THE USE OF NUDGES IN EXERGAMES TO MODERATE PLAYER EXERTION

by

Adrian L. Jessup Schneider

A thesis submitted to the School of Computing
In conformity with the requirements for
the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada
(May, 2016)

Copyright © Adrian L. Jessup Schneider, 2016
Abstract

Exergames are digital games with a physical exertion component. Exergames can help motivate fitness in people not inclined toward exercise. However, players of exergames sometimes over-exert, risking adverse health effects. These players must be told to slow down, but doing so may distract them from gameplay and diminish their desire to keep exercising.

In this thesis we apply the concept of nudges—indirect suggestions that gently push people toward a desired behaviour—to keeping exergame players from over-exerting. We describe the effective use of nudges through a set of four design principles: natural integration, comprehension, progression, and multiple channels.

We describe two exergames modified to use nudges to persuade players to slow down, and describe the studies evaluating the use of nudges in these games. PlaneGame shows that nudges can be as effective as an explicit textual display to control player over-exertion. Gekku Race demonstrates that nudges are not necessarily effective when players have a strong incentive to over-exert. However, Gekku Race also shows that, even in high-energy games, the power of nudges can be maintained by adding negative consequences to the nudges. We use the term "shove" to describe a nudge using negative consequences to increase its pressure.

We were concerned that making players slow down would damage their immersion—the feeling of being engaged with a game. However, testing showed no loss of immersion through the use of nudges to reduce exertion.

Players reported that the nudges and shoves motivated them to slow down when they were over-exerting, and fit naturally into the games.
Co-Authorship


- The study described in Chapter 5 is the subject of this paper.
Acknowledgements

I would like to thank my supervisor, Dr. Nicholas Graham, for mentoring me and assisting with my research. I would also like to thank Zi Ye and Dan Moran for creating the Gekku Race game I used in my second study, and Florian Eysseric, Hamilton Hernández, Luke Walker, and Liam Collins for creating and maintaining the FireANT software program used in both studies. I would like to thank my colleagues at the EQUIS lab for their participation in preliminary testing of my nudge feedback. I would like to thank my family and friends for their support throughout my degree, in particular my mother Sharon Jessup-Joyce for her work in proofreading this thesis.
# Table of Contents

Abstract ................................................................................................................................. ii
Co-Authorship ....................................................................................................................... iii
Acknowledgements ........................................................................................................ iv
List of Figures ....................................................................................................................... vii
List of Tables ......................................................................................................................... viii
Chapter 1 Introduction ....................................................................................................... 1
  1.1 Problem ......................................................................................................................... 2
  1.2 Solution .......................................................................................................................... 3
  1.2.1 Design Principles ...................................................................................................... 4
  1.3 Evaluation ...................................................................................................................... 6
  1.4 Contributions ................................................................................................................ 7
Chapter 2 Related Work ................................................................................................. 9
  2.1 Exergames ...................................................................................................................... 9
  2.1.1 Non-Commercial Exergames .................................................................................. 12
  2.2 Exertion ......................................................................................................................... 16
  2.3 Immersion ..................................................................................................................... 17
  2.4 Nudges .......................................................................................................................... 18
Chapter 3 Designing with Nudges .............................................................................. 21
  3.1 PlaneGame .................................................................................................................... 23
  3.1.1 Nudging in PlaneGame ........................................................................................... 25
  3.2 Gekku Race ................................................................................................................. 27
  3.2.1 Nudging and Shoving in Gekku Race ..................................................................... 27
  3.3 Design Summary .......................................................................................................... 30
Chapter 4 Design Principles ....................................................................................... 31
  4.1 Natural Integration ........................................................................................................ 32
  4.2 Comprehension ............................................................................................................. 33
  4.3 Progression ................................................................................................................... 34
  4.4 Multiple Channels ........................................................................................................ 35
  4.5 Iterative Design ........................................................................................................... 37
  4.6 Summary ....................................................................................................................... 38
Chapter 5 Proof-of-Concept Study: PlaneGame ...................................................... 39
  5.1 Study Design ................................................................................................................. 40
List of Figures

Figure 1: Dance Dance Revolution arcade cabinet ................................................................. 9
Figure 2: Wii Remote .................................................................................................................. 10
Figure 3: Wii Balance Board ................................................................................................... 11
Figure 4: Microsoft Kinect ...................................................................................................... 11
Figure 5: Namco Prop Cycle .................................................................................................. 12
Figure 6: Fisher-Price Smart Cycle .......................................................................................... 12
Figure 7: Growl Patrol Gameplay ............................................................................................ 13
Figure 8: Control configurations for Sportal ........................................................................... 14
Figure 9: Mandryk et al.'s visual overlays in the dirt-biking game "Nail'd" [54] ...................... 20
Figure 10: Logitech F710 controller used as input in PlaneGame and Gekku Race .............. 22
Figure 11: A player playing Gekku Race using the controller and exercise bicycle ............ 23
Figure 12: A screenshot of PlaneGame. The player’s plane avatar and an AI bird are both racing to capture the rainbow for points .................................................................................. 24
Figure 13: Nudges in PlaneGame .......................................................................................... 26
Figure 14: In Gekku Race, two players shoot cashews at each other while racing to the top of the track ......................................................................................................................... 28
Figure 15: Nudges and shove in Gekku Race ....................................................................... 29
Figure 16: Textual condition of PlaneGame. Target cadence at the top of screen, current cadence underneath plane ............................................................................................................. 42
Figure 17: Performance metrics for PlaneGame. Vertical bars show standard error. Horizontal hats indicate statistical significance at $\alpha = .05$ ........................................................................................................ 48
Figure 18: Questionnaire responses for PlaneGame. "Do you feel motivated to slow down when you’re pedaling too quickly?", rated on a 7-point Likert scale. Vertical bars show interquartile range. Horizontal hats indicate statistical significance at $\alpha = .05$ ........................................................................................................ 50
Figure 19: Performance metrics for Gekku Race. Vertical bars show standard error. Horizontal hats indicate statistical significance at $\alpha = .05$ ........................................................................................................ 59
Figure 20: Questionnaire responses for Gekku Race. Both were questions rated on a 5-point Likert scale. Vertical bars show interquartile range. Horizontal hats indicate statistical significance at $\alpha = .05$ ........................................................................................................ 61
Figure 21: Biri Brawl ............................................................................................................... 76
### List of Tables

Table 1: Latin square used for order balancing with *PlaneGame*. .................................................................45
Table 2: Pairwise t-tests for *PlaneGame* game score at 40rpm. ...........................................................................49
Table 3: Interview questions for *Gekku Race*. .................................................................................................56
Table 4: Pairwise t-tests for *PlaneGame*, 40 RPM. ...........................................................................................87
Table 5: Pairwise t-tests for *PlaneGame*, 60 RPM. ............................................................................................87
Table 6: Wilcoxon signed-rank tests for *PlaneGame*, 40RPM. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values. .................................................................88
Table 7: Wilcoxon signed-rank tests for *PlaneGame*, 60RPM. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values. .................................................................88
Table 8: Pairwise t-tests for *Gekku Race*. .........................................................................................................89
Table 9: Wilcoxon signed-rank tests for *Gekku Race*. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values. IEQ Quartiles are rounded to whole numbers. ...90
Chapter 1

Introduction

Exergames – video games with a physical exercise component – show promise for motivating fitness in people who would otherwise not be inclined toward exercise [27, 28, 33, 47]. To gain the benefits, however, exergame players must stay within a healthy range of effort, neither under- nor over-exerting themselves [2, 46]. In this thesis, we show that players' behaviour can be guided within the game by using indirect suggestions called "nudges". For example, to stop players from over-exerting in a cycling-based plane-racing game, the plane may visibly and audibly overheat. When players see the plane overheating, they naturally slacken their pedalling speed.

Sometimes, nudges are not enough. When people are particularly motivated to over-exert, nudges can be augmented with direct consequences for disobeying them, which we call "shoves". We show that shoves are effective for controlling player behaviour, but must be used carefully, as players can become frustrated at being forced to adapt.

We introduce principles for the design of nudge- and shove-based feedback systems, highlighting the value of using multiple channels and of taking advantage of players' knowledge of real-world behaviours and standard game mechanics.

We designed the nudge feedback to fit naturally into the game, expecting this natural integration to be closely linked to immersion. Immersion, the feeling of being engaged with a game, is an important part of video games according to many theories of game design [1, 53, 64]. However, while participants noted and expressed preference for the nudge-based natural integration during interviews, they showed no difference on questionnaires measuring immersion.
This suggests that natural integration is aesthetically preferred by players, but is not linked to their sense of immersion.

1.1 Problem

The foundational premise of exergames is that many people do not find exercise intrinsically motivating, but will often find playing video games to be compelling. For such people, video games that incorporate exercise can serve as an enjoyable way of being physically active [43, 48, 49].

The two sides of an exergame, the exercise and the game, need to support each other. If the player's perception of doing the exercise component as part of the game breaks down, players may become discouraged or bored. However, the opposite problem can sometimes occur: an exergame might motivate players to over-exert themselves, posing a risk to their physical health and failing to provide a well-balanced exercise session [27]. Such over-exertion might come from the players becoming especially stimulated by an exciting or difficult segment of gameplay, or by failing to go slowly enough during the warm-up or cool-down phase of an exercise session. In this thesis, we focus on this latter problem of over-exertion.

In theory, stopping players from over-exerting is simple: have the game’s user interface openly give them a message to slow down. However, having such a blatant reminder that the game is meant to get them to exercise instead of just have fun might reduce the value of embedding exercise into a game in the first place, risking the sense of fun of playing the game for its own sake. The challenge is how to prevent players from exceeding target levels of exertion without disrupting this sense of play.
1.2 Solution

We show that this can be accomplished through nudges: interface cues that gently push players in a desired direction, without explicitly instructing them to do so. This concept was originally described by Thaler and Sunstein in the book *Nudge* [55]. For example, serving food in smaller dishes increases a meal's apparent size and decreases consumption [60], while marking unsafe networks in red encourages computer users to use safe networks [14]. In this thesis we also demonstrate the use of shoves, less gentle nudges that carry consequences for disobeying them. While a nudge suggests, a shove insists, convincing players to slow down even if they previously ignored the consequence-free nudges.

To illustrate nudging for exertion, we have added nudge-based feedback techniques to two pedal-to-play exergames, *PlaneGame* and *Gekku Race*. Both feedback systems were designed to persuade players to slow down, while ensuring the feedback is naturally integrated into the game.

*PlaneGame* is a single-player racing game, in which the player takes control of an airplane and races against AI-controlled birds, who will steal rainbows if they reach them first. The speed of the plane is linked to the player's pace on a stationary exercise bicycle, so to beat the birds the player must pedal rapidly enough to outpace them. The game has an internal speed limit, after which the plane will no longer go any faster.

To persuade players to slow down when they exceed the speed limit, the plane begins to visibly and audibly malfunction as it overheats. Smoke and a rattling sound emerge from the plane's engine compartment, turning the edges of the screen grey. If the player continues to exceed the speed limit, the sound gets louder and the smoke thicker, until eventually the smoke is joined by fire and the screen begins to fade from grey to red.
*Gekku Race* is an existing multiplayer racing game [23, 30], in which players race as lizards climbing a wall to a finish line. The 'gekku' lizards are able to stun each other by shooting cashews from their mouths, allowing the attacking lizard to run ahead of its stunned opponent. Unlike *PlaneGame*’s plane, the lizards will continue to run faster even past the speed limit.

We enhanced *Gekku Race* to tell its players they're going too quickly by having the lizard become tired from the exertion. It begins panting audibly, and cartoon breath puffs from its mouth. The screen also begins fading to greyscale. If the player does not slow down, the lizard pants faster and the screen goes greyer.

Since the lizards continue to get faster as players' pace increases, players can get an advantage from ignoring the nudges and continuing to pedal quickly. We expected that at least some players would do just this, and so prepared shoves to make them slow down. If players pedal too quickly for too long, then the lizard not only pants for breath, but also intermittently collapses to the ground and gasps loudly for air. Since the lizard stops, it falls behind and the player cannot win the race without slowing down.

### 1.2.1 Design Principles

Using the concepts developed and employed in these designs, we also present a set of four design principles that can guide the creation of similar feedback to other games: integrate the feedback into the game world; ensure the feedback is easily understandable; intensify feedback incrementally to increase pressure to follow it; and use multiple channels of feedback to complement and supplement each other.

To keep the messages from being jarring to the player and risking the impression of playing for fun's sake, nudges can be integrated into game play. The feedback in both of our games has a core concept that is integrated into the game's internal fiction. In *PlaneGame*, the
plane starts breaking when it is pushed past its mechanical tolerances, and Gekku Race’s lizards get tired when they run too quickly.

Ensuring that feedback is comprehensible has two dimensions: clarity, and conspicuity. To be understood, the feedback must be noticed by the players as something significant, and be understood once they notice it. Feedback might not be conspicuous if its onset is gradual, so there ought to be at least one component that is clearly present when pedaling too quickly, but is otherwise absent. The clarity of feedback can be improved through a logical link between the effect and its cause, or by analogy to other games. The screen turns grey in both games, for example, resembling the tunnel vision caused by being severely short of breath; this effect is also used in several commercial games to indicate a threat to the character’s health, reinforcing the connection between a greying screen and an unhealthy speed.

Increasing feedback incrementally allows the game to increase the pressure on players who are not responding to the nudges. Increasing nudges progressively is also a good way to implement shoves. Accidentally drifting over the speed limit for a moment will not be punished, while players intentionally pushing the game can still be shoved back down. The progression also acts as a warning system, so players who already know the consequences of going too quickly can back off immediately.

Using multiple channels (for example, visual, auditory, and gameplay effects) allows nudges to combine. Having multiple channels can make the nudge feedback more robust against player distraction, by giving multiple things to be noticed; multiple channels allow specialized nudges, such as performance-affecting shoves, to complement each other. Combining channels helps satisfy other principles, such as increased conspicuity from several things happening at once.
1.3 Evaluation

To test the effectiveness of the nudges in the modified versions of PlaneGame and Gekku Race, we conducted a study using each game. The two studies shared two primary research questions:

- How effective are nudge techniques at convincing players not to over-exert?
- What effect do nudge techniques have on players’ immersion in the game?

In PlaneGame, we needed to establish a baseline for how effective nudge techniques are for controlling over-exertion. To do this, the version of the game with nudges was compared against a textual feedback version that conveyed speed information directly to the player, and a control condition that included neither nudge nor textual feedback. We also had participants play the game with two different speed limits, 40 rpm and 60 rpm. Pilot testers said 60 rpm was a comfortable pace, while the slower speed was used to see whether nudges could persuade players to pedal at an unpleasantly slow pace, if needed.

The expected result was that the textual condition would be most effective at preventing players from exceeding the speed limit, but that it would also cause the most loss in immersion. The control condition, we believed, would not control speed, but would also not cause a drop in immersion. We expected the nudge condition to be between the others on both axes.

The results of testing 24 participants showed that both the textual and nudge conditions were equally effective at persuading players to slow down, and we saw no significant differences in immersion between the conditions.

With the utility of nudges confirmed for slowing players, we conducted a second study with Gekku Race. In Gekku Race, players were able to continue speeding up even past the speed limit, so there was a tangible benefit to exceeding that limit. Gekku Race is also a multiplayer
game, and we suspected players would be more eager to win against a real opponent. We tested to see if nudges were enough to convince these players to slow down despite the advantages of ignoring them, or if stronger feedback in the form of shoves was needed.

We tested 20 participants, in 10 pairs, over three conditions: the full feedback that included shoves, a stripped version that included only the softer nudges, and a control condition with neither shoves nor nudges. We expected that the nudge condition would not be able to slow players, but that the shove condition would. We also expected that the forcefulness of the shove condition would cause a reduction in immersion.

Results showed that the shove condition succeeded in convincing players to slow down, and that the nudge condition did not. There was no measurable effect on immersion.

Together, these studies show that nudge techniques are useful for controlling players' exertion levels in exergames, whether alone (in some types of games) or in conjunction with shoves. We also saw that, despite our concerns, nudges caused no loss in immersion.

1.4 Contributions

The primary contributions of this thesis are:

- *Principles for the design of nudge-based feedback*: We provide a set of four design principles for designing or evaluating nudge-based feedback techniques: natural integration, comprehension, progression, and multiple channels.

- *Distinguishing between nudges and shoves*: We define the concept of a "shove" as a stronger nudge that punishes players for disobeying, and characterize when nudges are preferable and when shoves are necessary.

- *Examples of the use of nudges to control exertion in two concrete games*: We provide a full account of the design and implementation of nudge techniques in *PlaneGame*, and
both nudges and shoves in *Gekku Race*. These examples both provide guidance on how the design principles can be used, and how a pre-existing game can be adapted for use with nudges.

- *Evaluation of nudges as tools to reduce over-exertion*: We show that nudges and shoves are effective for keeping players from going too quickly in exergames, and that there is no cost in immersion for their use.

This thesis was aided by the contributions of other members of the EQUIS lab at Queen’s University. The game *Gekku Race* was originally created by Zi Ye [23, 30], and the iteration to which we added nudge techniques was created by Dan Moran.

Information about players’ pedaling speed was extracted and processed by the FireANT software program. FireANT receives network data from a magnet-based speed and cadence sensor attached to the stationary bicycle, and from those signals computes the speed at which the player is pedaling. FireANT then outputs that speed to the games so they can make use of it. FireANT was originally created by Florian Eysseric, and was subsequently updated and maintained by Hamilton Hernández, Luke Walker, and Liam Collins.

This thesis is organized as follows: first, we review background information and related work on exergames, exertion and over-exertion, immersion, and nudges. Second, we describe in detail the design and implementation of nudge feedback into *PlaneGame* and *Gekku Race*. Third, we describe the four design principles derived from our experiences in creating nudge feedback for two games. Subsequently, we detail the two studies testing our nudges, including both the results and the conclusions drawn from them. Finally, we discuss the results of the studies together, and what their implications are for the design of other types of games and for future work.
Chapter 2

Related Work

In this chapter, we give context to the contributions of this thesis through reviewing related work in four key areas. First, we review the field of exergames and the ways they have been used and designed during their history. Second, we examine the importance of maintaining a suitable level of exertion during exercise, and how exergames do not always meet these recommendations. Next, we review the concept of immersion, both its importance in games and how it can be measured. Finally, we explore examples of using nudges, both in general and in digital contexts.

2.1 Exergames

The term exergame is a portmanteau of "exercise" and "game", and refers to a digital game that includes physical exercise as part of its play. A well-known example of a commercial exergame is Konami's Dance Dance Revolution (Figure 1), in which players dance in time with music by hitting buttons with their feet on a purpose-built dance pad, cued by

![Figure 1: Dance Dance Revolution arcade cabinet.](image credit: True Tech Talk Time / CC-BY-SA-4.0, http://creativecommons.org/licenses/by-sa/4.0/)
the game's video component [31]. DDR has inspired other dancing rhythm games, such as StepMania, a DDR simulator that has been expanded and released as open-source software [16]. Despite being played on personal computers rather than an arcade dance surface, StepMania can still be played as an exergame using commercially available dance pads, such as Dance Pad Mania's D-Force.

Other commercial exergames include Wii Sports for the Nintendo Wii [40], in which players play simplified version of five sports (baseball, bowling, boxing, golf, and tennis) using a Wii Remote, an accelerometer-based handheld motion-sensing game controller (Figure 2). The controller is used to mimic the action of playing each sport, for example, by swinging the remote horizontally when playing baseball, or underhand vertically when bowling. Wii Sports boxing, in particular, has been used in exergaming research due to the relatively high level of exertion it provokes [6, 62]. A sequel, Wii Sports Resort, adds additional activities, including archery, canoeing, and parachuting [41].

The Nintendo Wii received additional exergame support with the release of the Wii Balance Board (Figure 3). Players use the board by standing on it and shifting their weight, which is detected by pressure sensors inside the board. The Wii Balance Board was developed in conjunction with the game Wii Fit, an exergame combining traditional exercise activities—yoga

**Figure 2: Wii Remote.**

image credit: Alphathon / CC-BY-SA-3.0
http://creativecommons.org/licenses/by-sa/3.0/
and strength training—with exergame mini-games, such as ski jumping and hula hooping [42].

*Wii Fit* has demonstrated utility for physiotherapy in outpatients [20].

Another commercial peripheral commonly used for exergames is the Microsoft Kinect, which combines a camera, a depth sensor, and processing software to read the position and posture of users (Figure 4). An example is the Ubisoft game *Your Shape: Fitness Evolved*, a fitness-training game that uses the Kinect's camera to show the player on-screen, allowing them to see directly how closely their movements match the current exercise [56].

Stationary bicycles are used in many exergames, as a way of getting players active while staying in place. An early cycling-based digital was Namco's Prop Cycle (Figure 5), an arcade game in which players controlled a virtual human-powered aircraft using the peddles and
handlebars of a stationary bicycle [37]. In 2007, Fisher-Price released the Smart Cycle (Figure 6), an educational exergame targeted at young children [19]. The Smart Cycle attaches to a television for display, and has the player pedal the bicycle while playing educational mini-games covering such topics as spelling, shape identification, and counting.

2.1.1 Non-Commercial Exergames

In addition to commercially-released exergames, many researchers have created exergames for purposes ranging from physical rehabilitation to enabling social interaction. An example is the game SilverPromenade, designed for "frail elderly persons suffering from age-related decrements in cognition and physical abilities with no or little previous gaming experience." [21] In SilverPromenade, up to three players use a Wii Balance Board and Wii Remotes to take a walk through a virtual park, completing mini-games along the way. SilverPromenade provided its players with an enjoyable, non-sedentary leisure activity, and provided an opportunity for increased social interaction among the players.
An exergame that uses off-the-shelf hardware rather than commercial exergaming equipment is *Growl Patrol* (Figure 7), in which audio cues help players find small animals in need of rescue while avoiding a hungry tiger [32]. *Growl Patrol* uses the player's actual position in the world, as recorded by a portable GPS device, as their avatar's position. The game therefore provides its exercise component by having the player run in the physical world.

Another example of a non-video exergame is *Jogging over a Distance*, played by jogging in the real world while conversing with an exercise partner [36]. *Jogging over a Distance* is a networked game that simulates the experience of jogging together, even if the players are not co-located. The game collects heart rate data from both players in real time, and alters the apparent

---

**Figure 7: Growl Patrol Gameplay**

image credit: "Hearing is believing: evaluating ambient audio for location based games" [32], reproduced by permission of author.
direction of the remote player based on relative heart rates: the player with a higher heart rate sounds as though they are further ahead, while a player with a slower heart rate sounds as though they are behind. Players will thus tend to match heart rates, simulating matching each other's pace in co-located jogging.

Heart rate control is used in other exergames. DanceBeat is a dance pad exergame, based on StepMania, that uses a heart rate monitor as part of the control scheme [5]. DanceBeat is designed to control the player's heart rate toward a desirable level by varying the tempo of the music, and thus the speed at which dance pad commands need to be inputted. A low heart rate triggers an increased tempo, intended to provoke more exertion and an increased heart rate, and vice-versa for a high heart rate.

Sportal (Figure 8) is an example of a converted exergame, where a non-exergame is converted into an exergame through the addition of physical input [59]. Sportal is an adaptation of the puzzle game Portal 2, created by Valve Corporation [58]. Sportal offers two possible
control configurations: walking in place, where physical controls are through accelerometers—in the form of Wii Remotes—attached to the player's shins; or suspended walking, in which the player is suspended in a ceiling-mounted harness with their movement captured by a Kinect. In both cases, the player's physical movements are mapped to the game character's actions, including walking to move in-game, and shoulder rotations to rotate the game view.

Converted exergames using indirect control mapping include Thighrim and Calf-Life [29], exergame adaptations of the action-RPG Skyrim [7] and the action-puzzle game Half-Life 2 [57]. Both Thighrim and Calf-Life achieve their exercise components through a stationary exercise bicycle, with players needing to pedal to control their character. The two games offer motivation to continue pedaling even when the game character is not moving through heart rate control, providing special in-game powers to the character when the player's heart rate is within a clinically significant exercise range.

Another cycling-based exergame is PaperDude, based on the Atari arcade game Paperboy [4], in which players deliver newspapers on a paper route [9]. PaperDude maps the player's cycling directly to the character's, and uses a virtual reality headset for display and a Kinect to read the newspaper-throwing motion used to deliver the papers.

The above exergames have a dedicated control scheme, whether using purpose-built hardware or extant technology. It is, however, possible to design an exergame to use any hardware that the player has available, as demonstrated by Brehmer et al.’s GAIM toolkit, used in their example exergame Racing Game [10]. In Racing Game, players steer a car around a racetrack, providing the car's power through their own power on either a stationary exercise bicycle or through a heart rate monitor while running in place. The GAIM toolkit permits
different physical control schemes to be used interchangeably, theoretically permitting exergames to be played with any physical input device, expanding the accessibility of exergames.

### 2.2 Exertion

The goal of exergames is to provide players with exercise, but the amount of exercise that exergames stimulate varies widely. Observed levels of exertion in exergame players range from exertion insufficient to promote health benefits, to over-exertion. The American College of Sports Medicine (ACSM) recommends a minimum of 30 minutes of moderate intensity exercise five times per week, or 20 minutes of vigorous exercise three times per week [2, 46]. Peng et al. [44] and Biddiss and Irwin [8] performed meta-analyses of 16 and 18 studies respectively on energy expenditure in exergames, and found the games frequently failed to meet these recommendations.

In contrast, Rhodes et al. found that affective attitude and adherence to cycling were greater when playing video games than when listening to music, and health benefits were seen after six weeks [47]. Leininger et al. compared playing DDR to exercising on a treadmill, and found that DDR resulted in as much oxygen consumption as walking on a treadmill, with a greater level of enjoyment [33]. Ketcheson et al. found in three different games that a player’s average heart rate was above the minimum threshold recommended by the ACSM, with players spending 79%-88% of their time above this threshold [27, 28].

Exertion levels must be moderated during exercise. In the *Complete Guide to Fitness & Health*, the ACSM emphasizes the importance of performing a warm-up before exercising [3]. Even after a warm-up, exercising too vigorously increases the risk of coronary events [35]. The ACSM defines exertion levels by heart rate reserve, the range between a person's resting and maximum heart rates. Vigorous exercise is defined as 60%-90% of heart rate reserve, with a heart rate surpassing 90% of heart rate reserve representing over-exertion [2].
Efforts to increase player exertion levels in exergames can sometimes result in over-exertion. An example is Ketcheson's use of special power-ups to encourage players to reach and maintain a high heart rate. 15% of participants exceeded the upper limit of vigorous exercise, and had to be instructed to slow down [27]. In this thesis, we focus on the problem of controlling such over-exertion.

2.3 Immersion

To counter over-exertion in exergames, it is necessary to give players some form of feedback to tell them to slow down. However, since the feedback given to players necessarily involves the exercise part of the exergames, we saw a risk of breaking the illusion of playing the game purely for its own sake. We expected this possibility to manifest in the form of reduced immersion in players. To address the necessity of keeping players focused on the gameplay part of an exergame, we integrated all feedback to the players within the internal fiction of the game. We expected this to have a positive impact on players' feelings of immersion.

Several theories of game design have identified immersion as a critical component for enjoyment and engagement with a game. In “Time flies when you’re having fun”, Agarwal and Karahanna list Focused Immersion as one of the five dimensions of cognitive absorption [1]. In the GameFlow model, an adaptation of the concept of flow specifically for games, immersion is listed as one of the eight essential elements of GameFlow [53].

In “Measuring presence in virtual environments: a presence questionnaire”, Witmer and Singer consider immersion a necessary component of presence [64]. In "Measuring and defining the experience of immersion in games", Jennett et al. define immersion as "the prosaic experience of engaging with a videogame", which is the definition of immersion used for this thesis [25]. These two papers also provide questionnaires used in our studies for measuring immersion.
Witmer and Singer's *Presence Questionnaire* measures presence and therefore by extension the component of immersion, while Jennet *et al.*'s *Immersive Experience Questionnaire* measures immersion directly.

We expected players of exergames to feel some degree of dissociation between the gameplay and exercise components, and we expected this dissociation to manifest in the form of reduced immersion in players. In fact, as shall be seen in Chapters 5 and 6, adding feedback to control exertion had little impact on immersion.

### 2.4 Nudges

In their book *Nudge*, Thaler and Sunstein define nudges as indirect suggestions toward a desired outcome [55]. As an example of the concept, the authors cite a tactic used by an airport in Amsterdam to improve the cleanliness of their washrooms. An image of a fly was added to the interior of the airport urinals, providing something to aim at. This nudge successfully reduced spills by 80%. Other examples of successful nudges include encouraging grocery shoppers to purchase local foods by adding a barcode scanner and light-up LEDs to shopping carts indicating how far the foods had to travel [26], encouraging a reduction in electricity consumption by delivering pamphlets to households comparing their energy consumption to that of their neighbours [13], and serving food in smaller dishes to increase the portion’s apparent size and decrease consumption [60].

The use of nudging techniques has also been applied in HCI contexts, though not always explicitly by that name. Coventry *et al.* describe a system for employing nudges in a cybersecurity context to encourage university computer users to use safe networks [14]. Nudges included in the system include providing incentives, such as free printing, to users of secure networks, or the use of emotive colours (such as marking unsecure networks in red). The authors
note that several nudging tools are already in use in HCI for ease-of-use design: "effective defaults, designing for error, understanding mappings, giving feedback, structuring complex choices, and creating incentives." Coventry et al. also draw parallels between nudges and the MINDSPACE (Messenger, Incentives, Norms, Defaults, Salience, Priming, Affect, Commitment, and Ego) framework of the most robust effects on behaviour, whose users include the UK government’s Behavioral Insight Team [18]. In “Emerging Patterns in Active-Play Video Games”, the authors investigate techniques for persuasive interfaces and draw a link between nudges and persuasion: "The immersive and interactive qualities of active-play video games could provide tools for people to nudge their health behaviors in positive ways." [24].

Existing exergame research contains some examples of how nudges might be applied to exergames. In ‘Ere be Dragons [17], players’ current heart rate is compared to their optimal value, and represented as one of five discrete bands of feedback, from low to high. When a player over-exerts, the environment becomes a dense forest accompanied by high-speed audio, while low exertion leads to a desert with quiet, slow sounds. In Balloon Burst [51], gameplay involves shooting down balloons; players can increase the rate at which these appear by pedaling faster. If players exceed the maximum speed, however, they are cued by haptic feedback in the game’s controller.

Outside of exergames, Waterhouse et al. find that playing music for cyclists and speeding up or slowing down the tempo of music has a corresponding effect on pedaling speed and power expenditure [61]. Mandryk et al. use a system of visual overlays to turn commercial games into biofeedback games [34]. The overlays are triggered by negative changes in EEG readings, and partially but progressively block vision of the game to encourage players to calm down. The
graphical effects of the overlays are chosen to match or complement the game; for example, mud splatters on the screen in a dirt-biking game (Figure 9).

There has been some limited investigation into the application of nudges to HCI, but to our knowledge there has been no study in digital contexts of how well nudging techniques work—separate from other aspects of design—or how digital nudges should be designed.
Chapter 3

Designing with Nudges

The purpose of this chapter is to show how nudges can be used to moderate exertion in exergames. For a game’s user interface to operate through nudges, the nudges must motivate players to correct their behaviour. However, the nudges must do this in a way that does not take away the feeling of playing the game for fun. A nudging interface should therefore guide players toward the behaviour required of them, and incent them to perform that behaviour, while fitting smoothly and believably into the fiction of the game world.

We designed nudging techniques following this line of reasoning for two cycling-based racing exergames. The racing game genre was selected with the idea that competing against other racers would drive players to greater levels of exertion in an attempt to outpace rival racers. The increased exertion would trigger more feedback from the nudge techniques, providing more data on how players responded to this feedback. PlaneGame is a single-player game designed specifically for this research, while Gekku Race is a pre-existing multiplayer game with the nudge techniques added afterward.

Both games employ the same two input devices for players: a cordless video game controller (Figure 10), and a recumbent exercise bicycle (Figure 11). The controller is used for the in-game avatar’s actions, while the exercise bike maps the rate at which the player pedals (called cycling cadence or just cadence) to the speed of the player’s avatar. The player’s cadence is read using a Garmin Speed and Cadence Sensor, a magnet-based sensor attached to the bicycle’s crankshaft. The sensor detects each time it passes one of three evenly-spaced magnets on the body of the exercise bike, making it possible to calculate the cadence in revolutions per
minute (rpm). This calculation is performed by a custom software tool called FireANT, and the cadence is passed from FireANT into the games.

Cadence is not the only possible measure for exertion level. Heart rate is a widely available measure of level of exertion. However, changes in heart rate lag behind changes in exertion, sometimes by as much as minutes [50]. Even halting exercise entirely causes heart rate to drop an average of only 17bpm per minute of rest [12]. This makes it difficult to give players immediate feedback when they exceed or return to the target level of effort. We used cycling cadence as our measure of exertion, since cadence can be changed rapidly in response to feedback, removing this obstacle.

Figure 10: Logitech F710 controller used as input in PlaneGame and Gekku Race. image credit: Amazon.com, modified to show controls used in PlaneGame and Gekku Race. http://www.amazon.com/Logitech-940-000117-Gamepad-F710/dp/B0041RR0TW
PlaneGame was designed as a proof-of-concept for nudging to control over-exertion. While succeeding in the game requires cycling sufficiently quickly, there is no advantage to exceeding the required pace. Since following the nudges has no downside, they do not conflict with the player's desire to win.

In PlaneGame, the player plays as an airplane and races against computer-controlled birds to collect floating rainbows. The player uses a stationary exercise bicycle to power their plane; pedaling faster increases the speed of the plane, up to its maximum speed. Each time one
of the rainbows appears on screen, a bird drops down from the top of the screen and races for it alongside the player. When either the player or the bird collects the rainbow, a new rainbow is spawned ahead. These mini-races produce a sequence of short competitions, ensuring the player cannot fall so far behind as to be unable to catch up or advance so far ahead as to have no incentive to pedal hard. That way, each mini-race both allows for and requires a new effort.

A single button on the controller controls the altitude of the plane, causing the plane to rise while the button is pressed and fall while it is not. The player’s cadence influences gameplay by how close it is to a target cadence. If the player is cycling at the target cadence or higher, the
plane will move at top speed. Otherwise, the plane will move more slowly in proportion to how far the player is below the target.

Since adhering to an exact cadence is difficult, the target cadence has a range of ±6rpm on either side of it, called the target range. Initial testing found a variation of 6rpm from target cadence to be the narrowest range that could be maintained by a PlaneGame novice. If players are pedaling below the target cadence, but still within the target range, the plane will still be faster than the birds. Likewise, players are not considered to be cycling too quickly unless they are exceeding the top of the target range.

3.1.1 Nudging in PlaneGame

To nudge players in the direction of reducing their cadence when above the target, we added a layer of feedback to the base game. This feedback layer had to make players understand that they needed to slow down, but also fit naturally into the game world.

The fundamental concept that makes the feedback fit into the game world is that it parallels what might happen to a real plane that flies too quickly: its engine overheats. We represent this visually through smoke billowing from the engine compartment, and by making the plane shudder back and forth (Figure 13 left). Audio feedback also suggests a problem with the engine, by having the sound of the engine become rougher and more grating.

To increase the appearance that the problem is related to the player cycling too quickly, the edges of the screen turn grey, suggesting tunnel vision caused by being out of breath. A motion blur effect complements the greying of the screen, further suggesting the problem is related to speed.

Finally, we wanted to increase the pressure on players who aren’t responding to the initial nudges. As players continue to cycle above the target cadence, the feedback effects become more
To tie increased deviation from the target range to increased response severity, we needed some measure of deviation from the target. For PlaneGame, this value, termed “Severity”, is a combination of the divergence between current cadence and the top of the target range (in rpm) and how long the player has been above the target range (in seconds). We define Severity = divergence/10 + time/5. This formula's constants are intended to ensure that divergence and time have approximately equal effect on Severity, and were derived through iterative refinement in pilot testing.
3.2 *Gekku Race*

*Gekku Race* is a racing game within the *Liberi* suite of exergames [23, 30], that we enhanced with nudge feedback. In *Gekku Race*, as in *PlaneGame*, a player’s cycling cadence determines the in-game speed of the player's avatar. Unlike *PlaneGame*, however, where in-game speed is capped at the target cadence, in *Gekku Race* no such cap exists: pedaling above target cadence continues to increase in-game speed. We expected that this, combined with playing in multiplayer against a real opponent, instead of just AI bots, would motivate greater levels of exertion than *PlaneGame*. *Gekku Race* would therefore provide a more difficult test of our nudge-based feedback design. To combat this increased challenge for our nudges, we introduce the concept of a "shove", a stronger nudge that punishes players for not obeying it.

Players of *Gekku Race* play as ‘gekku’ lizards climbing a wall, trying to be the first to reach the finish line at the top. Along the way, they are able to shoot projectiles at each other to slow down their opposition. *Gekku Race* is a multiplayer game and can have up to eight racers.

The left analog stick of the wireless controller controls the direction the gekku is facing. Pedaling moves the gekku forward at a speed proportional to cadence. Pedaling also charges up the gekku’s projectile attacks, which are activated by a single button press. The projectiles are launched in the same direction the gekku is facing (Figure 14).

3.2.1 Nudging and Shoving in *Gekku Race*

For *Gekku Race*, a new nudge concept needed to be created to match a running lizard, instead of an airplane, when the player over-exerts. Logically, a lizard that is running too quickly may run out of breath. We represented this by the sound of panting, and by puffs of air coming out of its mouth (Figure 15 left). This suggests panting for breath and fits with the game’s cartoon aesthetic.
The gekku panting is an intermittent rather than constant form of feedback, meaning that players might be momentarily unsure after it ends whether they’ve slowed down enough to prevent more panting. To ensure players always know whether they’ve successfully dropped below the target cadence, another form of feedback was added, and this remains as long as the player is pedaling too quickly. Having the screen gradually fade to greyscale was chosen for the same reasons as in PlaneGame’s similar approach (Section 3.1.1): it suggests the tunnel vision caused by being out of breath, and is used in other games to indicate acute health threats.

As in PlaneGame, we wanted to increase the pressure on players who weren’t responding to the feedback. We used the same Severity measure as in PlaneGame (Section 3.1.1) to link

Figure 14: In Gekku Race, two players shoot cashews at each other while racing to the top of the track.
severity of deviation to severity of response, making the screen become progressively greyer and
the panting more frequent.

Since we expected players to be even more inclined to pedal quickly than in *PlaneGame*,
and it would therefore be even harder to convince them to slow down, we added shoves to the
feedback to make it harder to keep moving forward. Since the term "nudge" refers to a
suggestion, this form of feedback that punishes players for ignoring it lies beyond the word's
usual meaning. For this reason, we call a nudge that has negative consequences on play a shove.

Because *Gekku Race* is a racing game, reducing the player’s effectiveness and making it
more difficult to win the race is a strong form of feedback. If players continue to pedal too
quickly, the gekku collapses and stops moving for a time, causing it to fall behind (Figure 15
right). This collapse is complemented by a new, stronger sound of gasping for breath.

---

**Figure 15: Nudges and shove in *Gekku Race*.**

left: Gekkus panting for breath, and the screen turning grey.
right: A gekku collapses and gasps for breath. 
Through iterative testing we found that having the gekku collapse intermittently (every three pants for breath), starting with the second pant event after crossing the target cadence threshold, gave players enough time to notice they were going too quickly and slow back down. If players did not obey the shove, and did not slow down, the gekku's collapse would make it impossible for them to win.

3.3 Design Summary

In PlaneGame, the nudge consists of the player's plane starting to malfunction, spewing smoke and fire, when the player pedals too quickly, spewing smoke and fire. In Gekku Race, the player's lizard gets tired and starts panting for breath. Our version of Gekku Race also includes shoves, where, in the face of persistent over-exertion, the lizard will sometimes fall and gasp for breath, making it impossible to win the race unless the player slows down enough to prevent another collapse.

In our proof-of-concept study with PlaneGame (Chapter 5), we show that nudges are effective at getting players to slow down, and players find them to be a natural fit with the game. In the second high-energy study with Gekku Race (Chapter 6), we find that ordinary nudges lose their effectiveness when players are highly motivated to over-exert, as we expected. However, we show that including shoves keeps players from pedaling too quickly, and these shoves, like the earlier nudges, still feel natural. In the next chapter, we present a set of design principles for the systematic design of nudges, intended to support the process of creating nudge-based feedback.
Chapter 4

Design Principles

As explored above in Section 2.4, existing games have used nudges to control player behaviour. For example, in 'Ere be Dragons, players are nudged toward an optimal heart rate by changing the background environment, both visually and audibly, according to how far players are from their target heart rate [17]. Balloon Burst players can make more balloons to burst by increasing their pedaling speed, but if they pedal too quickly the game controller will cue them using haptic feedback [51]. Visual overlays can be used to turn commercial games into biofeedback games, progressively blocking vision of gameplay in response to negative EEG readings [34].

However, while these games have used nudges to good effect, there is still a lack of principles supporting systematic design of nudge-based interfaces. In this chapter, we suggest design principles aiding in the creation of nudge-based feedback in video games. These principles were co-developed with the nudge versions of PlaneGame and Gekku Race, and describe the elements we found to be effective during design and early testing. The effectiveness of the nudge feedback for the two games was tested by the two studies described later in this thesis (Chapters 5 and 6).

The four design principles are: natural integration, making the feedback feel like it's part of the game; comprehension, making sure the feedback is both clear and conspicuous; progression, making the feedback become more severe as players continue to perform the negative behaviour; and multiple channels, making use of multiple forms of feedback to work together for a stronger overall system.

31
4.1 Natural Integration

Nudges should fit the game world: In order not to disrupt players' sense of playing an exergame for fun, nudges should feel as though they are a natural part of the game world. An element of the game that corresponds to an entity or concept inside the game’s environment should feel more natural than if it were apparently unconnected to the game’s reality. Any addition to the feedback system must therefore be chosen to avoid conflicting with aspects of the game environment, either native to the game itself or as introduced by other feedback mechanisms.

The problem of controlling players' behaviour, a problem that nudges try to solve, is that players may resent the game telling them what to do. Natural integration helps with softening the impact of making players change their behaviour by creating the impression that the feedback is simply part of the gameplay. Players normally want to win the game, so if part of the game involves avoiding a particular behaviour, then players will tend to be motivated to do what it takes to win.

The nudging interfaces for both PlaneGame and Gekku Race began from a natural element of the gameplay. In PlaneGame, the plane begins smoking and shaking, as though the engine is overheating. In Gekku Race, the lizard pants as though tired from its exertion. In both cases, the feedback shows that something is awry, as well as suggesting the cause: excessive speed.

In contrast, before our additions to Gekku Race, the method the game used to inform players they were over-exerting themselves was to literally tell them, via text across the screen reading “SLOW DOWN!” While this is an unambiguous message to players as to what is
required of them, it is separate from the game world, and risks distracting players or reminding them that they are exercising as well as playing a game.

Mandryk et al.’s system of putting overlays on the screen in biofeedback games [34] is a good example of natural integration. The overlays are designed to be thematically appropriate for the game being played, using mud splatters in a dirt-biking game, or branching frost in an ice hockey game.

4.2 Comprehension

Nudges should be clear and conspicuous: A nudge is unlikely to work if the recipient does not understand it. We consider comprehension to be made up of two components, both critical; clarity and conspicuity. By clarity, we mean that the player needs to understand what the feedback means. In Gekku Race, for instance, going too quickly causes the lizard to tire, suggesting that a solution might be to slow down and put less strain on the creature.

Another means of increasing the clarity of a nudge is to rely on conventions used in other games. In both PlaneGame and Gekku Race, excessive speed causes the screen to begin turning grey. This might suggest the visual effects of being severely short of breath, but also resembles a technique used in other games (such as Tomb Raider 2013 [15], DayZ [22], and the Uncharted series [38]), where acute health threats cause the screen to begin losing colour. If the player is familiar with this technique from other games, the player should clearly understand the feedback as something needing immediate attention for the sake of the game avatar’s well-being, even though there is no innately obvious connection between a grey screen and a malfunctioning plane.

By conspicuity, we mean that a nudge needs to be prominent enough that the player takes notice and can react. Too gradual an onset can make the feedback difficult to notice, thus delaying the player’s ability to react. In PlaneGame, the mechanical distress of the plane is
sharply delineated from its normal state by sounds, animation, and particle effects. Taken together, it’s difficult to ignore that something has gone wrong.

Another possible barrier to conspicuity is feedback that is intermittent or cyclic, rather than constant. In *Gekku Race*, for example, the need to slow down might not be obvious if the lizard has stopped panting, or if the panting will continue a second later. Adding continuous feedback in the form of whole-screen greying allows players to know the instant they cross the threshold between acceptable and excessive speed.

In summary, the principle of comprehension requires that nudges be both clear and conspicuous. To be conspicuous, feedback needs to be immediate and always present while the player continues the negative behaviour. To be clear, it needs to be obvious what the feedback means. This clarity might be innate to the type of feedback used, or else can leverage conventions from existing games.

### 4.3 Progression

Nudges should progress from low to high severity: To increase pressure on players who, intentionally or not, ignore the nudges, more serious deviation from the desired behaviour should be met with more severe feedback. The definition of greater deviation depends on the needs of the game. Both *PlaneGame* and *Gekku Race* use the Severity metric (Section 3.1.1), which comprises time and divergence away from the target cadence.

The way in which feedback becomes more severe should be based on the needs of the game; possibilities include intensifying an existing form of feedback or switching to a gameplay channel with more effect on the player. In *PlaneGame*, engine fire and red around the edges of the screen are introduced after a sufficiently high value of Severity, but increased severity of
feedback is primarily produced by intensifying the existing modes of feedback: more smoke and louder engine knock.

Such aesthetic feedback is suitable for PlaneGame’s comparatively low-intensity motivation toward increased exertion, but as we shall see in Chapter 6, the high-energy Gekku Race demands that the tangible benefits of over-exertion be met with tangible punishment for players who do not respond to the warning signs. By stopping the climbing lizard as it gasps for air when the player has been above the target cadence for too long, players are not able to dismiss this more severe feedback in favour of greater game performance. Once they realize there is an impediment to winning if they ignoring the feedback, even the earlier signs will become an effective discouragement, since the players now know the more severe subsequent consequences for continuing above the target range.

'Ere be Dragons's five discrete bands of heart rate feedback, from low to high [17], provides a good example of progression. At low heart rates, the visual background of the game is sparse and the audio slow and quiet, while high heart rates correspond to a densely wooded landscape with fast sounds. The five feedback bands tell players not only whether they are in their optimal range, but also provide information on how far away they are from their target. Mandryk et al.'s overlays [34] are an example of progression to a shove, as the overlays progressively block the player's view of the game world, making the game unplayable at high levels of deviation from the desired behaviour.

4.4 Multiple Channels

Nudges should employ multiple feedback channels: Though a single excellent form of feedback might be sufficient for some purposes, nudges can be made stronger by using multiple feedback channels. Examples of feedback channels include visual, auditory, haptic, or direct
gameplay effects. We have found it helpful to use multiple channels when designing nudges, for several reasons.

First, using redundant nudges through multiple channels can make the feedback system more robust against player distraction or inattention. If a player misses a visual nudge due to looking away or at another part of the screen, they might still notice and react to an auditory cue, while a player listening to music might miss an auditory cue, but still respond to a visual or haptic one.

Second, using multiple feedback channels allows specialized nudges to complement each other and produce a more effective feedback system overall. For example, the periodic stopping of the lizard in *Gekku Race* is a strong motivator, but it is not necessarily obvious to the player what they are doing wrong to cause this effect. The audio channel’s gasping sound conveys that the gekku is out of breath, making it obvious that it is stopping to catch its breath.

Third, having multiple feedback channels is helpful for satisfying other principles. A nudge made from three components happening simultaneously is more conspicuous than a unidimensional nudge, and adding more types of feedback can be an effective way of introducing the progression in severity. Because of this synergy with the other principles, it can be helpful to employ the principle of multiple channels early in the design process so that the available feedback mechanisms can be arranged to best meet all other goals.

*’Ere be Dragons* [17] offers an example of multiple feedback channels, using both visual (the nature of the game’s environment) and auditory (the tempo of the game’s background audio) cues for players. An example of haptic feedback is *Balloon Burst*, which adds controller rumble when players are above their target cadence [51].
4.5 Iterative Design

In our design of the nudges used in *PlaneGame* and *Gekku Race*, we carried out several iterations of design and informal testing. This process revealed that it is easy to violate these principles and that doing so can lead to negative user experience.

Our most common problem with nudges was insufficient clarity. What seemed obvious to the developer was at times perceived differently by players. In *Gekku Race*, for example, the shove of the gekku lizard collapsing and gasping for breath needed iterative refinement. The animation for the collapse was initially too subtle, giving players the impression that the lizard had glitched and stopped animating for a moment, rather than collapsing in exhaustion. The solution was to have the lizard close its eyes in fatigue, and exaggerate the animation of collapsing so its legs splayed far out to its sides.

The gekku collapsing animation is an example of a nudge that needed modification before it was suitable for inclusion in the game. Occasionally, a nudge may need to be removed entirely. In an early version of *PlaneGame*, the clouds in the background would rush backward when the player was above the target range, to make it more clear that the feedback was occurring because the player was pedaling quickly.

However, the speeding clouds made players feel as though they were speeding up. This again constituted a violation of the principle of comprehension due to insufficient clarity, since players did not understand that they were supposed to slow down; instead, they wanted to continue to over-exert to increase their apparent speed in the race. Removing the rushing clouds made players once again understand that the other feedback—the smoke and fire, the rattling engine—were a sign that something was going wrong, and this player comprehension improved the performance of the nudges.
These examples of problems and their solutions are only a sample of some of the largest and most visible complications. Just as with any interface, nudges built according to the design principles need to be tested and refined iteratively for best results.

4.6 Summary

Though nudges have been used before by other researchers, this chapter represents the first attempt, to our knowledge, to devise principles underlying their design. Drawing from our own experiences and those of other researchers, we created a list of four design principles that can be followed to create nudge-based feedback: the principle of natural progression, the principle of comprehension, the principle of progression, and the principle of multiple channels.

We also describe some of the particular challenges we encountered in iterative refinement and testing of our nudge-based feedback designs, in the hope that other designers can use them as examples of how to adapt when nudges have unexpected effects.

Having covered the design of both the specific nudges we used for testing, and the design principles that are derived from our experiences, the next two chapters cover the studies we conducted to test the nudge-based feedback in PlaneGame and Gekku Race.
Chapter 5

Proof-of-Concept Study: PlaneGame

We performed two studies of nudge-based feedback to address the following questions:

• How effective are nudge techniques at convincing players not to over-exert?
• What effect do nudge techniques have on players’ immersion in the game?

Each study also had its own individual goal. The proof-of-concept study, with PlaneGame, compares the nudge techniques against a text-based condition that makes no effort to maintain immersion, allowing us to compare how much impact designing with nudges has on the effectiveness and immersiveness of the game.

The high-energy Gekku Race study's particular goal was to compare regular nudges against shoves, to find out how much of a difference the more forceful aspects of shoves makes.

The nudge-based feedback techniques were designed to accomplish two things: to get players to slow down when they were working too hard, and to avoid reducing players' sense of immersion. To test whether players indeed slow down when they cross the target cadence threshold, each game was programmed to record two measures:

• Average cadence: player's average cadence over the course of the race. An effective system should get players on average near or below the target cadence.
• Time above target: how long players spent per race above the target cadence. If players spend very little time above the target, it implies the feedback prompted a rapid and lasting response.

For immersion, we used two measures: first, after each race, players filled out a validated questionnaire to measure their sense of immersion. Second, to measure natural fit of the nudges
within the game, in a closing interview with each participant, we asked whether they found any of the versions of the game to be a more natural fit to the game. Since our primary means of trying to preserve immersion was to integrate our feedback within the game world, we expected this question to correlate strongly with players' feelings of immersion.

In this chapter, we present the results of the first study. The second study and its results are described in the following chapter.

5.1 Study Design

In our first study, we wished to test our nudge-based feedback techniques against a version of the same game without any feedback telling players to slow down, to confirm that our techniques were effective. We also wanted to test our nudges’ effectiveness compared to a more straightforward text-based interface that gave the same information in a way not designed to be immersive.

This first study was designed to test the implementation of nudge-based feedback techniques according to the two previously-established metrics: effectiveness in preventing players from exceeding their target cadence, and immersiveness. To test these factors, three versions of the game were tested against each other:

- **Control**: The only indicator of current divergence from the target cadence is the speed of the plane relative to that of the birds. There is no direct feedback indicating whether the current cadence is correct or too high.
- **Nudge**: Nudge techniques inform players when they are above the target; as previously described in Section 3.1.1, the player's plane begins visibly and audibly malfunctioning. As in the control condition, there is no indication of how close the current cadence is to exceeding the top of the target range.
- **Textual:** Players are shown the target cadence at the top of the screen for reference, while their current cadence is displayed immediately underneath the plane (Figure 16). Whenever the player’s cadence is above the target range, the current cadence text turns red. Like the nudge condition, the textual condition lets the player know they are above their target range. In fact, it provides more information, as the textual condition allows players to know exactly where they are relative to the target cadence.

The second research goal was to ensure that the techniques were able to keep players pedaling slowly for a warm-up, and not just at a steady pace during the main part of an exercise session. We elected to simulate this by having two different target cadences for all participants: a moderate one to represent the main part of an exercise session, and a slower one to determine whether these techniques could persuade players to pedal more slowly than they wished, as would likely be the case during a warm-up.

The speeds chosen to represent warm-up and main-game speeds were 40rpm and 60rpm, respectively. Pilot testing showed 40rpm to be slow enough that it was slightly difficult to maintain. If we could get players to adhere to a low speed, that would be evidence that these techniques could support a warm-up phase of exercise; 60rpm, meanwhile, was a more comfortable speed that was easy to reach. Both cadence targets were low enough that pilot-test players had little difficulty exceeding them. This was by design, as it meant players would be more likely to exceed the cadence target and we would therefore have more data to analyze.
The textual condition was expected to be more effective at controlling exertion than the nudge condition, but less immersive for the player, and the control condition was expected to be ineffective at cadence control, but the most immersive due to freedom from any distraction.

5.1.1 Participants

24 students from the Queen's University community were recruited as participants. During pilot testing, we discovered that players both unfamiliar with video games and accustomed to maintaining a steady cadence on a bicycle had noticeably different behaviours during testing, compared to the target audience of video game players who did not exercise.
regularly. Accordingly, the 24 participants were screened to include only those who had at least 50 hours of lifetime experience with video games and who did not use a stationary exercise bike or go long-distance cycling regularly (defined as more than once per week). Before beginning, participants were given the PAR-Q+ physical activity readiness questionnaire [11] to ensure they had no medical conditions that counter-indicated the use of an exercise bike.

5.1.2 Data Collection

Three forms of data were collected. First, the game itself logged players’ average cycling cadence during gameplay, total seconds that cadence was above the top of the target range, and final game score.

Second, after each segment of gameplay, participants filled out a brief questionnaire. The questionnaire had a stand-alone question about motivation: “Do you feel motivated to slow down when you’re pedaling too quickly?” The remaining questions were the Involvement/Control subscale of Witmer and Singer’s Presence Questionnaire (PQ) [64], which was used to measure participants’ sense of immersion through the related concept of presence. All questions were answered on a seven-point Likert scale. PQ results are the sum of Likert responses for all questions, and therefore the results of the 11-question Involvement/Control subscale ranged from 11 to 77.

Finally, participants engaged in a semi-structured interview exploring their subjective impressions and rankings of the three versions of the game. The three conditions were referred to descriptively rather than by name (for example, the nudge condition was “the version with the smoke”) to avoid influencing the participants’ answers. Responses were coded into categories to allow comparisons between participants. For example, a participant who said they preferred the
nudge condition because it was "more visually appealing" and one who said it was because it was "more interactive and interesting" both fell under the code of "more fun or interesting".

5.1.3 Method

Each participant played the game under six conditions: three game conditions at each of the two target cadences. These conditions were order-balanced according to a Latin square to compensate for the effects of increasing skill at the game as participants progressed.

A Latin square is an $n$-by-$n$ matrix of $n$ items, arranged so that each item appears exactly once in each row, and once in each column. A Latin square can be used for order-balancing different conditions in a study, by assigning one condition to each of the $n$ items in the matrix. When running the study, the conditions are run in the order given by the rows of the matrix, for each row an equal number of times. In this way, each condition appears in every possible position—first, last, and all intermediate positions—an equal number of times. By doing this, any influence on the results caused by the position of a trial will be spread equally amongst all conditions. Therefore, the researcher can be confident that any perceived differences between conditions are truly caused by the conditions, rather than by the position in which the conditions were run.

This sort of order-balancing can also be accomplished by exhaustively running the conditions in every possible sequence. However, to do so requires $n!$ different iterations of the study, where $n$ is the number of conditions. A Latin square, meanwhile, always contains $n$ rows, and therefore requires only $n$ iterations of the study.

It is also possible, as we have done in this study (Table 1), to select Latin squares to balance for first-order carry-over effects [45]. A first-order carry-over effect, in this context, is an effect caused by one condition being run immediately before or after another. To balance against
these carry-over effects, Latin squares must be selected so that not only does each condition appear in each position exactly once, but it is preceded by and followed by every other condition exactly once. For studies with an odd number of conditions, a second complementary Latin square is needed, leading to $2n$ iterations. Even numbers of conditions still require only $n$ iterations. Order-balancing against first-order carry-over effects in this way is called a "Williams design" after its originator, E.J. Williams [63].

Before beginning the game, players were given an explanation of how to play. Participants familiarized themselves with the gameplay by playing a version of the game with no target cadence, in which the plane always moved at top speed, and no cadence feedback was provided.

After this practice round, participants were informed that they needed to reach the target cadence for the plane to reach full speed. They were cautioned against exceeding the target cadence, and the three conditions were introduced verbally and through printed screenshots. Players were also informed of the two speeds used in the game. Players were not forewarned which version of the game they would be playing, but had a three-second adaptation period.

<table>
<thead>
<tr>
<th>Control 60rpm</th>
<th>Nudge 60rpm</th>
<th>Textual 40rpm</th>
<th>Textual 60rpm</th>
<th>Nudge 40rpm</th>
<th>Control 40rpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nudge 60rpm</td>
<td>Textual 60rpm</td>
<td>Control 60rpm</td>
<td>Control 40rpm</td>
<td>Textual 40rpm</td>
<td>Nudge 40rpm</td>
</tr>
<tr>
<td>Textual 60rpm</td>
<td>Control 40rpm</td>
<td>Nudge 60rpm</td>
<td>Nudge 40rpm</td>
<td>Control 60rpm</td>
<td>Textual 40rpm</td>
</tr>
<tr>
<td>Control 40rpm</td>
<td>Nudge 40rpm</td>
<td>Textual 60rpm</td>
<td>Textual 40rpm</td>
<td>Nudge 60rpm</td>
<td>Control 60rpm</td>
</tr>
<tr>
<td>Nudge 40rpm</td>
<td>Textual 40rpm</td>
<td>Control 40rpm</td>
<td>Control 60rpm</td>
<td>Textual 60rpm</td>
<td>Nudge 60rpm</td>
</tr>
<tr>
<td>Textual 40rpm</td>
<td>Control 60rpm</td>
<td>Nudge 40rpm</td>
<td>Nudge 60rpm</td>
<td>Control 40rpm</td>
<td>Textual 60rpm</td>
</tr>
</tbody>
</table>

Table 1: Latin square used for order balancing with PlaneGame.
before the game began, allowing them to note both the condition and target speed, and to adapt accordingly.

A single game of PlaneGame is one minute long. To reduce the risk of players becoming fatigued in the later races and therefore potentially affecting our results by pedaling more slowly, the resistance on the exercise bike was set to the lowest setting.

Combined with the low cadence targets, this meant that players did not exert themselves hard enough to be in physical distress, as they would be in the case of genuine over-exertion. Given that actual exhaustion would provide additional incentive to slow down, we judged that any feedback capable of reducing the speed of non-tired players would also be effective for players who were working too hard. Accordingly, we saw no need to bring participants to a state of actual over-exertion.

5.2 Results

We first summarize the key findings of this study. More detail follows in Sections 5.2.1 Performance Metrics, and 5.2.2 Questionnaires. The results are discussed in Section 5.3 Discussion.

Both the nudge and textual conditions were better than the control condition at getting players to slow down, but with no significant difference between them.

At the slower speed only, players got a lower average score in the nudge condition than in either of the other conditions. At the faster speed, average scores were statistically indistinguishable.

Players reported feeling more motivated to slow down in the nudge and textual conditions than in the control condition, with no difference in motivation between the two. There were no statistically significant differences between any two conditions in PQ presence scores.
5.2.1 Performance Metrics

A one-way repeated measures ANOVA was conducted to compare the effects of the game conditions on players’ average cycling cadence. At a target cadence of 40rpm, a significant effect was found; Wilks’ Lambda=.310, $F(2,22)=24.16$, $p<.001$. Post hoc comparisons via paired samples t-tests with Bonferroni correction showed the control condition (M=57.84, SD=12.67) produced a significantly higher average cadence than the nudge condition (M=41.94, SD=2.55); $t(23)=6.53$, $p<.001$. The control condition also had a greater average cadence than the textual condition (M=43.24, SD=3.22); $t(23)=5.64$, $p<.001$. No significant difference was found between the nudge condition and the textual condition at the $\alpha=.05/3$ level; $t(23)=2.11$, $p=.046$ (Figure 17 left).

At a target cadence of 60rpm, there was again a significant difference in average cadence between game conditions; Wilks’ Lambda=.522, $F(2,22)=10.06$, $p=.001$. Post hoc t-tests showed higher average cadence in control (M=67.47, SD=10.65) than nudge (M=59.37, SD=1.84); $t(23)=3.99$, $p=.001$. Average cadence was also higher in the control condition than the textual condition (M=60.33, SD=1.45); $t(23)=3.40$, $p=.002$. Again, no significant difference was found between nudge and textual at the Bonferroni-corrected $\alpha=.05/3$ level; $t(23)=2.09$, $p=.048$ (Figure 17 left).

At 40rpm, an RM-ANOVA on time over target range showed significance; Wilks’ Lambda=.214, $F(2,22)=40.48$, $p<.001$. Post hoc t-tests showed significantly more time over target in the control condition (M=45.95, SD=23.13) than in the nudge condition (M=6.50, SD=6.37); $t(23)=8.36$, $p<.001$. Participants also spent more time above the target range in control than in textual (M=9.86, SD=14.40); $t(23)=6.41$, $p<.001$. No difference was found between the nudge and textual conditions; $t(23)=1.60$, $p=.123$ (Figure 17 right).
Time over target range also showed significance for 60 rpm; Wilks’ Lambda=.214, F(2,22)=40.48, p<.001. Time over target was significantly higher in the control condition (M=27.79, SD=24.52) than in the nudge condition (M=3.65, SD=3.00); t(23)=5.06, p<.001. Control was also higher than textual (M=2.29, SD=3.30); t(23)=5.30, p<.001. There was no significant difference between nudge and textual; t(23)=1.49, p=.150 (Figure 17 right).

Game score showed significant differences across conditions at a target cadence of 40 rpm; Wilks’ Lambda=.684, F(2,22)=5.08, p=.015. Scores in the control condition (M=24.13, SD=1.15) and textual condition (M=24.13, SD=1.42) were indistinguishable; t(23)=0.00, p=1.000. However, t-tests showed significantly lower scores in the nudge condition (M=22.79, SD=2.13) than the control condition; t(23)=3.00, p=.006. Scores were also lower in nudge than in textual; t(23)=2.82, p=.010 (Table 2).
The RM-ANOVA for game score at a target cadence of 60 rpm found no significance at the \( \alpha = .05 \) level; Wilks’ Lambda=0.797, \( F (2,22) = 2.80, p = .082 \). Given that the ANOVA was not significant, we did not analyze the pairwise results.

### 5.2.2 Questionnaires

A non-parametric Friedman test of the stand-alone motivation question showed statistical significance at 40 rpm; \( \chi^2 = 19.46, p < .001 \). Post hoc tests with the Wilcoxon signed-rank test showed that players felt less motivated to slow down in the control condition (Mdn=5.5, IQR=4-6) than in the nudge condition (Mdn=7, IQR=6-7); \( Z = -3.45, p = .001 \). They were also less motivated in the control condition than in the textual condition (Mdn=6.5, IQR=6-7); \( Z = -2.95, p = .003 \). There was no significant difference between the nudge and textual conditions; \( Z = -1.36, p = .175 \) (Figure 18).

For a target cadence of 60 rpm, the motivation question again showed statistical significance; \( \chi^2 = 21.00, p < .001 \). Players were less motivated in the control condition (Mdn=4.5, IQR=3-6) than in the nudge condition (Mdn=7, IQR=6-7); \( Z = -3.69, p < .001 \). Players also felt less motivated to slow down in control than in textual (Mdn=6, IQR=5-7); \( Z = -2.79, p = .005 \). No significant difference was found between nudge and textual; \( Z = -1.83, p = .067 \) (Figure 18).

For the presence score on the PQ, the Friedman test showed no significance on a target cadence of 40 rpm; \( \chi^2 = 4.69, p = .097 \). However, there was a significant difference at 60 rpm;

<table>
<thead>
<tr>
<th>Condition (Mean ±SD)</th>
<th>Comparison</th>
<th>t (23)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (24.13 ±1.15)</td>
<td>Nudge</td>
<td>3.00</td>
<td>0.006</td>
</tr>
<tr>
<td>Nudge (22.79 ±2.13)</td>
<td>Textual</td>
<td>2.82</td>
<td>0.010</td>
</tr>
<tr>
<td>Textual (24.13 ±1.42)</td>
<td>Control</td>
<td>0.00</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Table 2: Pairwise t-tests for PlaneGame game score at 40 rpm.**

The RM-ANOVA for game score at a target cadence of 60 rpm found no significance at the \( \alpha = .05 \) level; Wilks’ Lambda=0.797, \( F (2,22) = 2.80, p = .082 \). Given that the ANOVA was not significant, we did not analyze the pairwise results.
\( \chi^2 = 7.57, p = .022 \). Wilcoxon signed-rank tests showed no significant difference between the control (Mdn=68, r=51-75) and nudge (Mdn=69.5, r=50-76) conditions; \( Z = -0.56, p = .594 \). There was no significant difference between control and textual (Mdn=68, r=51-73); \( Z = -1.20, p = .237 \). No significance was found between the nudge and textual condition; \( Z = -1.31, p = .196 \).

Figure 18: Questionnaire responses for PlaneGame. "Do you feel motivated to slow down when you're pedaling too quickly?", rated on a 7-point Likert scale. Vertical bars show interquartile range. Horizontal hats indicate statistical significance at \( \alpha = .05 \).
5.3 Discussion

As we expected, players cycled closer to the target range in both the nudge and textual conditions than in the control condition, since the game was informing them when they were pedaling too quickly. All participants reported that the feedback in both the nudge and textual conditions was clearly understandable. However, there was no significant difference between nudge and textual in slowing players down. The only performance difference we saw between the two was in-game score, where players had lower scores in the nudge condition, but only for the lower target cadence of 40rpm.

It is not obvious that game score should be lower in the nudge condition than in either the control or textual conditions. But since the number of rainbows collected does not depend on staying below the top of the target range, it is possible that players were over-correcting in response to the nudges and therefore oscillating around the target cadence. This would cause them to sometimes be pedaling slowly enough that fewer rainbows had time to spawn, or else give enemy birds a chance to take some. Oscillating like this is easier to avoid in the textual condition, since the player knows exactly how close they are to the target.

The effects of the conditions on immersion are less obvious. We expected to find that immersion was highest in the control condition, due to a lack of any distractions that might break immersion; lowest in the textual condition, with its direct feedback not integrated into the game world; and between the two in the nudge condition, being more distracting than control but more integrated than textual. Players found the control condition less motivating than the other two, but most of them also found the nudge and textual conditions distracted from the gameplay at least a little: for the nudge condition, six said it was distracting and 10 a little distracting; for textual, seven said distracting and six a little distracting.
However, in the interviews the majority of participants said they found the nudge condition to be the most natural fit for the game (15, compared to five for textual and three for control). When asked to choose favourite (10 nudge, eight textual, six control) and least favourite (eight control, eight textual, six nudge) conditions, their reasons supported this result. Four of the eight participants who chose the textual condition as their least favourite reported that it was because it was a poorer fit for the gameplay. Conversely, many players who liked or disliked the nudge condition did so for game-related emotional reasons: they found the nudge either fun and interesting if they liked it (five of ten), or if they disliked it, a more visceral reminder of what had gone wrong when they made a mistake (four of six). However, the questionnaire we were using to measure immersion, the Presence Questionnaire, showed no differences between the conditions in the pairwise tests.

A possible reason that players do not feel more immersed, despite their feeling that the nudge condition suited the game best, is that we set the target cadence values deliberately low. This was to try to avoid exhausting players, as fatigue might affect their responses. We also wanted to increase the likelihood that players would experience feedback, giving them more opportunity to correct their performance and yield more data points. However, the need to avoid pedaling too quickly came up much more often at these lower cadences, and was always on players’ minds. It is possible that if the target were set higher, as it would be to control over-exertion in actual exercise conditions, players would have fewer encounters with the feedback, and would feel less distracted from the gameplay by it.
Chapter 6

High-Energy Study: Gekku Race

The first study demonstrated that nudge techniques are, in PlaneGame, as effective as direct textual feedback for controlling player exertion levels. In PlaneGame, players were willing to slow down when feedback directed them to, even when there was no direct consequence to continuing at an elevated pace. This was encouraging, but we suspected that feedback with no direct consequences for ignoring it would be less effective in a game that strongly motivated players to pedal rapidly.

In order to address this concern, as well as simply to confirm that our techniques worked in other games, we performed a second study using the game Gekku Race. Gekku Race is one of the games in the Liberi suite of exergames [23, 30]. Previous studies with Gekku Race have shown that some players can exceed safe exertion levels; Ketcheson reports over-exertion in as many as 15% of players [27].

Given the history of Gekku Race, we expected it to be more difficult to control player exertion levels than with PlaneGame. Accordingly, we expected that the previous "gentle" nudges would be less effective. However, we expected that "shove" nudges, with direct consequences that made the game more difficult to play when above the target exertion level, would effectively motivate players to slow down.

6.1 Study Design

Like the first, this study was meant to test both the effectiveness and immersiveness of our feedback techniques for controlling player exertion. The first study tested our nudge
techniques relative to a numeric feedback system; however, this time we wished to compare nudging techniques across the spectrum of nudges versus shoves. The conditions are as follows:

- **Control**: Players receive no feedback about whether they are pedaling too quickly.
- **Nudge**: Players’ lizard avatar intermittently pants for breath when players pedal above the target cadence, and the screen starts turning grey as described in Section 3.2.1.
- **Shove**: All feedback described in Section 3.2.1 is delivered, including all the effects of the nudge condition. Additionally, the lizard sometimes collapses and gasps for breath, preventing the player from continuing until the lizard recovers. This recovery takes long enough that the player will fall behind in the race, making it necessary to slow down to prevent the lizard from collapsing again.

In this study, we did not specifically test slower versus faster target cadences. However, we still wanted to ensure that participants would see the feedback so we could gauge their responses to it. We therefore set the target cadence to 75rpm, a speed that most preliminary testers considered to be neither slow nor fast, and which was easily exceeded when participants wanted to increase their speed to win the race.

We expected that, since it did not offset the benefits of pedaling rapidly, most players would ignore the feedback in the nudge condition. It would therefore not be significantly more effective at lowering exertion that the control condition, but would also not reduce immersion. We expected the shove condition to successfully cause players to slow down, but to be somewhat less immersive.

### 6.1.1 Participants

20 participants, in 10 pairs, were recruited using the same criteria as in the first study: members of the Queen's University community who reported 50+ lifetime hours playing video
games, and did not use an exercise bike or go long-distance cycling more than once per week. Mean age of participants was 19.7, and 65% were female. As before, participants were given the Par-Q+ questionnaire [11] to ensure they were able to exercise using a stationary bicycle.

6.1.2 Data Collection

Data was collected from three sources. First, the game logged the average cadence of both players, and the proportion of time per race they spent above the target cadence of 75rpm.

Second, participants filled out questionnaires between races. A custom questionnaire asked participants to rate their agreement with two statements on a five-point Likert scale: “The game made it clear when I was pedaling too fast”, and “It was frustrating when I was trying to win the race but the game was telling me to slow down”. The other questionnaire was the Immersive Experience Questionnaire (IEQ) [25]. We used the IEQ for this study, rather than the Involvement/Control subscale of Witmer and Singer’s Presence Questionnaire [64] that we had used in the previous study. This was because the IEQ was more directly targeted at immersion, and because the PQ is designed for virtual environments, whereas the IEQ is designed for use with video games generally. The 31 questions of the IEQ were also measured on a five-point Likert scale and then summed, giving a result range of 31 to 155.

Third, we conducted a semi-structured interview with participants, where they were asked to evaluate the three conditions by some specified measures (Table 3), and were also given a chance to express their opinions of the game and the study. As in the first study, conditions were referred to descriptively or by the order in which participants played them to avoid leading answers, and responses were coded into categories.
Of the three versions of the game, was there one you liked best? Why?

Was there a version of the game you liked least? Why?

Did one of the three versions seem like a more natural fit for the gameplay? Why?

Did you understand what the panting and gasping were telling you?

Did you feel motivated to slow down when you saw the lizard panting? When it stopped to gasp?

Did you find that either the panting or the gasping were a distraction from the gameplay?

Did you ever feel like you had to pedal faster than the game wanted you to in order not to fall behind?

<table>
<thead>
<tr>
<th>Table 3: Interview questions for Gekku Race.</th>
</tr>
</thead>
</table>

### 6.1.3 Conditions

In the first study, we used Latin square balancing to compensate for the effect of participants becoming more familiar with the game as they played. For the second study, there were two serious learning effects we anticipated and needed to avoid.

First, if participants played the shove condition before the nudge condition, they might obey the nudge condition because of their memory of the shove condition, and not because of anything innate to the nudge condition. Second, if they played the control condition after either of the others, any effect the other condition had might be retained even without feedback, since the target cadence was the same each time.

Both problems can be solved by running all participants in the order control-nudge-shove. However, we had to make sure that any observed differences were due to the conditions, and not the order in which they were played. We were concerned that players might develop better strategies as they played, or get tired and slow down. Therefore, we performed a pilot with 10 participants (recruited as in the full study, 50% female, mean age 20.7) in which they played the control condition three times in a row. Pilot testers were not eligible to be participants in the full study.
One-way repeated measures ANOVAs for this pilot test showed no difference in the populations for average cycling cadence (Wilks’ Lambda=.723, F(2,8)=2.55, p=.119) or for proportional time above target cadence (Wilks’ Lambda=.902, F(2,8)=.734, p=.458). According to a Friedman test, there was also no significant difference in immersion scores: $\chi^2=1.63, p=.458$.

This pilot showed no significant differences between play positions for any of our metrics, implying that the learning effect caused by the order of the conditions had no impact on our measures. Participants also displayed no tendency to change their strategies across races. Since the only change made to the full study from this pilot was the introduction of the nudge and shove conditions, any observed effect can be attributed to the differing conditions themselves, rather than to their play sequence.

6.1.4 Method

Each participant played *Gekku Race* three times, once under each of three conditions in the above order (control, nudge, shove). Participants raced against each other in pairs. Each race lasted approximately 40-50 seconds.

Before beginning, participants were allowed to play the game briefly in the control condition to familiarize themselves with the game controls. Afterward, they played two races in each of the three conditions.

Participants played two races each time, so that the first race in each condition served as a practice round to get accustomed to the condition, and the second race was the one that would be analyzed. This was to give the players time to become accustomed to the condition before data was collected.

Between races, participants were required to wait at least four minutes before starting the next condition, even if they had already completed the questionnaires. This ensured participants
were not fatigued for the second race. As with PlaneGame, resistance on the exercise bikes was set to its lowest setting, again to reduce the risk of excessive fatigue.

6.2 Results

We first summarize the key results of this study. Detailed results can be found in Sections 6.2.1 Performance Metrics, and 6.2.2 Questionnaires, and their implications are explored in Section 6.3 Discussion.

The shove condition was more effective than the control or nudge conditions at getting players to slow down. No difference was seen between control and nudge.

Asked if the game made it clear when they were pedaling too fast, participants rated the nudge condition as more clear than control, and shove as more clear than both the control and nudge conditions.

Players found it more frustrating to be told to slow down in the shove condition than in either the control or nudge conditions. No significant difference in frustration level was reported between control and nudge.

No significant differences were found in IEQ immersion scores between any two conditions.

6.2.1 Performance Metrics

A one-way repeated measures ANOVA comparing players’ average cadence across the second race in each condition showed a significant difference between populations; Wilks’ Lambda=.435, F(2,18)=18.30, p<.001. Post hoc comparisons via paired samples t-tests with Bonferroni correction showed that average cadence in the control condition (M=84.02, SD=13.26) was not significantly different than in the nudge condition (M=84.53, SD=15.83); t(19)=.256, p=.801. Average cadence in the shove condition (M=69.09, SD=5.81) was
An RM-ANOVA comparing proportion of time players spent pedaling in excess of the target cadence showed significance; Wilks’ Lambda=.311, F(2,18)=25.14, p<.001. Paired t-tests showed no difference between the control condition (M=.724, SD=.356) and the nudge condition (M=.643, SD=.401); t(19)=1.17, p=.255. Proportion of time over target was lower in the shove condition (M=.215, SD=.240) than in the control condition, t(19)=6.24, p<.001; and in the nudge condition, t(19)=5.35, p<.001 (Figure 19 right).

Figure 19: Performance metrics for *Gekku Race*. Vertical bars show standard error. Horizontal hats indicate statistical significance at α = .05.

left: Average cadence (rpm).
right: Proportion of time above target.

significantly lower than in the control condition, t(19)=4.93, p<.001; and also lower than in nudge, t(19)=4.42, p<.001 (Figure 19 left).
6.2.2 Questionnaires

Participants’ responses to the questionnaires were analyzed using Friedman non-parametric tests. The test for scores on the IEQ showed significant variation in participant responses; \( \chi^2 = 6.74, p = .031 \). However, pairwise comparison through Wilcoxon signed-rank tests showed no significant differences between IEQ scores for the control condition (Mdn=118, IQR=102-129) and for the nudge condition (Mdn=115, IQR=106-133); \( Z = -1.01, p = .327 \). Likewise, there was no significant difference between the shove condition (Mdn=122, IQR=103-135) and either the control condition, \( Z = -1.71, p = .089 \); or the nudge condition, \( Z = -.48, p = .644 \).

In response to whether the game made it clear when to slow down, the Friedman test showed significance; \( \chi^2 = 20.26, p < .001 \). Since there is a definite sequence we expected these values to follow, pairwise comparisons were single-tailed Wilcoxon signed-rank tests, which showed that the control condition (Mdn=1.5, IQR=1-2) was considered less clear by the participants than the nudge condition (Mdn=4, IQR=2-5); \( Z = -2.73, p = .002 \). The shove condition (Mdn=5, IQR=4-5) was considered clearer than both the control condition, \( Z = -3.77, p < .001 \); and the nudge condition, \( Z = -2.37, p = .008 \) (Figure 20 left).

To the question of whether being told to slow down was frustrating, responses were significantly different; \( \chi^2 = 16.36, p < .001 \). Single-tailed Wilcoxon signed-rank tests showed no significant difference in frustration between the control condition (Mdn=1, IQR=1-3) and the nudge condition (Mdn=2.5, IQR=1-3) after Bonferroni correction; \( Z = -1.75, p = .048 \). However, the shove condition (Mdn=4.5, IQR=2-5) was seen as more frustrating than either the control condition, \( Z = -3.43, p < .001 \); or the nudge condition, \( Z = -2.46, p = .006 \) (Figure 20 right).
6.3 Discussion

Our first research question for the study was about the ability of integrated feedback techniques like ours to control player exertion in a faster-paced multiplayer game like Gekku Race. We hypothesized that using only gentle nudges to inform players they were expected to slow down would be much less effective than in PlaneGame, but that adding performance-affecting shoves would succeed in getting players to slow down.

As we expected, the nudge condition was no more effective than the control condition at getting players to slow down, as measured either by average cadence throughout a race or by proportion of time spent pedaling above the target cadence. The shove condition, however,
reduced both measures significantly, bringing average cadence down below the target, and reducing time over target to one third of what it had been in the nudge condition.

The interviews revealed that the nudge condition was not considered to be as clear as the shove condition, with four participants saying they didn't clearly understand the nudges, and one saying it was only partly clear, while all participants said they clearly understood the shove condition. According to one participant, “I didn’t for the [nudge condition], I didn’t really clue in, but the [shove condition], I got it.” They also reported feeling less motivated to slow down in the nudge condition (three definitely motivating, two a little motivating) than in the shove condition (11 yes, five a little).

In addition to being less clear and motivating, there appears to be another cause for the nudge condition's ineffectiveness. In three of the 10 pairs of participants, one participant responded to the nudge condition and began slowing, but their opponent did not. Because pedaling above the target cadence still gave an advantage in the nudge condition, the participant had to disregard the feedback to avoid being left behind. This supports our expectation that multiplayer games present an additional problem for nudge techniques, since it only takes one player disregarding the feedback to make the feedback ineffective.

We expected the control and nudge conditions would have similar immersion scores, as suggested by the previous study with PlaneGame, but that immersion in the shove condition would be lower because of the more obvious and punitive nature of shoves. Counter to expectation, however, we saw that average immersion was not reduced in the shove condition, but was instead statistically indistinguishable from the other conditions.

Despite the pairwise comparisons showing no difference in immersion, the Friedman test reported significant variation across the three conditions. A possible explanation for this has to do
with the shape of the shove condition immersion scores between participants. When we examined the differences between the immersion scores participants gave the three conditions, we found that two participants gave exceptionally low immersion scores to the shove condition compared to their responses for the other conditions. Both these players said the shove condition was their least favourite because being forced to slow down was frustrating, and that control was both their favourite and the most natural condition. One player said of the control condition, “It was a fair competition… basically we’re going at the speed that we pedal, so it’s just fair for us to compete according to the speed that you pedal but in the [shove condition]… I was trying to pedal really hard, but it’s not letting me, it just stopped there, so I don’t think it’s fair.” It appears, then, that while on average the shove condition does not reduce immersion, a minority of players are so frustrated at having to slow down their immersion suffers considerably. This suggests that care should be taken to use shoves as lightly as possible without sacrificing effectiveness, to avoid alienating these sorts of players.

The additional questions from the questionnaires lined up with responses during the interviews. While participants found the shove condition to be clearer than the nudge condition in telling them when to slow down, both conveyed the message, while the control condition did not. As suggested by the in-game data, however, the shove condition motivated a reduction in exertion, whereas the nudge condition generally did not. Unfortunately, though, this increased motivation was coupled with participants feeling more frustrated at having to slow down. Again, this suggests that shoves should be used only when necessary to provide motivation, in order to avoid frustrating players.

Participants’ reasons for liking or disliking conditions tended to be similar. The nine who liked the control condition mostly liked it because it allowed them to set their own pace, while the
three who disliked it thought the lack of incentive to go at any speed other than maximum made it less interesting. Two participants liked the nudge condition because they thought it seemed the most natural, but five disliked it because they thought telling them to slow down without requiring it was confusing or pointless. Five of the six who preferred the shove condition said it was because it placed a priority on in-game strategy rather than physical supremacy, but those who disliked it were frustrated at being forced to slow down.

As in the previous study, the nudge techniques (both nudges and shoves) were judged to be more natural than conditions without nudges: eight of the 20 participants selected the shove condition as the most natural, four selected the nudge condition, and three preferred anything but the control condition, versus only four who chose the control condition. The four participants who said they found the control condition to be the most natural had all chosen shove as their least favourite, so possibly these participants found the sense of frustration itself made the feedback feel unpleasant and unnatural.

Surprisingly, the shove condition was considered more natural than the nudge condition. This is possibly due to two factors. First, the shove condition picked up on the natural consequence of getting tired when running too quickly, while also delivering realistic consequences for ignoring it; several participants found the nudge condition unintuitive because it suggested slowing down without doing anything to make it necessary: “It was just confusing… how there was no consequence to them. It just seemed like an annoyance because… I wouldn’t slow down, and it was just a random sound that would be made.” Second, some players answered in terms of Gekku Race’s purpose as an exergame, and felt that the shove condition’s stronger feedback was a more natural fit to the goal of getting players to slow down: “Usually in a game
you’ll see the [control condition], but then if you were to consider… health consideration, I think the [shove condition] would be a better option.”

Finally, there were a total of six (of 20) participants who said, unprompted, that they would have preferred the shove condition if the target cadence had been set higher. We previously suspected this might be a factor in the tests with PlaneGame, and these responses seemed to confirm it. This suggests that players would feel less frustrated if we were trying to control actual over-exertion, since the target cadence would be higher. While these techniques ultimately are designed to guide players into doing something they don't want to do, having the additional incentive of slowing down from a genuinely tiring pace would presumably reduce frustration at being compelled to slow down. This is encouraging in determining the usefulness of shove techniques, as this sense of frustration appears to be the major barrier to their acceptance.
Chapter 7

Conclusion

In this thesis, we have reviewed the fields of exergames and of nudges, and the applicable design considerations of exertion and over-exertion, and immersion. We present and explain two exergames, PlaneGame and Gekku Race, using nudge-based feedback designed to reduce over-exertion without reducing immersion.

We have shown a design process for creating nudge-based feedback in digital games, in the form of four design principles, derived from our own experiences in creating nudge feedback for PlaneGame and Gekku Race. The four design principles are: natural integration, making the feedback feel like it's part of the game; comprehension, making sure the feedback is both clear and conspicuous; progression, making feedback become more severe as players get further from the desired behaviour; and multiple channels, making use of multiple forms of feedback to complement each other and help satisfy the other principles.

We have also presented a pair of studies testing the nudge feedback in PlaneGame and Gekku Race. These two studies show that, overall, nudge-based user interfaces are effective at controlling player behaviour in exergames, without any measurable reduction in player immersion. Additional insights can be gained from examining both studies together, in terms of discoveries about immersion and frustration, implications for future design, and applications for other games.
7.1 Immersion and Frustration

We saw in the proof-of-concept study with PlaneGame (Chapter 5) that gentle nudges with no direct consequences are still able to persuade players to slow down in some circumstances; however, in the context of Gekku Race's powerful drive toward higher exertion (Chapter 6), we saw such techniques become ineffective. It took the addition of shoves with direct gameplay effects, which caused over-exertion to go from a winning move to a losing move, to motivate players to slow their pace. This effectiveness, though, came at the cost of increasing the amount of frustration experienced by players who still wanted to go past the limits imposed by the game, with two of the participants even experiencing a resulting serious loss in immersion.

These findings suggest that the principle of progression is important for balancing effectiveness against player frustration. In situations or games where gentle nudges are sufficient, the nudges are preferable, and frustration can be kept to a minimum by delaying shoves until it's clear players will not respond to the nudges. Players who push the game can still receive the forceful feedback needed to protect them from over-exertion, while players who respond to nudges are spared the possible frustration of shoves.

In terms of immersion, we saw little difference between conditions in either game, as measured by the Presence and IEQ instruments. We expected that adding feedback to control player behaviour would a cause a loss in immersion, from a small loss in nudge conditions to a sizeable one in the textual and shove conditions. Instead, average immersion was unchanged between conditions. A possible cause is simply that immersion is difficult to measure (this possibility is part of what led us to change questionnaires between studies), but the data from participant interviews suggest other causes as well.
Despite seeing no significant difference in overall immersion scores between conditions, a Friedman test showed differences among all participants (in Gekku Race, and at 60rpm in PlaneGame). This implies that, while no condition was overall more or less immersive than others, there were still significant differences in how different participants experienced a sense of immersion.

However, we evidently succeeded in our specific goal of making feedback fit naturally into the game world: participants said our nudge-based conditions felt more natural than the unaltered control conditions. The discrepancy between this result and the lack of significance in either measure of immersion may indicate that, contrary to our expectations, they are not the same thing. This natural fit to the gameplay, which we have called natural integration, may possibly contribute to immersion, but differences in natural integration alone do not appear to change how immersed players feel.

7.2 Implications for Design

One common factor for all the tests we performed was that the target cadence was low, ranging from 75rpm (a moderately slow pace for most cyclists) down to 40rpm (slow enough to require deliberate effort not to exceed it). The low cadences were intended to ensure that all participants would see the feedback we were testing, but they resulted in participants being told they were pedaling too quickly, when in fact they were not over-exerting themselves. Nearly a third of participants in the second study said they would have preferred the nudges if the target cadence had been higher. This suggests these techniques might be even more effective, or at least less frustrating, if the threshold had been set higher. In the case of an exercise warm-up, however, the nudge techniques have no such advantage. Since the warm-up period is by nature temporary,
it may be prudent to clearly indicate when the player is still in the warm-up period, letting players eager to work harder know they will soon be able to do so.

Another potential addition to feedback used to control actual over-exertion concerns the nature of having a target value, be it cadence, heart rate, or something else. In PlaneGame, players appeared to oscillate around the target cadence in the nudge condition. The same effect might be responsible for Gekku Race players, on average, still spending over 20% of their time above the target cadence, despite having a mean cadence lower than the target. A way to combat this oscillation could be to have positive feedback which tells players when they’re within their target range of exertion. Techniques already used in exergames for incenting exertion, rather than reducing it, would likely be suited to this task. For example, Ketcheson et al.'s use of special in-game powers for players whose exertion level is within their target range [29].

In designing the nudge conditions for PlaneGame and Gekku Race, we concentrated mainly on what would serve the design best, rather than giving consideration to reusing components for the sake of development speed. Reusing components, though, like the screen-greying used in both games, can be an effective strategy on its own. This tactic falls under the principle of comprehension (Section 4.2). It would certainly be possible to develop a suite of feedback tools that can be slotted into an appropriate game, making design faster and offering the benefits of familiarity. However, close fits of feedback to game, like Gekku Race’s panting lizard, would be less feasible under such a scheme, risking a less natural fit to the game world. A single core natural concept customized to the game, supported by multiple reused concepts, might be the most effective way to combine the principles of natural integration, comprehension, and multiple channels for best effect.
Aside from the nudging interfaces, there were design challenges in other parts of the studies as well. In an early version of the textual condition for PlaneGame, both the target and current cadences were displayed at the top of the screen. This kept the text from interfering with the game itself, but pilot testing showed that having the current cadence located away from where players were normally looking caused other problems. Players would forget to look at the text because they were concentrating on the altitude of the plane, or even fail to notice that the text was there at all. Moving the current cadence display to immediately under the plane allowed players to keep track of their status without harming their ability to concentrate on the gameplay. This is consistent with the tactic of embedding crucial game information into the player’s avatar, as described by Stach et al. in “Improving recognition and characterization in groupware with rich embodiments” [52].

7.3 Applicability to Other Types of Games

Nudging has been used in existing digital games. Indeed, we leveraged this fact in the principle of comprehension (Section 4.2), taking advantage of standard game mechanics for their recognizability. In the design of our nudges in PlaneGame and Gekku Race, we evoked commercial games that use colour desaturation to indicate character damage [15, 22, 38], in order to make our own feedback more recognizable. An example of a shove in commercial games is used in the racing game series Mario Kart [39]. To encourage players to stay on the track, the game slows players' vehicles when they leave the track on a course. Consistent with our own recommendations, nudges accompany this shove to accentuate it, and include an audible shift in the engine's tone, along with the wheels kicking up grass or dirt. This thesis’ contribution of formalizing nudges in games, and evaluating their success, has the potential to help expand the
concept to further areas of game design, by establishing the effectiveness of nudges and providing guidance for how they can be designed.

The idea of nudges can also be applied to other areas where their potential may not be obvious. In addition to convincing players not to do things they are not supposed to do, nudges might be used to discourage players from even attempting actions disallowed by the game. For example, even in an open-world game, players are not able to go into areas that have no content. In many games, they are kept out by literally making the player character unable to walk past a certain point. But this sort of “invisible wall” can be frustrating to players expecting to be able to go through the opening they perceive. If the same no-go area were instead a bricked-off archway with a “Keep Out” sign posted on it, players would be nudged away from even trying to pass. The four nudge design principles are useful for such non-obvious cases, as they can be applied preventatively to any interaction with the player to help maintain the natural logic of the game world.

This thesis represents the first attempt to consider specific principles for how nudges may be designed and where they may best be used. For example, the principles we developed allow us to characterize the strengths of Mandryk et al.’s work in using visual overlays to convert commercial games into biofeedback games [34], while suggesting opportunities for further improvement. The biofeedback overlays are triggered by negative changes in EEG readings that gradually block vision using thematically appropriate graphical effects (for example, mud splatters on the screen in a dirt-biking game). Seen through the lens of the design principles presented in Chapter 4, we can see that the overlays use two of the design principles to good effect. First, they use natural integration, by choosing thematic appearances for each game's overlays; second, they employ progression, with the overlays obstructing vision more and more
as the negative player input continues. However, the principle of comprehension is not entirely fulfilled: while the progressive overlays are conspicuous, it's not clear without being told that they are linked to the player's emotional state. Using the principle of multiple channels, we can imagine adding audio feedback to suggest dwindling emotional control: perhaps with a racing pulse, or the sound of static.

The above examples are only a sampling of the possibilities afforded by the concept of nudges. We are confident that game designers can apply the principles of nudge design to any number of occasions where games interact with players.

7.4 Analysis of Design Principles

The four design principles described in Chapter 4 were derived from our experiences in designing nudge feedback for *PlaneGame* and *Gekku Race*. The principles are: natural integration, comprehension, progression, and multiple channels. With the nudge-based feedback systems tested in the two studies presented in this thesis, the opportunity now exists to discuss how well the principles reflect the aspects of design that were most effective.

The principle of natural integration states that feedback to players should be integrated into the fiction of the game world. This is to reduce dissonance between gameplay and the player's physical reality, to prevent losing the sense of the game being for play's sake. The two games tested fulfilled this principle by providing in-universe explanations for excessive cadence having negative consequences. In *PlaneGame* the plane's engine overheats, causing smoke and fire to billow out and a rattling sound to replace the normal engine noise. In *Gekku Race* the lizard avatar gets tired and pants for breath, occasionally collapsing in the case of the shove condition.
In the studies, players responded positively to these integrated concepts. Players found that not only was the feedback not disruptive to the naturalness of the game world, but they reported that the nudge and shove conditions were also more natural than the unmodified control conditions. Players who favoured the nudge or shove conditions commonly gave natural integration as the reason. However, some players found the nudges to be somewhat distracting, possibly due to their abrupt onset. Further work in naturally-integrated feedback should bear this in mind, and try to find ways to avoid distracting players by even the small amount reported.

The principle of comprehension states that nudge feedback must be clearly understood to be useful. It has two components: clarity, and conspicuity. The nudges must be conspicuous enough to be noticed, and then clear enough to be understood. Conspicuity was easily achieved in PlaneGame and Gekku Race, with both visual and audio cues beginning as soon as players passed the target cadence.

Clarity came through two strategies. First, feedback implicitly suggested that a problem existed, and that it was connected to excessive speed. PlaneGame's rattling and engine overheating are both consequences that can be expected of a plane being pushed past its mechanical tolerances, and it is logical that the lizards of Gekku Race become tired and out of breath from running too quickly. Second, feedback drew from expectations that players may have had from other video games. Both games use greying effects on the screen, which is also used in commercial games to indicate a health threat to the player’s character.

During early testing of our nudge feedback in PlaneGame prior to the study, we encountered an unexpected obstacle to clarity. Testers only responded to the nudge feedback when they had been told we were investigating ways of getting players to slow down. However, when testers who did not slow down were asked afterward what they thought the smoke and fire
meant, they reported that they had perceived that they were supposed to slow down. These players understood what the game was telling them to do, but nevertheless did not slow down. This reluctance to act on the nudge may mean that even clear feedback needs a brief tutorial or explanation before players will feel confident about responding to it. To address this issue, during the study, players were primed to expect that the game might guide them away from overexertion.

This tactic of using similar feedback as in other games can be expanded, to benefit from being more recognizable to players of those games. If nudge-based feedback became more common in video games, then these games could begin borrowing from each other for increased clarity by analogy. This would create a standard library of effects suitable for different nudge contexts, speeding production time.

Conspicuity did not seem to be a major problem during design or during the studies for PlaneGame or Gekku Race, but should be kept in mind in case of more complicated or otherwise "busy" games. Also, if attempts are made to make feedback more natural by reducing distraction, it will be important to maintain a balance between being non-distracting but still conspicuous.

The principle of progression states that nudges should get more severe as players deviate further from the behaviour required of them. This allows players who already understand and follow the feedback to receive a gentle reminder if they accidentally cross the threshold, while players who are not following the feedback can be pushed harder until they do. This is especially valuable with shoves, since the harsher consequences should be reserved for when they are actually needed. PlaneGame uses progression by having the smoke get thicker and the engine rattling get louder over time, as well as by introducing fire and reddening of the screen at sufficiently high levels of Severity. Gekku Race increased the frequency of the lizards' gasps for
breath, and in the case of the shove condition did not cause a collapse until an ordinary pant had already happened.

The studies suggested that the principle of progression is important to prevent frustration from building, especially in the case of Gekku Race's shoves. If anything, progression may have been too rapid in the shove condition. We were concerned that, if the shoves took too long to appear, players would realize they could pedal above the target for several seconds to gain a brief speed advantage. However, we did not see such behaviour. Most players who experienced the lizard collapsing slowed down as soon as the nudges appeared, presumably to avoid any risk of encountering the shoves again. A more gradual progression might have helped reduce the frustration that some players experienced in the shove condition of Gekku Race.

The principle of multiple channels states that different channels of feedback should be combined in the creation of nudges, such as visual, auditory, haptic, and direct gameplay effects. Doing so allows the channels to combine and complement each other. PlaneGame and Gekku Race use multiple types of visual feedback, both particles (PlaneGame smoke and fire, Gekku Race puffs of breath) and overlays (greying screen), and both games also use audio feedback. Gekku Race also uses a direct gameplay channel for its shoves, stopping the lizard's movement for a second.

The principle of multiple channels was a contributor to the ease of increasing conspicuity for the nudges. Audio and visual feedback combined to make the nudges very difficult to ignore. Combining the direct gameplay shoves with audio and visual nudges in Gekku Race also helped make the shoves seem natural, since players could see and hear the in-world explanation for why the lizard had stopped running.
7.5 Future Work

As mentioned in Chapter 3, the original version of Gekku Race is part of an existing suite of exergames called Liberi [23, 30]. Part of the motivation for investigating nudges to control over-exertion was to be able to add such feedback to Liberi. Given the success of the nudges and shoves in Gekku Race, nudge feedback should likewise be successful in convincing players not to over-exert in the other games of Liberi. Adding nudges to the other games would also allow them to be tested in more longitudinal studies, in addition to the shorter lab-based studies described in this thesis. As an example, here is a possible implementation of nudge feedback in another Liberi game, Biri Brawl.

*Biri Brawl* (Figure 21) is a multiplayer fighting game in which players control microbes called 'biris' from an overhead perspective. The biris swim through a circular arena and attack each other. When a biri's health drops to zero, the player that delivered the final attack receives a

![Figure 21: Biri Brawl.](image)
point. A defeated biri becomes invulnerable to harm for a few seconds, and then enters play again at full health. Players compete to be the first to reach some number of points, set before the game begins.

The input devices of *Biri Brawl* are the same as in *Gekku Race*. Players pedal a stationary bicycle to move and to power their attacks. Players use a cordless video game controller to direct their biris, using an analog stick to steer and a button press to punch.

To implement anti-over-exertion nudges in *Biri Brawl*, we begin by considering the principle of natural integration. Since *Biri Brawl* takes place underwater, it does not make sense for tired biris to gasp and let out puffs of air. Instead, biris could begin to sweat when their players are pedaling too quickly, with cartoon sweat droplets flying out from their bodies.

Under the principle of multiple channels, we want to add some audio feedback to the visuals of the sweat droplets. Sweating does not have any obvious sound associated with it, but we can substitute another sound evocative of extreme exertion. A pounding heartbeat, distorted as though underwater, would suit the setting of the game.

With both visual and auditory cues that clearly suggest excessive exertion, the principle of comprehension is fairly well represented. However, we can easily add a third form of feedback by re-using the same greyscale overlay as in *Gekku Race*. This increases clarity by analogy to other games, not only to commercial games as with *Gekku Race*, but also by tying in to *Gekku Race* itself. Shared nudges between different *Liberi* games will help players comprehend what they are supposed to do, even if they have not played a particular game before, because they will have already learned the meaning of the nudge during previous games.

Finally, the principle of progression requires that the nudges become more severe as players continue to over-exert. To be intensified, the nudges can simply become stronger: more
and bigger sweat droplets, faster and louder heartbeat, greyer screen. Additionally, as we saw in *Gekku Race*, shoves are very powerful to have at the top end of the progression. Since biris have health bars, one possible shove would be for the biri to begin gradually losing health. However, this might not be conspicuous enough. Instead, the biri could take chunks of damage as though it were being attacked, synchronized with the sweat and heartbeat. This way, the biri is literally hurting itself because the player is going too fast, encouraging the player to slow down.

Alternatively, the biris could collapse intermittently as the gekku lizards do, interrupting the player's ability to fight, and introducing another instance of clarity by analogy.

### 7.5.1 Measurement of Exertion

As stated in Chapter 3, cadence is not the only possible measure of exertion for an exergame. We used cadence rather than heart rate in the research described in this thesis because of the latency between changes in exertion and changes in heart rate [12, 50]. However, heart rate is a better reflection of whether an exergame player is over-exerting, because much of the risk from over-exertion is from overtaxing the cardiorespiratory system.

A better solution for measuring exertion might be to combine both heart rate and cadence measurements. Over-exertion would be defined according to heart rate, reflecting the medical reality. However, the player would be able to quickly turn off the nudges and return to normal gameplay by reducing their cadence.

This hypothetical system would require an understanding of how much cadence needs to be lowered to drop heart rate to an acceptable value, and rapidly enough for safety. Unfortunately, it is not obvious how to determine what cadence is low enough to have this effect. More research and testing is required to create this sort of hybrid exertion metric.
Cadence is not an optimal measure of exertion for another reason. Technically, since we want to measure how much work the player is performing through their legs, and not simply how fast their legs are moving, we want to measure power, not cadence. Power can be approximated through cadence, assuming a fixed resistance, because increases or decreases in cadence necessarily imply increased or decreased pedaling power. However, it would be better to allow players to adjust the resistance of the exercise bicycle as they play, either upward as they warm up, or downward as they tire.

There are devices available that can be attached to an exercise bike to measure power directly, allowing resistance to be changed as needed. Examples of such devices include the Garmin Vector 2S\(^1\) and the PowerTap P1\(^2\) Pedals. Unfortunately, these devices are prohibitively expensive, costing in excess of 600 U.S. dollars. We expect that power-measuring devices will become more affordable over time, but for now, measuring cadence is more affordable.

7.6 Summary

An exergame is a helpful fitness tool only if people are willing to play it in the first place, and then use it in a way that supports their fitness needs and goals. While there is an ample body of research demonstrating methods to ensure exergame players meet target levels of exertion, there has been little research into ensuring that players do not exceed those levels.

Using the concept of nudges, we added feedback to two cycling-based exergames – one, Gekku Race, more fast paced, and the other, PlaneGame, slower paced – that pushed players toward slowing down whenever they passed a set exertion limit, while making an effort to avoid disrupting their sense of immersion. The design strategies used to add nudge feedback to the two

\(^2\) http://www.powertap.com/product/powertap-p1-pedals
games were isolated and written into a list of four design principles that can be followed to guide the creation of such feedback systems: natural integration, comprehension, progression, and multiple channels.

To test our feedback systems, each game was tested with three different conditions. Players of PlaneGame played the game with and without our nudge-based feedback mechanisms, and with a third condition using simple text to present the same information. In Gekku Race, participants played the game without feedback, and with two versions of our feedback system: one that made players slow down when going too fast by harming their ability to successfully play the game (shove), and another that did not (nudge).

Results of the two studies showed that the nudge-based feedback we designed was effective at persuading players to reduce their levels of exertion and, indeed, was as effective as a more traditional text-based interface. In slower-paced games, gentle nudges are effective, but when players are strongly motivated to work hard, shoves are required to make them slow down. There was no measurable difference in immersion between different conditions, but the nudge condition was considered to be a natural fit to the game environment.

When designed carefully, nudge techniques are highly effective at influencing player behaviour, and seem to cause no drop in immersion. We propose that nudge-based designs which include elements from our four design principles will create good feedback systems that motivate players to regulate their own actions according to the needs of the game, while keeping the game world consistent and not intruding on it with interface components that do not feel natural to players.
References


[22] D. Hall, DayZ, Game (Windows), Bohemia Interactive, Prague, Czech Republic, 2013.


83

[37] Namco, Prop Cycle, Game (arcade), Namco, Ōta City, Tokyo, Japan, 1996.


[39] Nintendo EAD, Mario Kart 8, Game (Wii U), Nintendo, Kyoto, Japan, 2014.

[40] Nintendo EAD, Wii Sports, Game (Wii), Nintendo, Kyoto, Japan, 2006.


[42] Nintendo EAD, Wii Fit, Game (Wii), Nintendo, Kyoto, Japan, 2007.


84


[56] Ubisoft, Your Shape: Fitness Evolved, Game (Xbox 360), Ubisoft, Rennes, France, 2010.


Appendix A

*PlaneGame* Statistics Tables

<table>
<thead>
<tr>
<th>Condition (Mean ±SD)</th>
<th>Comparison</th>
<th>t (23)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Cycling Cadence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (57.84 ±12.67)</td>
<td>Nudge</td>
<td>6.53</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nudge (41.94 ±2.55)</td>
<td>Textual</td>
<td>2.11</td>
<td>0.046</td>
</tr>
<tr>
<td>Textual (43.24 ±3.22)</td>
<td>Control</td>
<td>5.64</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Time Over Target Range</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (45.95 ±23.13)</td>
<td>Nudge</td>
<td>8.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nudge (6.50 ±6.37)</td>
<td>Textual</td>
<td>1.60</td>
<td>0.123</td>
</tr>
<tr>
<td>Textual (9.86 ±14.40)</td>
<td>Control</td>
<td>6.41</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Game Score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (24.13 ±1.15)</td>
<td>Nudge</td>
<td>3.00</td>
<td>0.006</td>
</tr>
<tr>
<td>Nudge (22.79 ±2.13)</td>
<td>Textual</td>
<td>2.82</td>
<td>0.010</td>
</tr>
<tr>
<td>Textual (24.13 ±1.42)</td>
<td>Control</td>
<td>0.00</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 4: Pairwise t-tests for *PlaneGame*, 40 RPM.

<table>
<thead>
<tr>
<th>Condition (Mean ±SD)</th>
<th>Comparison</th>
<th>t (23)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Cycling Cadence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (67.47 ±10.65)</td>
<td>Nudge</td>
<td>3.99</td>
<td>0.001</td>
</tr>
<tr>
<td>Nudge (59.37 ±1.84)</td>
<td>Textual</td>
<td>2.09</td>
<td>0.048</td>
</tr>
<tr>
<td>Textual (60.33 ±1.45)</td>
<td>Control</td>
<td>3.40</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Time Over Target Range</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (27.79 ±24.52)</td>
<td>Nudge</td>
<td>5.06</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nudge (3.65 ±3.00)</td>
<td>Textual</td>
<td>1.49</td>
<td>0.150</td>
</tr>
<tr>
<td>Textual (2.29 ±3.30)</td>
<td>Control</td>
<td>5.30</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 5: Pairwise t-tests for *PlaneGame*, 60 RPM.
<table>
<thead>
<tr>
<th>Condition (Median, IQR)</th>
<th>Comparison</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Do you feel motivated to slow down when you’re pedaling too quickly?”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (5.5, 4-6)</td>
<td>Nudge</td>
<td>-3.45</td>
<td>0.001</td>
</tr>
<tr>
<td>Nudge (7, 6-7)</td>
<td>Textual</td>
<td>-1.36</td>
<td>0.175</td>
</tr>
<tr>
<td>Textual (6.5, 6-7)</td>
<td>Control</td>
<td>-2.95</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table 6: Wilcoxon signed-rank tests for *PlaneGame*, 40RPM. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values.

<table>
<thead>
<tr>
<th>Condition (Median, IQR)</th>
<th>Comparison</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Do you feel motivated to slow down when you’re pedaling too quickly?”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (4.5, 3-6)</td>
<td>Nudge</td>
<td>-3.69</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nudge (7, 6-7)</td>
<td>Textual</td>
<td>-1.83</td>
<td>0.067</td>
</tr>
<tr>
<td>Textual (6, 5-7)</td>
<td>Control</td>
<td>-2.79</td>
<td>0.005</td>
</tr>
</tbody>
</table>

*Presence Questionnaire Score*

<table>
<thead>
<tr>
<th>Condition (Median, IQR)</th>
<th>Comparison</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (68, 63-73)</td>
<td>Nudge</td>
<td>-0.56</td>
<td>0.594</td>
</tr>
<tr>
<td>Nudge (69.5, 62-73)</td>
<td>Textual</td>
<td>-1.31</td>
<td>0.196</td>
</tr>
<tr>
<td>Textual (68, 63-72)</td>
<td>Control</td>
<td>-1.20</td>
<td>0.237</td>
</tr>
</tbody>
</table>

Table 7: Wilcoxon signed-rank tests for *PlaneGame*, 60RPM. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values.
Appendix B

*Gekku Race Statistics Tables*

<table>
<thead>
<tr>
<th>Condition (Mean, ±SD)</th>
<th>Comparison</th>
<th>t (19)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Cycling Cadence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (84.02 ±13.26)</td>
<td>Warning</td>
<td>0.256</td>
<td>0.801</td>
</tr>
<tr>
<td>Warning (84.53 ±15.83)</td>
<td>Punishment</td>
<td>4.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Punishment (69.09 ±5.81)</td>
<td>Control</td>
<td>4.93</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Proportion of Time Over Target Cadence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (0.724 ±0.356)</td>
<td>Warning</td>
<td>1.17</td>
<td>0.255</td>
</tr>
<tr>
<td>Warning (0.643 ±0.401)</td>
<td>Punishment</td>
<td>5.35</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Punishment (0.215 ±0.240)</td>
<td>Control</td>
<td>6.24</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 8: Pairwise t-tests for *Gekku Race*. 
<table>
<thead>
<tr>
<th>Condition (Median, IQR)</th>
<th>Comparison</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Immersive Experience Questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (118, 102-129)</td>
<td>Warning</td>
<td>-1.01</td>
<td>0.327</td>
</tr>
<tr>
<td>Warning (115, 106-133)</td>
<td>Punishment</td>
<td>-0.48</td>
<td>0.644</td>
</tr>
<tr>
<td>Punishment (122, 103-135)</td>
<td>Control</td>
<td>-1.71</td>
<td>0.089</td>
</tr>
<tr>
<td>“The game made it clear when I was pedaling too fast”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (1.5, 1-2)</td>
<td>Warning</td>
<td>-2.73</td>
<td>0.002</td>
</tr>
<tr>
<td>Warning (4, 2-5)</td>
<td>Punishment</td>
<td>-2.37</td>
<td>0.008</td>
</tr>
<tr>
<td>Punishment (5, 4-5)</td>
<td>Control</td>
<td>-2.73</td>
<td>0.002</td>
</tr>
<tr>
<td>“It was frustrating when I was trying to win the race but the game was telling me to slow down”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (1, 1-3)</td>
<td>Warning</td>
<td>-1.75</td>
<td>0.048</td>
</tr>
<tr>
<td>Warning (2.5, 1-3)</td>
<td>Punishment</td>
<td>-2.46</td>
<td>0.006</td>
</tr>
<tr>
<td>Punishment (4.5, 2-5)</td>
<td>Control</td>
<td>-3.43</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table 9:** Wilcoxon signed-rank tests for *Gekku Race*. Interquartile Range (IQR) is the lower and upper quartiles of data, delineating the middle 50% of values. IEQ Quartiles are rounded to whole numbers.