#### Entertainment Computing 19 (2017) 83-100

Contents lists available at ScienceDirect

**Entertainment Computing** 

journal homepage: ees.elsevier.com/entcom

# Nudging and shoving: Using in-game cues to guide player exertion in exergames

# Adrian L. Jessup Schneider\*, T.C. Nicholas Graham

School of Computing, Queen's University, 99 University Avenue, Kingston, Ontario K7L 3N6, Canada

#### ARTICLE INFO

Article history: Received 14 May 2016 Revised 21 December 2016 Accepted 20 January 2017 Available online 21 January 2017

Keywords: Exergames Game design Human-computer interaction Nudges Persuasive interfaces

# ABSTRACT

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 paper we apply the concept of nudges to keeping players from going too quickly, and describe the effective use of nudges through a set of four design principles. We demonstrate two exergames using nudges to persuade players to slow down. We show that nudges are highly effective in games where players are not strongly incented to work hard. We also show that, even in high-energy games, adding negative consequences to the nudges, creating "shoves", maintains the nudges' power. Players reported that the nudges and shoves motivated them to slow down, and fit naturally into the games.

© 2017 Elsevier B.V. All rights reserved.

# 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. To gain the benefits, however, exergame players must stay within a healthy range of effort, neither under- nor over-exerting. In this paper, 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 pedaling 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 guidelines for the design of nudge- and shovebased feedback systems, highlighting the value of using multiple channels and of taking advantage of players' knowledge of realworld behaviours and standard game mechanics.

The foundational premise of exergames is that many people do not find exercise intrinsically motivating, but do 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 [1–3].

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 [4]. Such overexertion 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 paper, 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, such a blatant reminder that the game is meant to get them to exercise instead of just having have fun might reduce the value of embedding exercise into a game in the first place, risking the 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.

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. These nudges are produced within the game, by adapting the game's mechanics and sensory presentation. For example, the player's avatar in a race begins to





2

Entertainme Computing

<sup>\*</sup> Corresponding author. *E-mail addresses:* adrian.schneider@queensu.ca (Adrian L. Jessup Schneider), nicholas.graham@queensu.ca (T.C.N. Graham).

visibly gasp and puff when the player is pedaling too hard. The nudges we designed convinced players to slow down, while being perceived as a more natural fit to the game than even the absence of any sort of feedback. In this paper 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; as with the gentle nudges, the shoves felt like a natural part of the game to players. Nudges were originally described by Thaler and Sunstein in the book *Nudge* as a means of helping governments achieve public policy [5].

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. Using the concepts developed and employed in these designs, we present a set of four design guidelines 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.

We tested our modified games under different conditions in a pair of studies. Our test of *PlaneGame* compared the nudge feedback against simple explicit textual feedback, and against no feedback at all. With the faster-paced *Gekku Race*, we found that gentle nudges were not sufficient to keep players from going too quickly, and so we employed shoves. The results of these tests showed that nudges are just as effective at telling players when to slow down as direct textual data. They also show that shoves can persuade players whose desire to over-exert is too strong for gentler nudges, by making excessive effort disadvantageous within the game.

We expected that our strategy of designing feedback to fit naturally into the game should be closely linked to immersion, the feeling of being engaged with a game. However, while participants noted and expressed in interviews a preference for natural integration of nudge-based feedback, participants indicated 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.

# 2. Background

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 is Konami's *Dance Dance Revolution* (DDR), in which players dance in time with music by hitting buttons with their feet on a purpose-built dance pad, cued by the game's video component [6]. An example of an exergame that uses commercial rather than custom hardware is *Growl Patrol*, in which audio cues help players find small animals in need of rescue while avoiding a hungry tiger [7]. *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.

A physical input device used by many games is a stationary exercise bicycle attached to a computer, so the player's pace can be read. For example, the *PaperDude* game uses both a recumbent exercise bike and a virtual reality headset to create a simulation of delivering newspapers on a paper route [8]. In this example, both the player and their character are cycling. But the use of the exercise bike can also be abstractly applied to power any kind of avatar movement, as in Brehmer et al.'s *Racing Game* [9]. The games investigated in this paper, *PlaneGame* and *Gekku Race*, fall into this category, as discussed in Section 3.

#### 2.1. 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 min of moderate intensity exercise five times per week, or 20 min of vigorous exercise three times per week [10,11]. Peng et al. [12] and Biddiss and Irwin [13] 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 [14]. 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 [15]. 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 [16,4].

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 [17]. Even after conducting a warm-up, exercising too vigorously increases the risk of coronary events [18].

Efforts to increase player exertion levels in exergames, however, 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 [4]. In this paper, we focus on the problem of controlling such over-exertion.

#### 2.2. 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 for 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 [19]. 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 [20]. In "Measuring presence in virtual environments: a presence questionnaire", Witmer and Singer consider immersion a necessary component of presence [21]. 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 paper [22].

The theory of immersion also provides a possible explanation for overexertion in exergame players. The dual flow model considers flow to apply in two different ways to exergames: the psychological flow of gameplay, and the physiological flow of how effective exercise is [2,23]. Under this model, just as psychological flow is produced by the balance of the game's challenge and the player's skill, physiological flow represents a balance of the exercise's intensity and the player's fitness. Overexertion can therefore be seen as a symptom of psychological flow occurring at a higher level of exertion than is appropriate to reach physiological flow. Under this view, nudges can be used to disrupt psychological flow, but in such a way that the disruption can be ended by lowering exertion, thus achieving both forms of flow simultaneously.

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 Sections 6 and 7, adding feedback to control exertion had little impact on immersion.

# 2.3. Nudges

In their book *Nudge*, Thaler and Sunstein introduce the term "nudge" to describe the use of indirect suggestions that guide people toward a desired behaviour. Key to nudges is that they hint rather than prescribe, preserving the person's ability to choose [5]. The original definition of a nudge introduced by Thaler and Sunstein is:

Any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives.

In this paper, we use a shortened version of the definition written by Hausman and Welch [24] and adapted by Hansen and Jespersen [25]:

A nudge is any attempt at influencing behaviour in a predictable way without forbidding any previously available courses of actions or making alternatives appreciably more costly.

A shove, as defined by the United Kingdom's Local Government Association (LGA) [26] is a stronger variant of a nudge that restricts, but does not eliminate, choice.

# 2.3.1. Examples of nudges

As an example of the concept of nudges, Thaler and Sustein cite a tactic used by an airport in Amsterdam to improve the cleanliness of their washrooms [5]. 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 [27], encouraging a reduction in electricity consumption by delivering pamphlets to households comparing their energy consumption to that of their neighbours [28], and serving food in smaller dishes to increase the portion's apparent size and decrease consumption [29].

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 cyber-security context to encourage university computer users to use safe networks [30]. 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 [31]. 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 behaviours in positive ways." [32].

# 2.3.2. Nudging in digital games

There are several channels through which nudges can be applied to digital games. As seen in the cyber-security example in Section 2.3.1 [30], sensory effects can serve as nudges, including both visual and auditory feedback and, for some games, touch feedback in the form of controller rumble. Visual feedback might include adding visual cues such as particle effects, modifying the whole display with a visual effect, or changing or introducing animations. Auditory nudges may include adding or changing sounds or music.

An approach unique to games is to nudge directly through game mechanics. Given that game mechanics form the foundation of the world in which the game takes place, nudging through mechanics has the potential to be very powerful. For example, the player's controls might temporarily become reversed – press left to go right, press up to go down, etc. – in response to colliding with the wall of a maze. While the player can still play the game, the confusion caused by the reversed controls provides a strong incentive to keep away from the walls. Game mechanic nudges are able to change the environment so that even such fundamental tasks as moving around can be made easier or harder depending on whether players are exhibiting behaviours desired by the developers.

In the case of exergames, there is an additional complication caused by the fact that the game responds to both in-game actions (such as the game avatar's movement) and out-of-game actions (such as pedaling a stationary exercise bicycle). A racing exergame might want to use nudges both to keep players on the road, and to avoid pedaling too quickly to the point of injury. Such a game would therefore need to ensure that players are able to clearly distinguish different nudge systems so they know what set of behaviours needs adjusting.

# 2.3.3. Examples of nudges in games

Using nudges in games is not a novel idea. Some example of nudges are well-established in commercial video games. One example is in how games represent to the player that their avatar is low on health. The traditional option is to have an explicit health display, such as a bar showing proportion of health remaining or a numeric display of exact remaining health. Increasingly, however, games such as Tomb Raider 2013 [33], DayZ [34], and the Uncharted series [35] use visual (muted colours, "tunnel vision") and auditory (pounding heart beat, ragged breathing) cues to inform players when their health is low. As opposed to the tradition hit point display, this is a nudge because it provides hints at wise behaviour (getting cover, getting healed) while fitting seamlessly into the game. The player "feels" through audio and visual feedback that their health situation is dangerous, rather than analytically deriving that information from a user interface display. Also, the nudge is not prescriptive: while the game encourages the player to seek cover from attacks, a player has the freedom to choose to continue fighting, perhaps being willing to risk dying, if they know the current enemy is close to defeat.

An example of a shove in commercial games is used in some racing games, for example in *Mario Kart 8* [36]. If players drive off the track, sensory effects are used to encourage players to return to the track, including an audible shift in the engine's tone, the wheels visibly kicking up grass or dirt, and haptic feedback in

the controller suggesting a rough driving experience. In addition to this pure nudging behaviour, the game slows players' vehicles when they leave the track. The added force classifies this as a shove because, while the player could choose to continue driving offroad, there is a penalty for doing so. The player will likely wish to return to the course unless there is a strong reason to continue weathering the shove, perhaps because they are taking an off-road shortcut.

A third commercial example is Pokémon Go [37], an augmented reality mobile game in which players catch virtual monsters called Pokémon. Pokémon Go's main means of encouraging players to perform desired actions is through gamification: the use of game design elements in non-game contexts [38]. In the case of Pokémon Go, this gamification takes the form of offering in-game rewards for behaviours like visiting certain real-world locations. For example, when a player walks to a PokéStop, they receive rewards such as Poké Balls and potions. Gamification techniques are distinct from nudges, in that they guide player behaviour through explicit rewards rather than implicit suggestions. However, Pokémon Go also includes a form of nudge in its display of nearby Pokémon. The display encourages players to continue walking in order to find where these Pokémon are hiding, but players are free to decide they do not want that particular Pokémon, or do not feel like walking that far.

Existing exergame research contains examples of how nudges might be applied to exergames. In '*Ere be Dragons* [39], 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* [40], 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 [41]. Mandryk et al. use a system of visual overlays to turn commercial games into biofeedback games [42]. 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).

While these examples help illustrate how nudges can be used in games, there has been, however, no comprehensive study of the effectiveness of nudges in games, or how they might be best used in games.

# 2.3.4. Nudge theory

In social policy, there has been some debate as to whether employing nudges is ethical. In particular, nudge-based policies can be perceived as being "paternalistic" [24,25]: restricting the freedom of the public in service of their best interest, as judged by the policy maker. This point was raised in the original *Nudge* [5], as well as in scholarly works such as "Social Nudges: Their Mechanisms and Justification" [43] and "Debate: to Nudge or not to Nudge" [24]. The critical argument in favour of nudges is that, by definition, nudges maintain the subject's ability to choose for themselves, reframing options rather than forbidding possibilities.

Nudges can also be viewed through the perspective of Herzberg's motivation-hygiene theory of job satisfaction [44], also called two-factor theory. In two-factor theory satisfying factors, or "motivators", are seen as positive contributors to job satisfaction, while dissatisfaction is due to a lack of "hygiene" factors that prevent irritations from diminishing the work experience. While two-factor theory was originally formulated in a workplace context, in games hygiene factors can also be seen as those elements which are necessary but not sufficient for enjoyment [45]. In two-factor theory, a nudge can be seen as a temporary irritation, for which the appropriate hygienic response is within the player's power to perform, which encourages them to do so.

In the exergame domain, there are strong health reasons for guiding players away from overexertion as presented is Section 2.1. However, using nudges allows the player ultimately to retain the ability to choose for themselves whether to go physically all-out, or to take the guidance of the nudges and slow down to a safe level of exertion.

There also exist frameworks for classifying different kinds of nudges. The Local Government Association (LGA) classifies intervention techniques on a spectrum of lower to greater levels of intervention [26], ranging, from hugs (strong positive incentives), through nudges and shoves, to smacks (eliminating choices entirely). Gamification techniques, such as in Pokémon Go [37] are examples of hugs, offering rewards for performing certain actions. A smack in terms of digital games is the complete forbiddance of an option, such as the threat of revoking achievements or other punitive action for players caught cheating.

Another framework introduced by Hansen and Jespersen [25] classes nudges into one of four quadrants, separated by two axes. Their first axis is the transparency of the nudge to the nudge subjects, on the basis that a more transparent nudge is less manipulative. The second axis is what the authors term "type 1" versus "type 2" nudges, where type 2 nudges aim to influence active, reflective thinking (i.e., choices), while type 1 nudges aim to influence automatic behaviours made without thinking. The type 1 versus type 2 axis can be seen as analogous to the LGA's spectrum of level of intervention, with type 1 nudges corresponding to lower intervention, while type 2 nudges involve greater intervention that may, in fact, tip over into being shoves.

Since shoves involve greater levels of intervention and so come closer to paternalistic manipulation than do nudges, it is important not to use a shove when a nudge will accomplish the goal. For example, Benhassine et al. investigated the use of cash transfers conditional upon being used for a certain purpose (in this case, education) [46]. The authors found that by providing cash unconditionally and merely labeling what its intended use was, recipients used the money as desired nearly as often as when they were explicitly required to.

The idea of nudging bears similarity to the operant conditioning concepts of reinforcement and punishment [47]. Reinforcement refers to consequences that encourage a behaviour by adjusting stimuli when the behaviour is performed, either through providing a pleasant stimulus (positive reinforcement) or removing a negative stimulus (negative reinforcement). Punishment, oppositely, discourages behaviour by providing an unpleasant stimulus (positive punishment) or by removing a pleasant stimulus (negative punishment). Nudges can operate using reinforcement or punishment. For example, the indirect visual and auditory warnings of low health used by some games [33-35] is a form of negative reinforcement, as they encourage players to seek cover or heal through the promise of removing the negative cues. However, some nudges cannot sensibly be characterized in terms of operant conditioning. The urinal fly mentioned in [5], for example, functions without being either pleasant or unpleasant and without coming into or leaving existence.

Nudges have been experimentally verified in some contexts by previous authors. In addition to individual studies, such as those seen in Section 2.3.1, there have been larger-scale tests of nudge concepts. For example, Arno and Thomas conducted a metareview of 42 cases of nudge-based intervention to positively influ-

ence healthful dietary behaviour in adults [48]. They found that nudges had positive effectiveness, resulting in an average 15.3% increase in incidence of healthier dietary choices.

However, while there has been some limited investigation into the application of nudges to HCI, 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.

# 3. Designing with nudges

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 motivate 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 singleplayer 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, and a recumbent exercise bicycle (Fig. 1). 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).

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 [49]. Even halting exercise entirely causes heart rate to drop an average of only 17 bpm per minute of rest [50]. 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.

# 3.1. Planegame

*PlaneGame* (Fig. 2) 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 the 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  $\pm 6$  rpm on either side of it, called the target range. Initial testing found a variation of 6 rpm 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.



Fig. 1. A player playing Gekku Race using the controller and exercise bicycle.



**Fig. 2.** A screenshot of *PlaneGame*. The player's plane avatar and an Al bird are both racing to capture the rainbow for points.

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 (Fig. 3 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 evident: more smoke and louder engine knock. Additionally, fire is added to the smoke coming from the engine, and the edges of the screen start turning red (Fig. 3 right).

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, a formula whose constants were derived through iterative refinement in pilot testing.

#### 3.2. Gekku Race

*Gekku Race* is a racing game within the *Liberi* suite of exergames [51,52], 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 (Fig. 4).

# 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 (Fig. 5 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 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 (Fig. 5 right). This collapse is complemented by a new,



Fig. 3. Nudges in PlaneGame. Left: Early feedback, billowing smoke and motion blur. Right: Advanced feedback, fire added to smoke and screen reddening.



**Fig. 4.** In *Gekku Race*, two players shoot cashews at each other while racing to the top of the track.

stronger sound of gasping for breath. We anticipated that players could be frustrated by this mechanic, as it makes it difficult to go at a flat-out pace, and so we needed it not too occur too frequently or to begin too soon.

This question of frustration is addressed by our study (Section 7).

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* (Section 6), 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* (Section 7), 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.

# 4. Design guidelines

As explored above in Section 2.3.3, existing games have used nudges to control player behaviour. While these games provide compelling examples, there is still a lack of guidelines supporting systematic design of nudge-based interfaces. In this section, we present design guidelines aiding in the creation of nudge-based feedback in video games (Table 1).

In creating the games described in the last section, we performed extensive iterations of design, user testing and redesign. This process provided experience in what design techniques led to nudges that players found easy to understand and compelling to follow. While based on our experience, the guidelines are ultimately grounded in the framework for nudges presented by Hansen and Jespersen, as discussed in Section 2.3.3 [25]. This framework places nudges on two axes, the transparency axis and the type 1 (adjustment of automatic behaviours) versus type 2 (manipulation of active choices) axis. Transparent nudges and type 1 nudges are generally preferable because they are less manipulative. We describe below how the guidelines reflect the lessons arising from this framework.

In terms of operant conditioning [47] The nudges we designed can also be characterized through both reinforcement and punishment. The nudge design we employed demonstrates both positive punishment, in that there are negative consequences when players start going too fast, and negative reinforcement, because these negative consequences are removed when they slow down.

The effectiveness of the nudge systems for the two games were tested by the two studies described later in this paper (Sections 6 and 7).

# 4.1. Natural integration

<u>Nudges should fit the game world</u>: 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



Fig. 5. Nudges and shove in Gekku Race. Left: Gekkus panting for breath, and the screen turning grey. Right: A gekku collapses and gasps for breath.

Table 1Summary of design guidelines.

Design guidelines						
Natural integration	Comprehension	Escalation	Multiple channels			
Nudges should fit the game world	Nudges should be clear and conspicuous	Nudges should escalate from low to high severity	Nudges should employ multiple feedback channels			

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.

Viewed through Hansen and Jespersen's nudge framework, natural integration is a technique intended to make nudges part of the gameplay, allowing players to respond to the nudges on the same mental level as they do other gameplay elements. A naturally integrated nudge is therefore more type 1 than type 2 because it does not require or compel the player to consciously react to the nudges separately from the game.

#### 4.2. Comprehension

<u>Nudges should be clear and conspicuous</u>: 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, where acute health threats cause the screen to begin losing colour (see Section 2.3.2). 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 ava-

tar'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, it 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.

Comprehension maps to the axis of transparency in Hansen and Jespersen's nudge framework. By making the nudge clearly comprehensible to players, it becomes more transparent.

#### 4.3. Escalation

<u>Nudges should escalate from low to high severity</u>: 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 — in the form of shoves — 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.

In terms of Hansen and Jespersen's nudge framework, escalation represents a gradual transition from type 1 to type 2 nudging, as the feedback becomes more insistent after failing to get an automatic response from the player. This transition provides a theoretical foundation for why it is important to use shoves sparingly, and only when it becomes clear they are needed.

#### 4.4. Multiple channels

<u>Nudges should employ multiple feedback channels</u>: 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 guidelines. 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 escalation in severity. Because of this synergy with the other guidelines, it can be helpful to employ the guideline of multiple channels early in the design process so that the available feedback mechanisms can be arranged to best meet all other goals.

The guideline of multiple channels addresses both axes of Hansen and Jespersen's nudge framework. Providing information over multiple channels helps to increase the transparency of the nudge. Increasing the insistence of the nudge through increasing the number of channels effectively migrates the nudge along the type 1/type 2 axis.

#### 5. Research questions

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-ofconcept 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.

## 5.1. Measures

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. In the first study with *PlaneGame*, we had players complete the involvement/control subscale of Witmer and Singer's Presence Questionnaire [21]. For *Gekku Race*, we used the *Immersive Experience Questionnaire* published by Jennett et al. [22]. 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.

#### 6. Proof-of-concept study: PlaneGame

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.

A second research goal was to ensure 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 players could be persuaded to pedal more slowly than they wished by these techniques, as would likely be the case during a warm-up.

# 6.1. Study design

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 (Fig. 6). 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 speeds chosen to represent warm-up and main-game speeds were 40 rpm and 60 rpm, respectively. Pilot testing showed 40 rpm 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; 60 rpm, meanwhile, was a more comfortable speed that



Fig. 6. Textual condition of *PlaneGame*. Target cadence at top of screen, current cadence under plane.

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.

#### 6.1.1. 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. For example, some testers who were complete novices with video games were unable to play effectively even when no feedback was present, making it impossible to note any changes in their behaviour when feedback began. Highly experienced cyclists, meanwhile, were able to find and maintain the optimal cadence within a few seconds, and so we could not examine their responses to feedback because they did not encounter it. Accordingly, the 24 participants were screened to include only those who had at least 50 h 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 [53] to ensure they had no medical conditions that counter-indicated the use of an exercise bike.

# 6.1.2. 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) [21], 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.

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 by the researcher overseeing the trials. Codes were derived by grouping answers that were similar to each other. For example, participant responses indicating that the nudge condition was "more visually appealing" and "more interactive and interesting" were assigned the code "more fun or interesting".

# 6.1.3. 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.

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 before the game began, allowing them to note both the condition and target speed, and to adapt accordingly.

#### 6.2. Results

We first summarize the key findings of this study. More detail follows in Sections 6.2.1 and 6.2.2. The results are discussed in Section 6.3.

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.

#### 6.2.1. 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 40 rpm, a significant effect was found; Wilks' Lambda = 0.310, F(2,22) = 24.16, p < 0.001. 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 < 0.001. The control condition also had a greater average cadence than the textual condition (M = 43.24, SD = 3.22); t(23) = 5.64, p < 0.001. No significant difference was found between the nudge condition and the textual condition at the  $\alpha = 0.05/3$  level; t(23) = 2.11, p = 0.046 (Fig. 7 left).



**Fig. 7.** Performance metrics for *PlaneGame*. Vertical bars show standard error. Horizontal hats indicate statistical significance at  $\alpha$  = 0.05. Left: Average cadence (rpm). Right: Total time over target range (seconds).

At a target cadence of 60 rpm, there was again a significant difference in average cadence between game conditions; Wilks' Lambda = 0.522, F(2,22) = 10.06, p = 0.001. 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 = 0.001. Average cadence was also higher in the control condition than the textual condition (M = 60.33, SD = 1.45); t(23) = 3.40, p = 0.002. Again, no significant difference was found between nudge and textual at the Bonferroni-corrected  $\alpha$  = 0.05/3 level; t(23) = 2.09, p = 0.048 (Fig. 7 left).

At 40 rpm, an RM-ANOVA on time over target range showed significance; Wilks' Lambda = 0.214, F(2,22) = 40.48, p < 0.001. 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 < 0.001. 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 < 0.001. No difference was found between the nudge and textual conditions; t(23) = 1.60, p = 0.123 (Fig. 7 right).

Time over target range also showed significance for 60 rpm; Wilks' Lambda = 0.214, F(2,22) = 40.48, p < 0.001. 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 < 0.001. Control was also higher than textual (M = 2.29, SD = 3.30); t(23) = 5.30, p < 0.001. There was no significant difference between nudge and textual; t(23) = 1.49, p = 0.150(Fig. 7 right).

Game score showed significant differences across conditions at a target cadence of 40 rpm; Wilks' Lambda = 0.684, F(2,22) = 5.08, p = 0.015. 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 = 0.006. Scores were also lower in nudge than in textual; t(23) = 2.82, p = 0.010 (Table 2).

The RM-ANOVA for game score at a target cadence of 60 rpm found no significance at the  $\alpha$  = 0.05 level; Wilks' Lambda = 0.797,

Table 2							
Pairwise	t-tests	for	PlaneGame	game	score a	t 40 rpm.	

Condition (Mean ± SD)	Comparison	t (23)	р
Control (24.13 ± 1.15)	Nudge	3.00	0.006
Nudge (22.79 ± 2.13)	Textual	2.82	0.010
Textual (24.13 ± 1.42)	Control	0.00	1.000

F (2,22) = 2.80, p = 0.082. Given that the ANOVA was not significant, we did not analyze the pairwise results.

# 6.2.2. Questionnaires

A non-parametric Friedman test of the stand-alone motivation question showed statistical significance at 40 rpm;  $\chi^2 = 19.46$ , p < 0.001. 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 = 0.001. They were also less motivated in the control condition than in the textual condition (Mdn = 6.5, IQR = 6–7); Z = -2.95, p = 0.003. There was no significant difference between the nudge and textual conditions; Z = -1.36, p = 0.175 (Fig. 8).

For a target cadence of 60 rpm, the motivation question again showed statistical significance;  $\chi^2 = 21.00$ , p < 0.001. Players were less motivated in the control condition (Mdn = 4.5, IQR = 3–6) than



**Fig. 8.** 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  $\alpha$  = 0.05.

in the nudge condition (Mdn = 7, IQR = 6–7); Z = -3.69, p < 0.001. Players also felt less motivated to slow down in control than in textual (Mdn = 6, IQR = 5–7); Z = -2.79, p = 0.005. No significant difference was found between nudge and textual; Z = -1.83, p = 0.067 (Fig. 8).

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 = 0.097. However, there was a significant difference at 60 rpm;  $\chi^2 = 7.57$ , p = 0.022. 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 = 0.594. There was no significant difference between control and textual (Mdn = 68, r = 51–73); Z = -1.20, p = 0.237. No significance was found between the nudge and textual condition; Z = -1.31, p = 0.196.

# 6.3. 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 40 rpm.

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.

# 7. High-energy study: Gekku Race

The Performance Metricstechniques 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 [51,52]. 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 [4].

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.

# 7.1. 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 75 rpm, 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.

# 7.1.1. Participants

20 participants were recruited using the same criteria as in the first study: university students 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 [53] to ensure they were able to exercise using a stationary bicycle.

# 7.1.2. 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 75 rpm.

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 first statement was "The game made it clear when I was pedaling too fast", to learn if participants found the nudges comprehensible. The second statement, "It was frustrating when I was trying to win the race but the game was telling me to slow down", was designed to test the possibility that participants found the constraints imposed by the shove condition to be frustrating. The other questionnaire was the Immersive Experience Questionnaire (IEQ) [22].

We used the IEQ for this study, rather than the Involvement/ Control subscale of Witmer and Singer's Presence Questionnaire [21] 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.

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 grouped into simplified categories.

# 7.1.3. 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 = 0.723, F(2,8) = 2.55, p = 0.119) or for proportional time above target cadence (Wilks' Lambda = 0.902, F(2,8) = 0.734, p = 0.458). According to a Friedman test, there was also no significant difference in immersion scores:  $\chi^2$  = 1.63, p = 0.458.

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.

#### 7.1.4. 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.

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.

# 7.2. Results

We first summarize the key results of this study. Detailed results can be found in Sections 7.2.1 and 7.2.2, and their implications are explored in Section 7.3.

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.

#### Table 3

Interview questions for Gekku Race.

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?

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.

## 7.2.1. 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 = 0.435, F(2,18) = 18.30, p < 0.001. 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) = 0.256, p = 0.801. Average cadence in the shove condition (M = 69.09, SD = 5.81) was significantly lower than in the control condition, t(19) = 4.93, p < 0.001; and also lower than in nudge, t(19) = 4.42, p < 0.001 (Fig. 9 left).

An RM-ANOVA comparing proportion of time players spend pedaling in excess of the target cadence showed significance; Wilks' Lambda = 0.311, F(2,18) = 25.14, p < 0.001. Paired t-tests showed no difference between the control condition (M = 0.724, SD = 0.356) and the nudge condition (M = 0.643, SD = 0.401); t(19) = 1.17, p = 0.255. Proportion of time over target was lower in the shove condition (M = 0.215, SD = 0.240) than in the control condition, t(19) = 6.24, p < 0.001; and in the nudge condition, t(19) = 5.35, p < 0.001 (Fig. 9 right).

# 7.2.2. 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 = 0.031. 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 = 0.327. Likewise, there was no significant difference between the shove condition (Mdn = 122, IQR = 103–135) and either the control condition, Z = -1.71, p = 0.089; or the nudge condition, Z = -0.48, p = 0.644.

In response to whether the game made it clear when to slow down, the Friedman test showed significance;  $\chi^2 = 20.26$ , p < 0.001. 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 = 0.002. The shove condition (Mdn = 5, IQR = 4–5) was considered clearer than both the control condition, Z = –3.77, p < 0.001; and the nudge condition, Z = –2.37, p = 0.008 (Fig. 10 left).

To the question of whether being told to slow down was frustrating, responses were significantly different;  $\chi^2 = 16.36$ , p < 0.001. 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 = 0.048. However, the shove condition (Mdn = 4.5, IQR = 2–5) was seen as more frustrating than either the control condition, Z = -3.43, p < 0.001; or the nudge condition, Z = -2.46, p = 0.006 (Fig. 10 right).

#### 7.3. 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 *Plane-Game*, 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).



**Fig. 9.** Performance metrics for *Gekku Race*. Vertical bars show standard error. Horizontal hats indicate statistical significance at α = 0.05. Left: Average cadence (rpm). Right: Proportion of time above target.



**Fig. 10.** 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  $\alpha = 0.05$ . Left: *"The game made it clear when I was pedaling too fast"*. Right: *"It was frustrating when I was trying to win the race but the game was telling me to slow down"*.

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 annovance 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 overexertion, 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.

# 8. Synthesis

Overall, these two studies show that nudge-based user interfaces are effective at controlling player behaviour in exergames, without any measurable reduction in player immersion. However, 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.

#### 8.1. Immersion and frustration

We saw in the proof-of-concept study with *PlaneGame* (Section 6) 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 (Section 7), 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 guideline of escalation (having feedback become more severe as players get further from the desired behaviour) is extremely 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 60 rpm in *Plane-Game*). 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 even 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.

# 8.2. Implications for design

One common factor for all the tests we performed was that the target cadence was low, ranging from 75 rpm (a moderately slow pace for most cyclists) down to 40 rpm (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 were 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.

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, as mentioned under the guideline of comprehension (Section 4.2): understanding an effect from parallel experience is just as valuable as if it were innately obvious. It would certainly be possible to develop a suite of feedback tools that can be slotted into an appropriate game, making design faster and getting 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 best way to combine the guidelines 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 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 in "Improving recognition and characterization in groupware with rich embodiments" [54].

This paper represents the first attempt to consider specific guidelines for how nudges in digital games may be designed and where they may best be used. For example, the guidelines 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 [42], while suggesting opportunities for further improvement. Seen through the lens of the design guidelines presented in Section 4, we can see that the overlays use two of the design guidelines to good effect. First, natural integration, by choosing thematic appearances for each game's overlays; second, escalation, with the overlays obstructing vision more and more as the negative player input continues. However, the guideline of comprehension is not entirely fulfilled: while the progressive overlavs are conspicuous, it's not clear without being told that they are linked to the player's emotional state. Using the guideline of multiple channels, we can imagine adding audio feedback to suggest dwindling emotional control: perhaps with a racing pulse, or the buzzing of agitated insects.

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 guidelines of nudge design to any number of occasions where games interact with players.

#### 9. Conclusion

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 cyclingbased 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 immersion. The design strategies used to add nudge feedback to the two games were isolated and written into a list of four design guidelines that can be followed to guide the creation of such feedback systems: natural integration, comprehension, escalation, 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, and another that did not.

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. The nudge condition was considered to be a natural fit to the game environment, but there was no measurable difference in immersion between different conditions.

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 guidelines 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.

#### Acknowledgements

Proofreading for this paper was provided by Sharon Jessup Joyce.

This research was funded by the GRAND Network of Centres of Excellence, and by the Natural Science and Engineering Research Council of Canada.

# References

- T. Park, U. Lee, I.S. MacKenzie, M. Moon, I. Hwang, J. Song, human factors of speed-based exergame controllers, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York City, New York, USA, 2014, pp. 1865–1874.
- [2] J. Sinclair, P. Hingston, M. Masek, Exergame development using the dual flow model, in: Proceedings of the Sixth Australasian Conference on Interactive Entertainment, New York City, New York, USA, 2009.
- [3] J.M. Silva, A.E. Saddik, An adaptive game-based exercising framework, in: 2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems (VECIMS), Ottawa, Ontario, Canada; 2011, pp. 1–6.
- [4] M. Ketcheson, Designing for exertion: using heart rate power-ups to improve energy expenditure in exergames M.Sc. thesis, Queen's University, 2016.
- [5] R.H. Thaler, C.R. Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness, Yale University Press, New Haven, Connecticut, USA, 2008.
- [6] L.M. Mealey, Effect of Dance Dance Revolution on Energy Expenditure and Enjoyment in Severely Overweight Children Compared to Alternative Forms of Activity Ph.D. thesis, University of Pittsburgh, 2008.
  [7] J. Kurczak, T.C.N. Graham, C. Joly, R.L. Mandryk, Hearing is believing:
- [7] J. Kurczak, T.C.N. Graham, C. Joly, R.L. Mandryk, Hearing is believing: evaluating ambient audio for location based games, in: Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, New York City, New York, USA, 2011.
- [8] J. Bolton, D. Lirette, M. Lambert, B. Unsworth, PaperDude: a virtual reality cycling exergame, in: CHI '14 Extended Abstracts on Human Factors in Computing Systems, ACM, New York City, New York, USA, 2014, pp. 475–478.
- [9] M. Brehmer, T.C.N. Graham, T. Stach, Activate your GAIM: a toolkit for input in active games, in: Proceedings of the International Academic Conference on the Future of Game Design and Technology, Vancouver, British Columbia, Canada, 2010, pp. 151–158.
- [10] R. Rahl, Physical Activity and Health Guidelines, Human Kinetics, Champaign, Ilinois, USA, 2010.
- [11] American College of Sports Medicine, ACSM's Guidelines for Exercise Testing and Prescription, Lippincott Williams and Wilkins, Philadelphia, Pennsylvania, USA, 2013.
- [12] W. Peng, J.-H. Lin, J. Crouse, Is playing exergames really exercising? A metaanalysis of energy expenditure in active video games, Cyberpsychol., Behav. Soc. Netw. 14 (2011) 681–688.
- [13] E. Biddiss, J. Irwin, Active video games to promote physical activity in children and youth, Arch. Pediatr. Adolesc. Med. 164 (2010) 664–672.
- [14] R.E. Rhodes, D.E.R. Warburton, S.S.D. Bredin, Predicting the effect of interactive video bikes on exercise adherence: an efficacy trial, Psychol., Health Med. 14 (2009) 631–640.
- [15] L.J. Leininger, M.G. Coles, J.N. Gilbert, Comparing enjoyment and perceived exertion between equivalent bouts of physically interactive video gaming and treadmill walking, Health & Fitness J. Canada 3 (2010) 12–18.
- [16] M. Ketcheson, Z. Ye, T.C.N. Graham, Designing for exertion: how heart-rate power-ups increase physical activity in exergames, in: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, New York City, New York, USA, 79–89, 2015.
- [17] American College of Sports Medicine, Complete Guide to Fitness & Health, Human Kinetics, Champaign, Illinois, USA, 2011.
- [18] S. Möhlenkamp, N. Lehmann, F. Breuckmann, M. Bröcker-Preuss, K. Nassenstein, M. Halle, T. Budde, K. Mann, J. Barkhausen, G. Heusch, K.-H. Jöckel, R. Erbel, Running: the risk of coronary events, Eur. Heart J. 29 (2008) 1903–1910.
- [19] R. Agarwal, E. Karahanna, Time flies when you're having fun: cognitive absorption and beliefs about information technology usage, MIS Quarterly 24 (2000) 665–694.
- [20] P. Sweetser, P. Wyeth, GameFlow: a model for evaluating player enjoyment in games, ACM Comput. Entertain. 3 (Jul. 2005) 1–24.
- [21] B.G. Witmer, M.J. Singer, Measuring presence in virtual environments: a presence questionnaire, Presence 7 (1998) 225–240.
- [22] C. Jennett, A. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, A. Wilson, Measuring and defining the experience of immersion in games, Int. J. Hum Comput Stud. 66 (2008) 641–661.
- [23] J. Sinclair, P. Hingston, M. Masek, Considerations for the design of exergames, in: Proceedings of the 5th international conference on computer graphics and interactive techniques in Australia and Southeast Asia, New York City, New York, USA, 2007, pp. 289–295.

- [24] D.M. Hausman, B. Welch, Debate: to nudge or not to nudge, J. Polit. Philos. 18 (2010) 123–136.
- [25] P.G. Hansen, A.M. Jespersen, Nudge and the manipulation of choice: a framework for the responsible use of the nudge approach to behaviour change in public policy, Eur. J. Risk Regul. 4 (2013) 3–28.
- [26] Local Government Association, Changing Behaviours in Public Health: To Nudge or to Shove?, England, London, 2013
- [27] V. Kalnikaitė, Y. Rogers, J. Bird, N. Villar, K. Bachoul, S. Payne, P.M. Todd, J. Schöning, A. Krüger, How to nudge in situ: designing lambent devices to deliver information salience in supermarkets, in: UbiComp 2011: Ubiquitous Computing, Beijing, China, 2011.
- [28] D.L. Costa, M.E. Kahn, Energy conservation 'nudges' and environmentalist ideology: evidence from a randomized residential electricity field experiment, J. Eur. Econ. Assoc. 11 (2006) 680–702.
- [29] B. Wansink, Environmental factors that increase the food intake and consumption volume of unknown consumers, Annu. Rev. Nutr. 24 (2004) 455–479.
- [30] L. Coventry, P. Briggs, D. Jeske, A. van Moorsel, SCENE: a structured means for creating and evaluating behavioural nudges in a cyber security environment, in: Design, User Experience and Usability, Heraklion, Greece, 2014, pp. 229– 239.
- [31] P. Dolan, M. Hallsworth, D. Halpern, D. King, R. Metcalfe, I. Vlaev, Influencing behavior: the MINDSPACE way, J. Econ. Psychol. 33 (2012) 264–277.
- [32] Institute for the Future, Emerging patterns in active-play video games, in: The Power of Play: Innovations in Getting Active, San Francisco, California, USA, 2011.
- [33] Crystal Dynamics, Tomb Raider, Game (PS3, PS4, Xbox 360, Xbox One, Windows, OS X, Linux), Square Enix, Shinjuku, Tokyo, Japan, 2013.
- [34] D. Hall, DayZ, Game (Windows), Bohemia Interactive, Prague, Czech Republic, 2013.
- [35] Naughty Dog, Uncharted: Drake's Fortune, Game (PS3, PS4), Sony Computer Entertainment, San Mateo, California, USA, 2007.
- [36] E.A.D. Nintendo, Mario Kart 8, Game (Wii U), Nintendo, Kyoto, Japan, 2014.
- [37] Niantic, Pokémon Go, Game (iOS, Android, Apple Watch), Niantic, San Francisco, California, USA.
- [38] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining "gamification", in: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, New York City, New York, USA, 2011.
- [39] S.B. Davis, M. Moar, R. Jacobs, M. Watkins, C. Riddoch, K. Cooke, 'Ere be dragons: heart and health, in: Pervasive, Dublin, Ireland, 2006, pp. 157–164.
- [40] T. Stach, T.C.N. Graham, Exploring haptic feedback in exergames, in: Interact, Lisbon, Portugal, 2011, pp. 18–35.

- [41] J. Waterhouse, P. Hudson, B. Edwards, Effects of music tempo upon submaximal cycling performance, Scand. J. Med. Sci. Sports 20 (2010) 662– 669.
- [42] R.L. Mandryk, S. Dielschneider, M.R. Kalyn, C.P. Bertram, M. Gaetz, A. Doucette, B.A. Taylor, A.P. Orr, K. Keiver, Games as neurofeedback training for kinds with FASD, in: Interaction Design and Children, New York City, New York, USA, 2013.
- [43] M. Nagatsu, Social nudges: their mechanisms and justification, Rev. Philos. Psychol. 6 (2015) 481–494.
- [44] F. Herzberg, B. Mausner, B.B. Snyderman, The Motivation to Work, John Wiley & Sons, New York City, New York, USA, 1959.
- [45] E.H. Calvillo-Gámez, P. Cairns, A.L. Cox, Assessing the core elements of the gaming exeperience, in: R. Bernhaupt, Game user experience evaluation, Springer Publishing, New York City, New York, USA, 2015.
- [46] N. Benhassine, F. Devoto, E. Duflo, P. Dupas, V. Pouliquen, Turning a shove into a nudge? a "labeled cash transfer" for education, Am. Econ. J.: Econ. Pol. 7 (2015) 86–125.
- [47] W.B. Baum, Understanding Behaviorism: Behavior, Culture, and Evolution, Wiley-Blackwell, Hoboken, New Jersey, USA, 2004.
- [48] A. Arno, S. Thomas, The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis, BioMed Central Public Health 16 (Jul.) (2016) 11.
- [49] T.B. Stach, Heart Rate Balancing for Multiplayer Exergames Ph.D. thesis, Queen's University, 2012.
- [50] C.R. Cole, E.H. Blackstone, F.J. Pashkow, C.E. Snader, M.S. Lauer, Heart-rate recovery immediately after exercise as a predictor of mortality, N. Engl. J. Med. 341 (Oct.) (1999) 1351–1357.
- [51] H.A. Hernandez, Z. Ye, T.C.N. Graham, D. Fehlings, L. Switzer, Designing actionbased exergames for children with cerebral palsy, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), ACM, New York City, New York, USA, 2013, pp. 1261–1270.
- [52] S. Knights, T.C.N. Graham, L. Switzer, H.A. Hernandez, Z. Ye, B. Findlay, W.Y. Xie, D. Fehlings, An innovative cycling exergame to promote cardiovascular fitness in youth with cerebral palsy, Developmental Neurorehabilitation (2016) 135–140.
- [53] Canadian Society for Exercise Physiology, PAR-Q+: The Physical Activity Readiness Questionnaire for Everyone <http://www.csep.ca/cmfiles/ publications/parq/parqplussept2011version\_all.pdf> (Accessed May 13, 2016).
- [54] T. Stach, C. Gutwin, D. Pinelle, P. Irani, Improving recognition and characterization in groupware with rich embodiments, in: Proceedings of the 2007 Conference on Human Factors in Computing Systems (CHI), San Jose, California, USA, 2007, pp. 11–20.