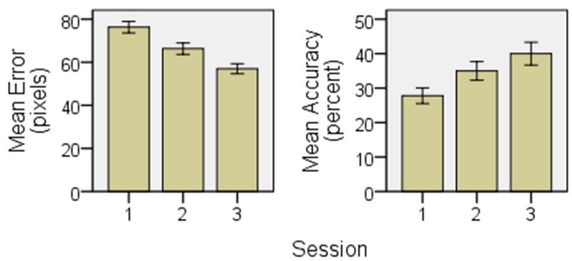

Figure 10. Mean error from target (left) and mean accuracy (right) (± s.e), by session.

Subjective Results

Participants completed questionnaires after each task. Wilcoxon tests showed that impulse was rated significantly higher than rate for physical fatigue (Z=-5.454, p<0.001), effort required (Z=-7.139, p<0.001), and potential for improvement with practice (Z=-3.856, p<0.001). Figures 11 and 12 show these ratings. We also asked participants which control type they found more boring and frustrating; see Table 1.

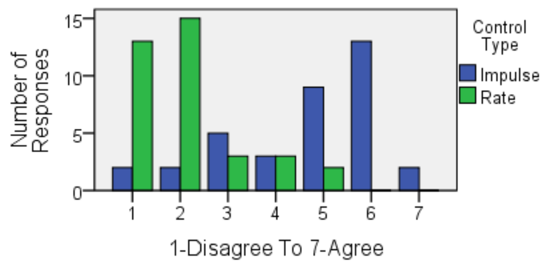

Figure 11. Responses to "the task was physically demanding".

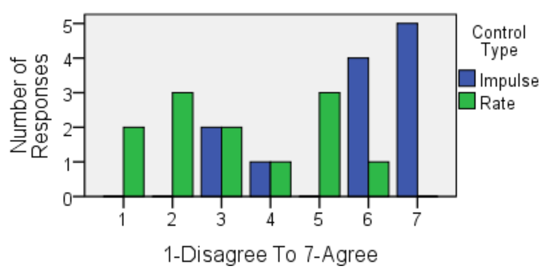

Figure 12. "I could still get faster at moving my character".

	Rate	Impulse	Equal
More boring	7	2	3
More frustrating	2	8	2

Table 1. Subjective responses (number of participants).

STUDY 2: GAMEPLAY STUDY

We carried out a second study to follow up on the results of study 1 in a real game environment. We had first year university student participants play a game using both control schemes to show that individual variability, expertise development, and fatigue show similar results to study 1 in terms of the difference in control type.

Study Methods

We ran a Jelly Polo league with four teams of three participants. Movement and passing controls were as described for the first study. All participants had experience playing video games with standard controllers. Each team played six games: half played the first 3 games with rate-based movement, and the second half with impulse-based. The other two teams reversed this order. The study ran for six weeks. Every week, four teams played in two 10-minute games.

Results

For our data analysis, we only include the data of 12 participants (10 male, 2 female) who played regularly (there were some absences from week to week). All participants included in the analysis played 5 or more of the 6 games.

Because of the variable and unpredictable nature of the game (i.e., players play different positions, do not move at top speed the entire game because of varying factors, etc.) and the fact that participants traded off because of absences, we did not record which participant was which character every game. This made it difficult to track expertise development.

We did find individual variability, however. For example, the max speeds for impulse-based games were highly variable (12.19% difference from the mean), whereas they were not in rate-based games (0.04% difference). Average speed was variable in both control types (impulse=12.75%, rate=10.09%), but this could be attributed to ordinary gameplay (e.g., players moving from forward to defense or goalie). Figure 13 shows variability differences between the control types for max speed and average speed.

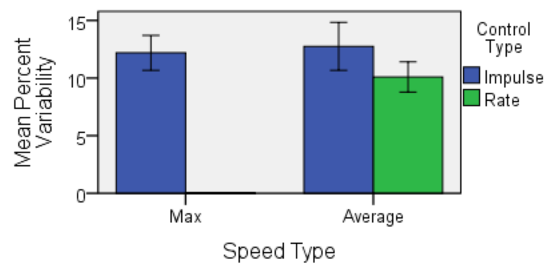

Figure 13. Variability in max and avg speed, by control type.

Fatigue can also be suggested from our data. Players were much more likely to move near their max speed in the rate-based version than the impulse-based version (although this is a weaker measure of fatigue than it was in the running race, since there are more reasons in a real game as to why people might move at different speeds). Players were above 90% of their max speed 60.24% of the time for rate-based and only 0.53% for impulse-based. In comparison, the time spent below 10% of max speed was roughly equal – 8.57% for rate-based control and 10.58% for impulse-based.

We compared enjoyment ("Regardless of outcome, Jelly Polo was fun to play") and engagement ("Jelly Polo was engaging") scores collected after each game (7-point scale), using Mann-Whitney U tests. For enjoyment, there was no difference found between impulse and rate control (impulse mean 6.16; rate mean 6.35; Z=-0.87, p=0.38). For engagement, there was a significant difference (impulse mean 6.31; rate mean 5.71; Z=2.12, p=0.034). Finally, preferences at the end of the study were evenly split between rate and impulse control (6 for each).

DISCUSSION

Our studies provide the following main findings:

- Individual variability in movement speed was much higher with impulse control than with rate-based control. Where rate control varied only by about 1% across participants, impulse control varied by about 10%.
- Skill development in speed was also much greater with impulse control – people improved by more than 10%, whereas with rate control they improved by only 1-2%.
- There was substantial skill development in passing (error decreased by 23%; accuracy increased by 44%).
- There were substantial fatigue effects within races for impulse control (significantly more speed drops, and significantly less time spent near max speed), whereas rate control showed essentially no effects of fatigue. We did not see longer-term effects of fatigue (between races).
- Variability results were confirmed in the gameplay study, and player ratings of engagement were significantly higher for the small-scale exertion approach.

In the following sections we interpret these results and discuss their importance for informing the design of future games based on the idea of small-scale exertion.

Explanation and interpretation of results

Movement variability

The goals of small-scale exertion interfaces are to allow for higher individual variability and to enable greater skill development over time. Our studies showed that these goals were met for most of the measures in the study. The main reason why impulse-based control had these characteristics is that it provides a higher-bandwidth control scheme that depends more on player actions than on an algorithm in the game. Human physical abilities of all kinds are highly variable, and impulse-based control superimposes this variability onto the actions of the on-screen character. Similarly, because of the underlying human variability, there is more room for improvement in most players – whereas with rate-based control, everyone is already near the top of the performance curve, so there is little opportunity for change. In addition, in both the running race and obstacle course, players in the rate condition almost never changed their speed from maximum (even though the controls allowed it, and although we demonstrated this to participants). This result suggests that even when a control scheme has a large expressive range, people may not use that range if it is too easy to simply maximize the variable at all times. The impulse control scheme, in contrast, imposed a physical limit on the duration of maximum speed, and so players were forced to make use of a larger portion of the control's range.

Although the differences in variability were expected, the amount of variability and improvement that we would see was unknown. The ranges that we found in the study – variability of approximately 10% from the mean, and of approximately 12% for skill improvement – are interesting from a game-design perspective because the amount of variability in the basic game mechanics can help to determine the audience for the game. For example, Jelly Polo is designed to be an accessible game for a wide audience – and this suggests that the control schemes should allow a wide range of people to play in a walk-up-and-use scenario.

Ten percent variability means that there are noticeable differences among players, but that no one person will be able to completely dominate the game. The initial variability and the opportunity for improvement are similar to the idea of "floor and ceiling" that has been proposed for user interfaces [14] – that is, the amount that people can do when they first start with a system, and the highest level of performance that they can achieve. In these terms, small-scale exertion controls are interesting because they have a higher floor (i.e., they are more difficult to begin with); in contrast, rate-based control and managed throwing schemes start all players near the performance ceiling. It may be true that, for example, a more difficult obstacle course could cause more variability with rate-based controls. However, the variability displayed would mostly be in maneuverability, not speed. It is possible that the lack of variability in sports video games arose from an intention to provide equality for all players – but our participants' subjective responses suggest that taking over too much of the action leads to a reduction in engagement.

In addition, the amount of potential improvement for impulse control in Jelly Polo is roughly in line with the amount of individual variability – meaning that with practice, players will experience an improvement that is similar to the differences that they see between players.

Other games may have different characteristics in terms of these issues. For example, the designer of a competitive first-person shooter may want a wider range of possible expertise and a much larger amount of possible improvement, in order to keep players interested and give them more to strive for.

Variability and the value of unpredictability

The higher engagement scores that the impulse-based version of Jelly Polo received compared to the rate-based version may arise from the variability that is inherent in the scheme. Game researchers have noted that games are more enjoyable and engaging when there is greater suspense about the outcome [1] – for example, Mueller states "uncertainty contributes to an element of suspense and facilitates surprise in games through random or chance events, which can play an important part in what makes a game engaging" [13]. The uncertainty that arises from the variability in the impulse-based control scheme (both in terms of individual differences and changes over time) is one kind of randomness that may contribute to our positive engagement results.

Fatigue and maneuverability results

We saw local effects of fatigue in both movement activities of the study (the race and the obstacle course). The movement patterns for impulse control (i.e., speed over time) follow the expectations suggested by physiological research on muscle fatigue. The muscles of the thumb and hand contain fast-twitch fibres that are susceptible to short-term fatigue, but can recover quickly [2].

Our measure of fatigue used performance data rather than physiological data (e.g., we did not test for lactic acid levels or other physiological occurrences). However, the simple nature of the running race and our clear instructions to move as fast as possible suggest that either true fatigue was occurring, or that people were adjusting their behavior in order to avoid future fatigue (e.g., slowing down to conserve energy). From a game design perspective, both of these results are valuable – people cannot continuously maintain their maximum movement frequency, but can increase speed again after a short recovery period. Nevertheless, we plan future studies that measure fatigue more directly (e.g., with a maximum voluntary force test after each race [10]).

We did not observe longer-term effects of fatigue (i.e., across laps of the obstacle course, or between races). This may be due to the shorter duration of the activities, compared to previous studies that reported substantial fatigue [16]; longer races might likely show more global fatigue effects. In addition, increasing expertise (or better strategy) may have counteracted any long-term fatigue effects. Finally, subjective responses showed that people felt fatigued during the activities; from a design perspective, creating the perception of fatigue may be as important as the actual phenomenon.

Finally, our maneuverability results did not show differences between impulse and rate control – and in fact, collisions increased for impulse control in the third session. From our observations, we believe that this is a result of participants becoming better at movement, and then attempting to go faster and tighter around the obstacles. This may have created a speed/accuracy trade-off and led to an increase in collisions. However, our obstacle course was relatively simple, and did not demand complex turns. It is possible that with a more challenging course, we would begin to see variability in the rate-based control – because maneuvering involves steering as well as movement, and players will be able to improve this skill for both movement-control schemes.

Further work and application to sports video games

Our studies provide evidence that small-scale exertion is a feasible design approach for team sports games and one that can provide reliable effects on both performance and engagement. Developers of other games should be able to take up this idea and use our results to create certain kinds of experiences based on individual variability, skill development, and fatigue. There are several ways, however, that our work can be extended to further explore the generalization of the idea.

First, it is not known whether small-scale exertion can be implemented in different ways, particularly with a standard game controller – but popular games such as *Joust* or *Flappy Bird* show that other instances of the idea are certainly possible, and there are now several sensors built into some controllers in addition to joysticks and buttons. A wider range of techniques may be desirable in order to enable different skills and difficulties in different games (e.g., avatar running in a role-playing game might require a different kind of small-scale exertion than the impulse-based control that worked well for the jellyfish characters we tested with).

Second, little is known about the limits on difficulty. An underlying principle in small-scale exertion is to provide a higher-bandwidth control interface; but the limits of these interfaces are not well studied. The game *QWOP* (foddy.net) is an example of an extraordinarily difficult control scheme that has become popular primarily because of its difficulty. Further work is needed on ways that game designers can balance difficulty, learnability, and engagement for small-scale exertion interfaces.

Third, the engagement of small-scale exertion does not necessarily translate into overall preference. In real-world sports games, some players may prefer being able to operate the game without a great deal of effort. Therefore, more work is needed to find ways of allowing different play styles to coexist – for example, we plan to consider how a game can be balanced if some players are using an exertion interface and others are using a more traditional control scheme.

CONCLUSIONS

Sports video games use mechanics for skills like running and throwing that are nothing like those skills in real sports. Small-scale exertion is one way to re-introduce some degree of physicality into sports video games – but there is little quantitative knowledge about how small-scale exertion changes games, and how it compares to traditional mechanics. This is critical because these factors could be used in game design (e.g., fatigue can force teams to change strategy; individual differences increase unpredictability; expertise development rewards practice). Without knowing how variable individuals are, how quickly expertise builds up, or when fatigue becomes an issue, designers do not know whether small-scale exertion is a viable mechanic for their game, or how it can be used in game design. Although trial-and-error approaches and play-testing can work, existing small-scale exertion games (such as *Track and Field*) show that small-scale exertion has been used in ways that lead to relatively simple games (i.e., bursts of button mashing).

We carried out two studies to provide this quantitative information. Our studies show that impulse-based movement control is significantly and substantially different from rate based control, and that both movement and passing skills show significant increases with practice. Our work provides valuable information about magnitude and change of expertise development, individual differentiation, and fatigue. For example, it is shown that for impulse-based movement, individual variability exists but is not so large that it would make games unplayable. We also show that the increase in expertise roughly matches the initial individual differences – designers could use this to help balance games. This knowledge will help designers decide when and how to use small-scale exertion, and provides an empirical basis for the design of new game interaction techniques.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada, by the GRAND NCE, and by the SurfNet Strategic Research Network.

REFERENCES

1. Abuhamdeh, S., and Csikszentmihalyi, M. The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities. *Pers. & Soc. Psych. Bulletin.*, 38 (3), 2012, 317-330.
2. Allen, D., Lamb, G., Westerblad, H. Skeletal Muscle Fatigue: Cellular Mechanisms. *Physiol Rev*, 88, 2008, 287-332.
3. Dörrfuß, K. Bader, F. Wegener, R. Siemon, A. Schwake, J. Horman, F. Hieber, T. Schmid, U. Video Games can Improve Performance in Sports – An Empirical Study with Wii Sports Bowling. *KI 2009*.
4. Castañer, M. Camerino, O. Parés, N. Landry, P. Fostering Body Movement in Children Through an Exertion Interface as an Educational Tool. *Procedia – Social and Behavioral Sciences*, 2011, 236-240.
5. ESA. *Essential Facts about the Video Game Industry,* http://www.theesa.com/facts/pdfs/esa_ef_2014.pdf.
6. Enoka, R.M. Stuart, D.G. Neurobiology of muscle fatigue. *J. Applied Physiology*, 72(5), 1992, 1631-48.
7. Heathcote, A. Brown, S. Mewhort, D. J. K. The power law repealed: the case for an exponential law of practice. *Psych. Bulletin & Review,* 7(2), 2000, 185-207.
8. Jennett, S. *Dictionary of Sport and Exercise Science and Medicine*. London: Churchill Livingstone, 2008.
9. Mattel Electronics. *Football Manual*, www.handheld museum.com/Manuals/Mattel-Football.pdf.
10. Merton, P.A. Voluntary Strength and Fatigue. *J. Physiol.* 123, 1954, 553-564.
11. Mueller, F., Agamanolis, S., Gibbs, M., and Vetere, F. Remote impact: shadowboxing over a distance. *Proc. CHI 2008*, 2291-2296.
12. Mueller, F., Cole, L., O'Brien, S., Walmink, W. Airhockey over a distance: a networked physical game to support social interactions. *ACE 2006*, 1133-1138.
13. Mueller, F., Edge, D., Vetere, F., Gibbs, M., Agamanolis, S., Bongers, B., Sheridan, J. Designing sports: a framework for exertion games. *CHI 2011*, 2651-2660.
14. Myers, B., Hudson, S., and Pausch, R. Past, Present, and Future of User Interface Software Tools. *ToCHI*, 7 (1), 3-28.
15. Schmidt, R., and Lee, T. *Motor Control and Learning: A Behavioral Emphasis*. Human Kinetics, 2005.
16. Sheinin, M. and Gutwin, C. Exertion in the Small: Improving Differentiation and Expressiveness in Sports Games with Physical Controls. *CHI 2014*, 1845-1854.
17. Stein, A. and Mitgutsch, K. Who are sports gamers? A large scale study of sports video game players. *Convergence*, 19(3), 2013, 345-363.
18. Wikipedia, *Daley Thompson's Decathlon*, 2014, en.wikipedia.org/wiki/Daley_Thompson's_Decathlon.
19. Wikipedia, *Mario & Sonic at the Olympic Games*, 2015, en.wikipedia.org/wiki/Mario_&_Sonic_at_the_Olympic_Games.
20. Wikipedia, *Track & Field (video game)*, 2015, en.wikipedia.org/wiki/ Track_&_Field_(video_game).