Effects of Balancing for Physical Abilities on Player Performance, Experience and Self-Esteem in Exergames

Kathrin M. Gerling[1], Matthew Miller[1], Regan L. Mandryk[1], Max Birk[1], Jan Smeddinck[2]
[1]University of Saskatchewan, Saskatoon, SK, CA, firstname.lastname@usask.ca
[2]University of Bremen, Germany, smeddinck@tzi.de

ABSTRACT
Game balancing can help players with different skill levels play multiplayer games together; however, little is known about how the balancing approach affects performance, experience, and self-esteem–especially when differences in player strength result from given abilities, rather than learned skill. We explore three balancing approaches in a dance game and show that the explicit approach commonly used in commercial games reduces self-esteem and feelings of relatedness in dyads, whereas hidden balancing improves self-esteem and reduces score differential without affecting game outcome. We apply our results in a second study with dyads where one player had a mobility disability and used a wheelchair. By making motion-based games accessible for people with different physical abilities, and by enabling people with mobility disabilities to compete on a par with able-bodied peers, we show how to provide empowering experiences through enjoyable games that have the potential to increase physical activity and self-esteem.

Author Keywords
Motion-based games; balancing; exergames; player experience; physical abilities.

ACM Classification Keywords
K.8.0 [Personal Computing]: General - Games.

INTRODUCTION
Many computer and video games include the element of competition between multiple players. Competition is optimal when the outcome is uncertain because the players are well matched, complicating multiplayer games where players have different skill levels. Commercial games provide balancing options so that people with different skill levels can compete, including asymmetric driver and gunner roles in Mario Kart: Double Dash, 'rubber banding' in Mario Kart, and bullet magnetism in the Halo series.

Previous research has supported game balancing. In particular, Bateman et al. [2] showed that the reduced score differential of balanced competitions increased the fun that assisted players experienced, while not changing the fun that non-assisted players had. However, these results derive from a genre (target shooting) where assistance was hidden from players. There are many genres where balancing cannot be hidden. For example, in symmetric performance games like the Guitar Hero or the Dance Central series, competition is balanced by having players explicitly select different difficulty levels. Thus players are aware that assistance is provided to one player in terms of reduced difficulty. How players respond to explicit balancing is an open question; however, the work by Bateman et al. [2] gives some insight. In their study, the majority of assisted players felt that assisting a player is unfair because the win does not reflect actual ability, the win does not count if it was actually due to computer assistance, and the scores would be meaningless. These results suggest that if weaker players win through assistance known to both players, the experience may not be positive.

Balancing competition in games is further complicated when the relative strength of the players is a matter of innate ability, rather than skill acquired through practice. For example, dominating a competition based on practiced skills (e.g., professional golfer trounces amateur friend) does not carry the same stigma as dominating a competition based on given traits (e.g., 6-foot man crushes 5-foot man in one-on-one basketball) or ability (e.g., able-bodied child routs child with mobility impairment in race). The reason for the disparity in player strength (e.g., due to skill, ability, age, fitness level) may differentially affect how players feel about games that balance competition by providing assistance to weaker players. This is important because the effects of providing assistance to players of different abilities may not only affect play experiences, but may also affect how the player feels about themselves – *empowering players through balancing assistance is desirable*, whereas, *harming players' self-esteem is not*.

To explore balancing strategies in multiplayer games, we studied the effects of balancing symmetric competition in a motion-based game. We focus on symmetric competition to explore explicit versus hidden balancing approaches and on motion-based games to explore balancing in dyads with a range of game skills and physical abilities. In Study 1, we explore the effects of 4 balancing approaches (input, time, score, and no balancing) on performance, experience, and self-esteem in a custom-built dance game. We also compare physical play to play with a gamepad to see whether differences arose when disparity in player strength was

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada
Copyright 2014 ACM 978-1-4503-2473-1/14/04…$15.00.
http://dx.doi.org/10.1145/2556288.2556963

more related to physical ability than game skill. In Study 2, we explored balancing for dyads with vastly different physical abilities–comprised of dyads where one player was able-bodied and the other used a wheelchair. Based on results from Study 1, we adapted our game to use a combined balancing approach so that the able-bodied players used the dance mat, whereas the participants in wheelchairs played by moving around with the wheelchair.

Results of our studies show that: 1) explicit input balancing, similar to the approach used in commercial games, reduced feelings of self-esteem in both the weaker and stronger players, and resulted in lower relatedness between the pairs; 2) hidden time balancing best balanced score differential, but did not change the game outcome (i.e., stronger player still won more often), and also improved self-esteem for both players; 3) weaker players felt less guilty competing against their partner when using the motion-based controls than the sedentary controls, where disadvantage was more related to skill than physical ability; 4) people who use wheelchairs performed well in the game and enjoyed competing in a motion-based game against their able-bodied friends; and 5) the able-bodied competitors were concerned about ensuring a level playing field, which is possible by tuning the balancing parameters in our game.

Balancing motion-based exergames is an important step toward engaging people with different skills and abilities in physically-challenging play, which can reduce sedentary behaviour and benefit health. By making motion-based games accessible for people with different physical abilities, we can provide enjoyable games with the benefit of increasing physical activity. More importantly, by enabling joint motion-based play using hidden balancing, we can help increase self-esteem by providing empowering experiences for players.

RELATED WORK

Defining Skills and Abilities

Skills generally refer to a set of learned abilities to perform a task. In the physical context, Merriam-Webster defines skills as "dexterity or coordination especially in the execution of learned physical tasks"[1]. In contrast, abilities are referred to as "the quality or state of being able"[2], emphasizing that certain abilities are given, and not acquired through practice. Thus, individual performance depends on the basic ability to participate, e.g., ability to run, and on the acquired skills, e.g., dribble a soccer ball.

Physical Competence and Self-Concept

Cairney et al. show that physical activity can improve well-being [5], but positive effects can be lost in competitive settings. On a basic level, perceived physical competence influences whether people enjoy physical activity, and follow through with physical activity in adulthood [5], but can depend on comparisons to peers. This may lead to negative self-perception, despite increased objective performance [16]. These findings are supported by research by Goni et al. on sport in schools [9], showing that sports have a positive impact on students' self-concepts, but that balanced competition leads to greater improvements. If perceived and objective relative physical skill and ability in sport influences an individual's self-concepts, we must explore whether these effects prevail in video games that require physical activity, and which balancing strategies can be applied to ensure a positive player experience.

Balancing Skills and Abilities in Sports and Games

In sport, balancing is often applied to improve competition between players with different abilities (physical, cognitive) or skill to compete. Handicapping is an approach that determines the outcome based on a combination of the player's performance and their skill level or physical ability, and is used with both able-bodied players (e.g., in golf [22]) and people with disabilities (e.g., the system of the Paralympic Games [12], where athletes with disabilities are classified based on the kind and severity of disability, and scoring parameters are adjusted). Likewise, balancing has been applied in video games to support player of varying skill. The most common way of adjusting games to the skills and abilities of the players are different levels of difficulty settings, based on predefined difficulty parameters, determined via extensive testing. However, such difficulty levels frequently prove too coarse. *Dynamic difficulty adjustments* (DDA) can foster more personalized and subtle balancing. In DDA, adjustments to the gameplay variables happen in real time based on the player's performance [1]. The physical nature of motion-based games adds another level of complexity to balancing: Sinclair et al. show that games do not only have to account for different skill levels of players in relation to mastering in-game challenges, but also in relation to their physical skills and abilities [20].

Balancing Multiplayer Experiences

In addition to player balancing to adapt in-game challenges to the skill level of an individual player, balancing has been applied in multiplayer settings to account for differences in skills and abilities between players, a process that affects the aspect of perceived fairness. Strategies for multiplayer game balancing summarized by Bateman et al. include asymmetric roles, matchmaking, and individual difficulty adjustment [2]. The authors examined aim assist techniques to balance shooting-gallery games played in a co-located setting using the Nintendo Wii Remote [2]. Their results show that players generally accept skill balancing in group settings, and that the player experience of less skilled players can be improved by providing assistance. Multiplayer game balancing in exertion-based games is complicated by the need to consider both player skill and fitness. Although the game in [2] was played with a motion-

[1] http://www.merriam-webster.com/dictionary/skills
[2] http://www.merriam-webster.com/dictionary/ability

based controller, the physical effort of pointing input does not compare to that required by many exergames. Different approaches for balancing exertion games have emerged in literature: The game Life is a Village [23] implements asymmetric player roles (cycling on a recumbent bike and using the Nintendo Wii Remote for gestural input), whereas Heart Burn [21] uses difference from target heart rates to balance for people with difference fitness levels. Likewise, Mueller et al. [15] demonstrated that heart rate balancing contributes to player experience in a remote jogging application, because it allows players to focus on their own fitness levels while still engaging with another person.

We present a game that is played with sedentary or motion-based controls in either unbalanced or balanced settings to examine player performance, experience, and self-esteem.

A SYSTEM TO STUDY PLAYER BALANCING IN MOTION-BASED GAMES

To study the effects of balancing in motion-based play, we implemented a dancing game similar to Dance Dance Revolution – a popular motion-based game [11] – that is playable by able-bodied persons (using dance mats) and persons with mobility disabilities (wheelchair input). We chose this game for two reasons. First, the game draws much of its appeal from symmetric player roles where competing players are required to complete similar challenges; therefore, it is interesting to study balancing. Second, performance in the game is determined by a combination of the player's level of coordination and their physical ability; therefore it is interesting for balancing a broad range of intra-dyad differences in skill and ability.

Gameplay

The gameplay consists of performing steps synchronously to the beat of a song. There are four different types of steps: up, down, left, and right. The game uses arrows that point in the direction of the steps to indicate which steps the user has to perform. The arrows appear at the top of the screen and move downward. A stationary set of targets (one for each type of step) is displayed near the bottom of the screen. The player aims to perform the move indicated by each arrow at the moment when the arrow is in line with the target. The player is awarded points based on how close the arrow was to the target at the moment they performed the indicated step. For the two-player competitive mode, the screen is divided in half, with each half having its own set of targets and arrows. We included dance music similar to the music used in commercially available dance games.

Game Input

The game was implemented in C# using Microsoft XNA Game Studio 4.0 and the Microsoft Kinect SDK alongside the KINECT[Wheels] toolkit [7] for wheelchair input.

The game includes input modes for motion-based (dance mat or wheelchair) and sedentary (game pad) play.

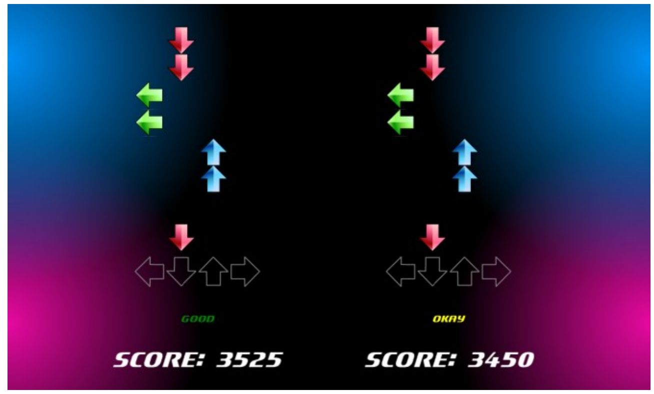

Figure 1. Screen of the two-player dancing game we created.

Dance Mat. The dance mat is a 1 square meter pad, divided into a 3-by-3 grid (see Figure 2). Players stand in the center cell of the grid, and perform steps by stepping on cells with their feet. The cell in front of the player corresponds to up arrows, the cell behind the player corresponds to down arrows, and the cells to the left and to the right of the player indicate left and right arrows, respectively.

Game Pad. The sedentary input mode uses the four action buttons (arranged in a diamond pattern) on a standard game controller. Players press the top button to perform an up-step, the lower button to perform a down-step, and the left and right buttons to perform left- and right-steps.

Wheelchair. The wheelchair mode emulates the dancing approach of the dance mat version, by requiring players to move around with the wheelchair. The player performs an up-step by moving the wheelchair forward, a down-step by moving backward, and left and right steps by turning the wheelchair to the left and right, respectively. Wheelchair movements are captured by a Microsoft Kinect sensor.

Balancing Parameters and Approaches

We implemented three balancing approaches that varied in perceptibility and interaction with game's mechanics and adapted parameters addressing game accessibility guidelines by Yuan et al. [24]. Each balancing approach varies features three game parameters. (1) *Step count*. Each step is associated with a specific time in the song at which it must be performed and coincides with a beat. Nine step charts were created for each song – the first step chart contains a large number of steps, and each successive chart removes a constant number of steps from the previous chart. (2) *Hit interval*. Steps must coincide with the moment when the arrow hits the stationary target. The timing window around this moment can be adjusted, giving players more or less time to obtain full points for performing the correct move. (3) *Score multiplier*. The number of points awarded for steps of varying quality can be changed to increase or decrease the final score. By varying these parameters, we created three balancing approaches.

Input Balancing. Input balancing uses a different step chart for each player. In order to account for score discrepancies resulting from fewer total steps, the score multiplier is also adjusted so that both players have an equal chance of obtaining the same score. Hit intervals are kept constant. Input balancing is visible to players because the interface displays fewer falling arrows for one player.

Time Balancing. Time balancing adjusts the hit interval for each player. Step count and score multiplier are kept constant. There are four possible results: input is too late and will not be scored, it is rather late and players will receive "okay" as feedback, it is slightly late and will be rated as "good", it is right on time and can be classified as "perfect". In the time balancing condition, we extended the time interval during which input would be recognized as "perfect" to give the weaker player an advantage, but we did not adapt the overall interval during which input was recognized. Time balancing is most hidden from players.

Score Balancing. Score balancing sets individual score multipliers for each player. Step count and hit interval are kept constant. Balancing is hidden from players in the sense that in-game challenge appears to be similar. However, it may become apparent when players compare their scores.

STUDY 1: BALANCING FOR ABLE-BODIED PLAYERS

In the first study, we examined the effectiveness of the balancing approaches, and explored how player balancing affected player experience, self-esteem, and relatedness.

Balancing Conditions

In addition to the input, time, and score balancing approaches described previously, we also included a control condition, i.e., no balancing, and a sedentary condition, i.e., no balancing, played with a game pad.

Participants and Procedure

The experiment was conducted in a university laboratory; the game was presented on a 42 inch flat-screen. 34 students (24 female, mean age=25.24 years, SD=4.79) participated in the study. 58.7% of the participants indicated that they play games at least a few times per month; 47% were experienced with the gamepad, 55% with a dance mat. Dyads were composed of friends and acquaintances.

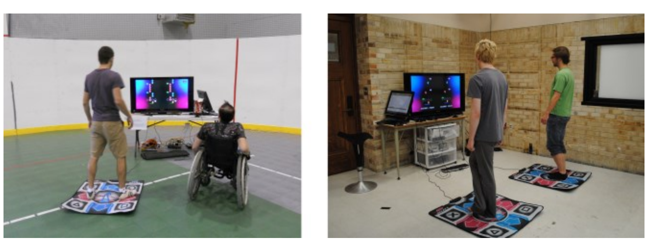

Figure 2. Study setup for able-bodied players and players in wheelchairs (left) and able-bodied dyads (right).

After providing informed consent, one participant completed an introductory questionnaire about their current affective state, while the other participant completed an ability pretest in a separate room. They then switched roles. The ability pretest was comprised of a training song followed by an assessment sequence of increasing difficulty and speed, and was scored by the number of completed moves. The score differential between the two players that resulted from this ability test determined the amount of balancing applied in the main part of the experiment. In addition, the score differential was used to set group values for step count in the control condition. Participants were not informed about their performance relative to their partner. During the main part of the experiment, participants played the game together in each of the five balancing conditions. Each condition was comprised of three songs (consistent across conditions and participating dyads) resulting in six minutes of play per condition. The order of presentation of the motion-based conditions was balanced using a Latin square. The sedentary condition was presented first for half of the participants and last for the other half, to avoid switching controllers multiple times during the experiment. After each condition the participants completed questionnaires assessing their affective state, experience with the game, their self-esteem, and their perception of the balancing techniques. At the end of the experiment, participants completed a demographic questionnaire, which included questions about sex, age, and game experience.

Measures

We collected dependent measures using in-game logging and validated scales. Cronbach's-α for all constructs by conditions was acceptable ($\alpha > .693$), except for relatedness in score balancing which was unacceptable ($\alpha > .392$).

Performance Logging. The system logged a variety of performance measures. In our analysis, we discuss the following values: *step count*, i.e., the number of correct steps players made during gameplay, and the information on the *overall score*, i.e., the points a player achieved regardless of their step count. Overall score is used to determine the winner of the game.

Player Experience Measures. We collected a variety of measures to assess player experience. These are all validated scales that are common for assessing play experience (e.g., [3, 19]). Please see [3] for details on the scales used. We assessed emotional valence using the 5-point Positive Affect Negative Affect Schedule (PANAS) [25]. To deconstruct player experience into its underlying constructs, we employed the 5-point Player Experience of Need Satisfaction Scale (PENS) [19], which assesses five constructs: *Competence*, *Autonomy*, *Relatedness*, *Immersion*, and *Intuitive Control.*

Player Self-Perception. To assess self-esteem after play we used the Rosenberg Self-Esteem Scale (RSE) [17], which has been used in games user research before [13]. In the RSE participants are asked to rate on their agreement on ten statements using a 4-point scale ranging from 1 (strongly disagree) to 4 (strongly agree), e.g., "On the whole, I am satisfied with myself".

Data Analyses

We categorized each member of the dyad as either the weaker or stronger player, depending on his or her ability pretest score. Player experience data were analyzed per player using RM-ANOVA with balance type (input, score, time, control) as the within-subjects factor and player strength (weaker, stronger) as a between-subjects factor; performance data were analyzed per group using RM-ANOVA with balance type as the within-subjects factor. All parametric tests were performed after validating the data for assumptions of ANOVA use. Degrees of freedom were corrected using the Huynh-Feldt method, if the assumption of sphericity was not satisfied. Pairwise comparisons used the Bonferroni method. Perception data were analyzed with non-parametric tests (as described in the results section) separately for the weaker and stronger players. Significance was set at α=0.05. The sedentary condition was analyzed separately.

Results

We present results for performance, player experience and then player perceptions of game balancing.

Table 1. Means (SD) for player experience and performance data by player strength. **Values are in thousands of points*

	Weaker Players				
	Input	Score	Time	Unbalanced	Sedentary
	mean *(SD)*	mean *(SD)*	mean *(SD)*	mean *(SD)*	mean *(SD)*
Self-Esteem	29.88(2.83)	32.76(3.61)	31.76(3.93)	31.18(4.93)	30.35(4.43)
Competence	3.41(0.79)	3.49(0.94)	3.65(0.84)	3.55(0.93)	3.59(0.86)
Autonomy	3.04(0.85)	3.02(0.92)	2.96(0.89)	2.76(0.96)	3.18(0.57)
Relatedness	3.25(0.86)	3.33(0.63)	3.27(0.6)	3.25(0.66)	3.27(0.58)
Immersion	2.85(0.9)	2.91(0.88)	2.84(0.86)	2.84(0.91)	2.96(0.67)
Control	3.59(0.89)	3.71(0.83)	3.63(0.87)	3.65(0.98)	3.39(0.81)
Neg. Affect	1.35(0.37)	1.42(0.45)	1.46(0.46)	1.44(0.46)	1.63(0.53)
Pos. Affect	3.35(0.77)	3.46(0.73)	3.19(0.87)	3.19(0.93)	3.22(0.66)
Score* (m)	11.6(2.5)	12.2(2.8)	9.7(3.1)	7.7(3.3)	10.1(2.5)
Hits (m)	106.7 (36.5)	111.7 (37.7)	118.9 (32.3)	109.1 (37.4)	139.7 (20.8)
	Stronger Player				
Self-Esteem	29.94(4.15)	31(5.92)	31.71(5.68)	31.76(6.13)	32(4.46)
Competence	3.67(0.79)	3.78(0.68)	3.82(0.71)	3.88(0.6)	4.14(0.5)
Autonomy	3.27(0.86)	3.33(0.75)	3.35(0.79)	3.31(0.78)	3.35(0.84)
Relatedness	3.27(0.78)	3.61(0.62)	3.43(0.72)	3.45(0.73)	3.45(0.7)
Immersion	3.19(0.69)	3.05(0.69)	3.06(0.72)	3.09(0.78)	3.05(0.82)
Control	3.73(0.71)	3.84(0.59)	3.94(0.6)	3.98(0.64)	4.06(0.79)
Neg. Affect	1.31(0.32)	1.38(0.46)	1.21(0.24)	1.42(0.57)	1.18(0.2)
Pos. Affect	3.03(0.94)	3.02(0.85)	3.14(0.86)	3.28(0.86)	3.19(0.89)
Score* (m)	12.6(2.5)	11.8(2.5)	11.0(2.8)	9.1(3.6)	11.5(1.8)
Hits (m)	132.3(41.8)	118.8(34.2)	123.4(30.8)	119.6(34.7)	150.2(10.6)

Performance and Outcome

Did balancing approach affect performance? A RM-ANOVA on the score differential (calculated as the difference between the sum of each player's scores over the three songs in a single balancing condition) and the step count differential (calculated in a similar fashion for step count) revealed that balancing type changed both score differential ($F_{1.94,\ 31.11}$=4.50, p=.020, η^2=.22) and step count differential ($F_{1.69,27.1}$=11.61, p≈.000, η^2=.42). Pairwise comparisons show that input balancing – which balanced the number of arrows given to participants – yielded a greater step count differential than the other three approaches (score, p=.012; time, p=.009; control, p=.018). In terms of the resulting score differential, time balancing resulted in a lower score differential than input (p=.011) or score (p=.033) balancing. There were no other differences.

Did the balancing approach change the outcome (win/lose)? We counted the number of songs won by each player for each balancing type (range: 0-3). A Friedman test showed that the number of songs won by the weaker player was different across the balancing types (χ^2_3=16.7, p=.001). Pairwise Wilcoxon tests showed that a greater number of games were won by the weaker player with score balancing (30/51) than input (15/51, p=.002), time (17/51, p=.008), or no balancing (14/51, p=.008).

As Figure 3 shows, compared to the control (no balancing), the time balancing resulted in more wins by the weaker player (more blue bars), but closer games (more lower bars); however, the score approach resulted in many wins by the weaker player with the weaker player greatly outscoring the stronger.

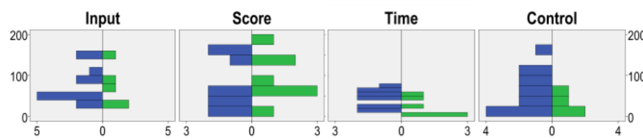

Figure 3. Histograms for score differential (y-axis x10²). Green bars are wins by weaker player; blue bars are wins by stronger players. Lower bars show a smaller score differential and higher bars show a higher score differential

Player Self-Esteem

Did balancing type affect post-play self-esteem? There was a main effect of balance type on self-esteem ($F_{3,96}$=9.10, p≈.000, η^2=.22). Pairwise comparisons showed that input balancing resulted in lower ratings of self-esteem than score (p≈.000), time (p=.001), or no (p=.018) balancing. However, a significant interaction with player strength ($F_{3,96}$=2.87, p=.040, η^2=.08) revealed that the input approach resulted in lower self-esteem than score (p≈.000) or time (p=.021) approaches for weaker players but only the time (p=.036) approach for stronger players (see Table 1).

Player Experience

Did balance approach affect the player experience? There was a main effect on relatedness ($F_{3,96}$=2.70, p=.050, η^2=.08). Pairwise tests showed that score balancing resulted in slightly higher relatedness than input balancing, although this comparison failed to reach statistical significance (Bonferroni-corrected p=.052). There were no other significant main effects of balance type or player strength.

There was a significant balance type by player strength interaction on increase in positive affect ($F_{3,96}$=4.10,

$p=.009$, $\eta^2=.11$) showing that score balancing resulted in a bigger increase in positive affect than time balancing for weaker players ($p=.037$), but not stronger players ($p\approx1.00$).

Player Perception of Balancing

Was balancing perceived? We asked players to report after each balancing condition on whether or not one of the players had been given assistance (for frequencies see Table 2). In general, players assumed that balancing had not been applied. Chi-squared tests show that this position was significant for weaker players for score balancing and no balancing and for stronger players in time balancing.

Table 2. Frequencies per balancing type. Grey rows: choice that balancing was *not* applied (/17). White rows: choice of who was stronger player (me, no difference, my partner; /17).

	Input			Score			Time			Control		
	N	χ^2	*p*	*N*	χ^2	*p*	*N*	χ^2	*p*	*N*	χ^2	*p*
Weaker	12	2.88	.090	13	4.77	.029	12	2.88	.090	14	7.12	.008
Stronger	11	1.47	.225	10	0.53	.467	14	7.12	.008	12	2.88	.090
Weaker	2,2,13	14.2	.001	4,5,8	1.53	.465	3,4,10	5.06	.080	2,2,12	12.5	.002
Stronger	12,0,5	2.88	.090	6,6,5	1.53	.943	8,5,4	1.53	.465	13,1,3	14.6	.001

How did players perceive their relative performance? After each balancing condition, we asked players to choose who was the stronger player: themselves, their partner, or neither. Chi-squared tests showed that weaker players thought their partner was stronger when input balancing or no balancing was used. Stronger players thought that they were the better player with no balancing (see Table 2).

How did players feel about beating or losing to their partner? We asked players to rate their agreement with the statement "I tried my best to beat the other player" and Competing against the other person made me feel guilty". Friedman tests showed no differences in both questions for either the weaker or stronger players (see Table 3). Interview results show that players were comfortable competing against their friends. When reflecting upon situations in which they would feel uncomfortable winning and consider letting another person win, many players referred to young children, a player group with largely different abilities. Additionally, player statements suggest that dyads usually were aware of each other's physical abilities, and had expectations of who would win a game.

Table 3. Means (SD) for agreement on a 5-point scale (1=strongly disagree, 5=strongly agree), and Friedman tests.

		Input	Score	Time	Control	Friedman
Tried Best	Weaker	3.8 (.73)	3.7 (.92)	3.8 (1.1)	3.6 (1.0)	$\chi^2=3.00$, p=.392
	Stronger	3.7 (.77)	3.7 (.92)	4.0 (.61)	3.8 (.88)	$\chi^2=3.84$, p=.279
Feel Guilty	Weaker	1.9 (.86)	1.7 (.69)	1.8 (.81)	1.6 (.62)	$\chi^2=3.44$, p=.329
	Stronger	2.1 (.75)	1.8 (.64)	1.8 (.75)	2.1 (1.2)	$\chi^2=2.35$, p=.503

Results for Balancing the Sedentary Game

Performance. The ability pretest was based on the motion-based version of the game, so we first looked to see whether there were performance differences between the control and sedentary conditions. Paired-samples t-tests showed no difference in the score differential ($t_{16}=1.46$, $p=.886$) or number of songs won by the weaker player ($t_{16}=0$, $p=1.00$).

Self-Esteem and Player Experience. We conducted a RM-ANOVA on all experience measures with player strength as a between-subjects factor. There was no effect of controller ($F_{1,32}=0.15$, $p=.697$), or interaction of controller and player strength ($F_{1,32}=0.50$, $p=.484$) on self-esteem. There was an interaction of controller and player strength on negative affect increase ($F_{1,32}=7.5$, $p=.010$, $\eta^2=.19$), showing that the weaker players had negative affect drop more in the motion-based condition than the sedentary condition, whereas the stronger players displayed the opposite pattern. There were no other differences.

Perception of Balancing. When asked to report on whether assistance was provided to a player, weaker players chose 'no' a significant percentage of the time in both the control (14/17, $\chi^2_1=7.12$, $p=.008$) and sedentary (14/17, $\chi^2_1=7.12$, $p=.008$) conditions, whereas stronger players chose 'no' more in the sedentary (16/17, $\chi^2_1=13.2$, $p=.000$) condition, but not significantly so in the control (12/17, $\chi^2_1=2.9$, $p=.090$) condition. When asked to report on who was the stronger player, weaker players chose their partner significantly more in both (control: 12/17, sedentary: 11/17) conditions, and stronger players chose themselves significantly more in both (control: 13/17, sedentary: 13/17) conditions. Chi-squared tests show that these choices are all significant (all $p<0.029$). When asked to report on whether they tried their best to beat the other player, controller made no difference in the ratings for weaker ($\chi^2=1.6$, $p=.107$) or stronger ($\chi^2=0.5$, $p=.603$) players. When asked to report on whether they felt guilty competing against the other player, controller made no difference for the stronger players ($\chi^2=0.9$, $p=.366$); however, weaker players felt less guilty in the motion-based game (mean=1.59, SD=0.62) than in the sedentary game (mean=2.06, SD=0.66) ($\chi^2=2.3$, $p=.021$).

Summary for Sedentary Interaction

The use of motion-based or sedentary controls made little difference to how balancing affected play. It is interesting that weaker players felt less guilty competing against their partner when using the motion-based controls (where they were at a disadvantage) than the sedentary controls (where any disadvantage was not related to physical ability).

Summary of Findings for Study 1

Our experiment showed a variety of differences between the balancing approaches. **Input balancing** best balanced step count, which is not surprising because fewer arrows were presented, thus it was more likely that the arrows would be hit. However, input balancing was a visible approach, making the difference in ability between the participants explicit, and it negatively affected the weaker participants' feelings of self-esteem and of their own performance compared to their partner's. This visible approach also negatively affected the stronger player in terms of reduced self-esteem and the dyad in general in terms of reduced feelings of relatedness (note that Cronbach's-α for score balancing relatedness was low). **Time balancing** best balanced the score differential (i.e., brought the score closest), but did not result in more games actually won by the weaker player. This was the least visible method and least perceived by stronger players. Time balancing also had good effects on self-esteem for both the weaker and stronger players. **Score balancing** resulted in more games won by the weaker player; however, score balancing may have overbalanced the game in that it not only resulted in more games won by the weaker player, but also the highest score differential, where the weaker player often outscored the stronger player. This resulted in higher self-esteem and higher positive affect for the weaker player, likely due in part to the fact that score balancing was the least perceived balancing approach for weaker players.

In Study 1, the players had different physical abilities in terms of fitness and ability to interact with the game; however, these differences were not drastic. Both participants were generally healthy individuals able to participate in the game. Had one member of the dyad been physically unfit due to injury, health, or general wellness, the effects of balancing may have been different. To investigate the situation where two members of a dyad playing an exergame have vastly different physical abilities, we conducted a second study where one member of the dyad had a mobility disability and was using a wheelchair.

STUDY 2: BALANCING ABLE-BODIED PLAYERS AND PLAYERS WITH MOBILITY DISABILITIES

In the second study, we investigate whether motion-based games can be balanced in a way that provides an enjoyable experience for people with extreme differences in physical ability. Based on the findings from Study 1, we created a new version of our game that combines the hidden balancing strategies (time balancing and score balancing).

Participants, Procedure, and Methods

This study was conducted in conjunction with the Canadian Paraplegic Association's wheelchair relay, an annual family sports event. Four dyads of people using wheelchairs and an able-bodied friend (mean age=34, 5 male) participated. All play video games; 4 had dance mat experience. After providing informed consent (one dyad consented to use only their interview data), participants completed a questionnaire about their affective state and self-esteem, followed by the ability pretest. Participants then played two songs in the game using the combined balancing approach, followed by a post-play questionnaire. Observations and interviews provided qualitative data. Measures were the same as in Study 1, except that emotion was assessed using the shorter Self Assessment Manikin (SAM) [4].

Results

Because of the small number of participants, we do not make statistical comparisons between the groups.

Playability

Was the game suitable for people using wheelchairs? Yes. Players using wheelchairs scored between 5662 and 8163 points per song. Observations suggested that wheelchair game controls were intuitive. We observed a wide range of interaction styles, with some players moving the wheelchair in a slow, careful way, and others trying to keep up with the pacing of the game. Feedback suggests that the type of assistive device may have influenced the way players interacted, *"if I had my sports wheelchair, I would just have to move my arms for the side turns."* (G1).

Are there performance differences between persons using wheelchairs and able-bodied players? Although players using wheelchairs were able to participate, their able-bodied opponent generally outscored them. We compared results to the mean score (10600) from Study 1. One-sample t-tests show that players using a wheelchair had a lower score (mean=7061, SE=737, t_2=4.80, p=.041) than the able-bodied participants in Study 1, but the scores for able-bodied players show no differences (mean=10292, SE=2324, t_2=0.13, p=.907). In one pair, the players scored the same number of points in the pre-test, so no balancing was applied.

Player Experience

Are there any differences in the experience of persons using wheelchairs and able-bodied players? We performed one-sample t-tests separately for able-bodied players and participants using wheelchairs against the neutral rating for PENS data (3), see Figure 4. The results yield no significant differences for able-bodied players, but wheelchair users showed heightened satisfaction of autonomy (t_2=8, p=.008), relatedness (t_2=3.464, p=.037), enjoyment (t_2=8.5, p=.007), and competence (mean=4.0, t not computed because SD=0). This is supported by statements showing that players enjoyed how the game integrated the wheelchair, with one stating that *"it is nice to see my wheelchair in the game instead of being an object that stands between me and the world"* (G2, person using wheelchair).

How did players feel about competing against a person with different physical abilities? Player feedback on competing against a person with radically different abilities suggests that balanced competition is not only beneficial for the experience of the player with a mobility disability, but also affects the able-bodied participant. Players using wheelchairs generally did not seem to mind losing to the able-bodied person, stating that *"[Competing against an able-bodied person is] fine, doesn't really bother me cause it was fun so it was cool beans"* (G2, person in wheelchair), and that *"I didn't care if I was losing or winning, I never played the game before"* (G4, person using wheelchair). In contrast, only one able-bodied person stated that she felt *"comfortable competing with a person in a wheelchair."* (G1, able-bodied person). Other able-bodied participants were concerned about the fairness of competition: *"I don't think it is fair. We don't seem to be on an equal playing*

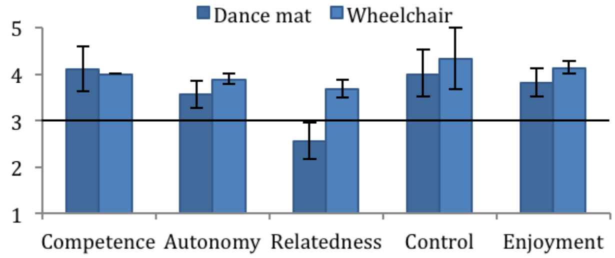

Figure 4. Mean (±SE) scores for Study 2 by input device. The line represents a neutral response.

field." (G4, able-bodied person), and the impact of their physical abilities, claiming that they "*felt bad that [the player using the wheelchair] did such a good job and I still beat him*" (G2, able-bodied person). This reinforces comments from Study 1 where participants mentioned that beating certain groups of people that differed radically in ability (e.g., young children) would not be acceptable. The expected level of physical abilities plays a role in the experience of the stronger participant; however, it is likely that an experienced player who uses a wheelchair could be stronger than their able-bodied opponent using a dance mat.

Interestingly, one able-bodied participant (G4) who claimed that she felt uncomfortable winning against her partner in the motion-based game pointed out that competition would be *"completely fair if we were playing on a hand-held"*, suggesting that skill differences as a result of practice are acceptable, but differences in physical abilities as a result of a disability that lead to different in-game scores are not. Finally, one participant pointed out that she had lower expectations of winning, stating that she *"didn't expect to do really well, he's an athlete"* (G3, person using wheelchair). In addition, she noted that she'd been more competitive with another person in wheelchair, saying that if she *"was competing with someone else in a chair, I may have felt bad losing"*, suggesting that she does not perceive an able-bodied person a contender she could beat.

Summary of Findings for Study 2

Our game was accessible to people using wheelchairs, and they enjoyed playing. Experience ratings and feedback show that people using wheelchairs enjoy the integration of their assistive device into the game. However, performance differences within dyads show that balancing needs to be increased to bridge score gaps. Although players using wheelchairs did not mind losing to the able-bodied player, able-bodied persons felt uncomfortable beating a person with a disability; therefore, balancing is not only important to empower the weaker player, but also to improve the experience of the stronger player.

DISCUSSION

The work presented in this paper investigates balancing strategies for co-located motion-based games with a focus on the effects of player performance, experience, and self-esteem. In Study 1, we showed differences in player experience depending on whether the balancing strategies were explicit or hidden; in Study 2, we included people with mobility disabilities in competition, and outlined how player experience is affected by extreme differences in player abilities. In this section, we discuss our findings in the context of game design, we tie them back to findings from sports psychology, and we outline how they can be applied to design better motion-based game experiences for able-bodied people and people using wheelchairs.

Effective Balancing for Motion-Based Video Games

In this section we discuss our findings with a focus on balancing strategies for co-located motion-based games.

Choosing Suitable Balancing Approaches

Results of our study suggest that explicit input balancing – the kind of balancing commonly applied in commercially available dance games such as *Dance Dance Revolution* or *Dance Central* – poses the biggest risk in terms of player experience because it is clearly visible to both players. In contrast, hidden strategies offer the opportunity of subtly mitigating skill differences between players. Our work shows that score balancing is a suitable means of closing extreme performance gaps between players, whereas time balancing can be applied to reduce smaller differences and to account for the impact of asymmetric physical input (e.g., if one player uses a wheelchair).

Accommodating Extreme Ability Differences

Large score gaps in Study 2–despite the combined balancing approach–suggest that optimization is necessary to enable people with extreme ability difference to compete. Our observations showed that the characteristics of wheelchairs led to differences in input difficulty, causing some in-game sequences to be easier for the able-bodied player than for the player using the wheelchair (e.g., moving from a side turn into a forward motion). To optimize our approach, it is necessary to consider difficulty levels of input sequences as balancing factor: By adapting transitions between movements in a way that is suitable for players using wheelchairs (e.g., giving more time between turns), it can be adapted to their needs without affecting the game experience of able-bodied players.

The Effects of Player Differences in Skills and Abilities in Co-Located Motion-Based Games

In this section, we discuss the effects of skills, abilities and balancing on player experience and self-esteem.

Explaining the Effects of Balancing on Player Experience

The results of our studies suggest that balancing strategies for co-located motion-based games do not only affect the in-game performance of the weaker player, but extend into the experience of both members of the dyad. Locus of control theory [18] suggests that the attribution of failure and success plays a major role in how people perceive themselves. In the context of game balancing, this means that adaptation should allow players to attribute their performance to the relationship between task difficulty and their own ability. This is difficult if players are aware of game balancing: with explicit balancing, the visibility of the adaptation is likely to hinder the attribution of game success to the weaker player's abilities. Likewise, if difficulty is drastically lowered in hidden balancing scenarios, weaker players might attribute a win to a source outside of their control. At the same time, the stronger player can attribute a loss to the software support of the weaker player. If the stronger player is able to beat a supported player, additional challenge might increase the value of a win, while also introducing guilt. Our results also show that the weaker player has a negative experience when aware of assistance, yet unable to beat the stronger player. Hindering the stronger player introduces a similar problem by allowing for outcome attribution to his or her lowered ability.

Explaining the Effects of Balancing on Self Esteem

The results of Study 1 suggest that players are aware of skill differences between themselves and the other player, that they have an intuitive understanding of their own physical abilities, and that they are able to perceive game balancing under certain circumstances. People constantly compare themselves with others to estimate their own abilities in a social context [14]. The effect of balancing mechanisms is intertwined with the process of social comparison by artificially equalizing skill differences and making an initial estimate meaningless. In Study 1, self-esteem is conceptualized as a state variable, and measured depending on balancing approach. The underlying assumption is that people update themselves in comparison with a self-standard: if actual performance differs too much from expected performance, self-esteem and well-being are threatened. This explains how balancing mechanisms affect self-esteem and player experience of interdependent dyads by narrowing a natural gap of ability.

Understanding the Effects of Perceived Ability Differences

Our results show that perceived differences in player abilities and expectations regarding player performance play a role in the experience of both weaker and stronger players. This is particularly interesting in the context of balancing dyads with extreme ability differences. When comparing Study 1 and Study 2, there seems to be a shift in player expectations and experience: Our results suggest that persons with disabilities have lower expectations regarding their performance, leading to a positive experience even when losing. Able-bodied players perceive the same ability differences, and report negative emotions despite winning against the person using a wheelchair. We believe that this difference in the evaluation of the outcome of competition may be caused by the extreme difference in abilities and general perceptions of persons with disabilities: if, in the eyes of the able-bodied person, the player using the wheelchair never had a chance of winning, this reduces the value of their victory and may induce feelings of guilt.

Ability Differences and the Risk of Overbalancing

The previous sections have addressed the effects of balancing on player experience and self-esteem, and how perceived ability differences affect competition if performance gaps cannot be closed in a suitable way. However, designers also need to consider the risk of overbalancing. If games are overbalanced and give a large advantage to the weaker player, this may affect the player experience of the stronger participant (as demonstrated by results for score balancing in Study 1). Interestingly, this is in line with discussions related to the participation of persons with disabilities in sports tournaments directed towards able-bodied participants, where some athletes claimed that certain assistive devices lead to an unfair advantage [6]. This suggests that balancing intensity needs to be adapted carefully to strike an acceptable balance between enabling the weaker player to win while allowing the stronger player to experience in-game competence. From a cognitive perspective, it is important that the gap closed by adaptation is reasonable and accepted by both players; only then it is possible for players to develop positive relationships and have rewarding experiences.

LIMITATIONS AND FUTURE WORK

There are several limitations and opportunities for future work. To study the long-term effects of ability differences and balancing strategies on player experience – particularly when able-bodied persons and persons with disabilities compete against each other – a longitudinal study would be necessary. Additionally, participants in Study 1 were asked to bring a friend. In a follow-up, it would be interesting to control for the kind of relationship to investigate the effect, e.g., if romantic partners or parents and children play together. Studying a broader range of able-bodied players and players using wheelchairs would provide insights into how physical abilities affect competition and experience. Studying the effects of disabilities beyond the use of manual wheelchairs would be valuable, e.g., whether users of power wheelchairs could play the game, either in a motion-based or in a sedentary setting using the joystick of their wheelchair as input device. Finally, there are aspects related to the impact of player performance on player experience that have not yet been explored. For instance, facing challenges (e.g., competing with a stronger opponent) may at first result in negative affect, but release a strong positive influence once the challenge is overcome. Such time-dynamics are very hard to predict but might play an important role and have not yet been investigated.

CONCLUSION

Developing suitable balancing strategies for co-located motion-based games is an important step to mitigate negative effects of player performance on player experience and self-esteem, enabling players of different abilities to engage in playful competition. In this paper, we demonstrate the effects of explicit and hidden balancing strategies, and we provide insights into the application of balancing strategies to bridge extreme ability gaps between players. By keeping motion-based game controls accessible for persons with different levels of physical abilities – including people with wheelchairs – co-located motion-based games can be a means of motivating physical activity among broad audiences. Additionally, such games may have other benefits: joint play between able-bodied persons and persons with disabilities can help provide empowering experiences for both players, and contribute to the development of inclusive motion-based games.

ACKNOWLEDGMENTS

We thank the Canadian Paraplegic Association for supporting our research, and the participants for their contribution during the evaluation phase. We thank all of our CHI reviewers for their helpful comments and NSERC and the GRAND NCE for funding.

REFERENCES

1. Adams, E. *Fundamentals of Game Design*. Berkeley: New Riders, 2010.
2. Bateman, S., Mandryk, R.L., Stach, T., and Gutwin, C. Target Assistance for Subtly Balancing Competitive Play. In: *Proc. of CHI 2011*, ACM (2011), 2355-2364.
3. Birk, M. & Mandryk, R.L. Control Your Game-Self: Effects of Controller Type on Enjoyment, Motivation, and Personality in Game. In: *Proc. of CHI 2013*, ACM (2013), 685-694.
4. Bradley, M.M. and Lang, P.J. Measuring emotion: the self assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49–59.
5. Cairney, J., Kwan, M.Y.W., Veldhuizen, S., Hay, J., Bray, S.R., and Faught, B.E. Gender, perceived competence and the enjoyment of physical education in children: a longitudinal examination. *Int Journal of Behavioral Nutrition and Physical Activity* 9, 26 (2012).
6. Epstein, D. Fair or foul? Experts split over whether Pistorius has advantage. Available at http://sportsillustrated.cnn.com/2012/olympics/2012/writers/david_epstein/08/03/oscar-pistorius-london-olympics/index.html, last access: 10/09/13.
7. Gerling, K.M., Kalyn, M.R., and Mandryk, R.L. KINECTWheels: Wheelchair-Accessible Motion-Based Game Interaction. In: *EA of CHI 2013*, ACM (2013).
8. Gerling, K.M., Mandryk, R.L., and Kalyn, M.R. Wheelchair-Based Game Design for Older Adults. In: *Proc. of ASSETS 2013*, ACM (2013).
9. Goni, A. and Zulaika, L. Relationships Between Physical Education Classes and the Enhancement of Fifth Grade Pupils' Self-Concept. *Perceptual and Motor Skills* 91, (2000), 246-250.
10. Hernandez, H.A., Ye, Z., Graham, T.C.N., Fehlings, D., and Switzer, L. Designing Action-based Exergames for Children with Cerebral Palsy. *Proc. of CHI 2013*, ACM (2013).
11. Höysniemi, J. International Survey on the Dance Dance Revolution Game. *Comp in Entertainment* 4, 2 (2006).
12. IPC. *Paralympic Classification*. Available at http://www.paralympic.org/Classification/Introduction/, last access: 21/08/2013.
13. Jung, Y., Li, K. J., Janissa, N. S., Gladys, W. L. C., & Lee, K. M. Games for a better life: effects of playing Wii games on the well-being of seniors in a long-term care facility. In: *Proc. of IE '09,* (2009), 5.
14. Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of personality and social psychology* 68, 3 (1995), 518.
15. Mueller, F., Vetere, F., Gibbs, M.R., Edge, D., Agamanolis, S., Sheridan, J.G., and Heer, J. Balancing Exertion Experiences. *Proc. of CHI 2012*, ACM (2012).
16. Ninot, G., Billard, J., and Delignières, D. Effects of integrated or segregated sport participation on the physical self for adolescents with intellectual disabilities. *Journal of Intellectual Disability Research* 49, 9 (2005), 682-689.
17. Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
18. Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *Journal of consulting and clinical psychology* 43, 1 (1975), 56.
19. Ryan, R. M., Rigby, C. S., Przybylski, A. K. Motivational pull of video games: A self-determination theory approach. *Motivation and Emotion* 30, (2006).
20. Sinclair, J., Hingston, P., and Masek, M. Exergame development using the dual flow model. In: *Proc. of IE '09*, ACM (2009).
21. Stach, T., Graham, N., Yim, J., and Rhodes, R.E. Heart Rate Control of Exercise Video Games. In: *Proc. of GI 2009*, ACM (2009), 125-132.
22. Swartz, T.B. A New Handicapping System for Golf. *Journal of Quantitative Analysis in Sports* 5, 2 (2009).
23. Yim, J. and Graham, T.C.N. Using Games to Increase Exercise Motivation. *Proc. of FuturePlay 2007*, ACM (2007), 166-173.
24. Yuan, B., Folmer, E., Harris Jr., F.C. Game Accessibility: a Survey. *Univ Access Inf Soc* 10, 1 (2011), 81-100.
25. Watson, D., & Clark, L. A. *The PANAS-X: Manual for the positive and negative affect schedule-Expanded Form*. Iowa City: University of Iowa (1994).