The Trick is to Stay Behind?: Defining and Exploring the Design Space of Player Balancing Mechanics

David Gonçalves
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
dmgoncalves@fc.ul.pt

Daniel Barros
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
fc55957@alunos.ciencias.ulisboa.pt

Pedro Pais
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
pgpais@fc.ul.pt

João Guerreiro
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
jpguerreiro@fc.ul.pt

Tiago Guerreiro
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
tjguerreiro@fc.ul.pt

André Rodrigues
LASIGE, Faculdade de Ciências, Universidade de Lisboa
Lisboa, Portugal
afrodrigues@fc.ul.pt

ABSTRACT

In multiplayer gaming, skill disparity can lead to frustrating and excluding experiences. Balancing approaches exist to level the playing field (e.g., providing aim assistance to low-performing players), but it is unclear how different design choices affect individual player experience. We first introduce a design space for balancing mechanics encompassing six categories: Determination, Timing, Targeting, Effect, Feedback, and Information. We then present a mixed-methods study, focused on the effect of two subcategories: Targeting Direction and Effect Dependency on skill. In this study, eight pairs of participants played a game prototype and experienced seven balancing mechanics. We collected data from questionnaires and group interviews, revealing implications for future designs, including the importance of 1) merited victory that does not ignore individual achievements, 2) sense of agency when determining the balancing before and during gameplay, and 3) balancing as an intrinsic part of the game that does not disrupt the core gameplay.

CCS CONCEPTS

• **Human-centered computing** → *HCI theory, concepts and models*; **Empirical studies in HCI**; • **Applied computing** → **Computer games**.

KEYWORDS

gaming, multiplayer, balancing, difficulty adjustment, design space, competition, fairness

ACM Reference Format:
David Gonçalves, Daniel Barros, Pedro Pais, João Guerreiro, Tiago Guerreiro, and André Rodrigues. 2024. The Trick is to Stay Behind?: Defining and Exploring the Design Space of Player Balancing Mechanics. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3642441

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642441

1 INTRODUCTION

Ensuring an adequate level of challenge in multiplayer experiences is a research problem not fully addressed. Given the range of skill levels and abilities that can differ between players, the challenges posed by a game (or by co-players) may not conform to everyone's level of performance. In competitive games, players with higher skill level may dominate the game, leaving other players feeling frustrated and discouraged [15, 18]. In collaborative games, stronger players may complete the challenges alone, leaving others feeling left out or unchallenged [18, 63].

Player balancing and difficulty adjustment were previously proposed to level the playing field, catering to variable skill level and ability [1–5, 14, 20, 28, 29, 35, 36, 52]. These often provide advantages or hindrances based on player performance (e.g., in shooter games, assisting players with low performance by providing aim assistance) [5, 20, 61]. Research suggests that balancing can be applied in gameplay without harming the experience [5, 14, 20].

Yet, player perceptions (e.g., fairness) may vary depending on the particularities of the balancing, as well as the context in which it occurs. Past research gives important but limited insight on how player experience and performance are affected by different approaches of balancing (e.g., adjusting difficulty versus directly manipulating the score) [1, 28, 35, 36, 52]. We argue that, to craft more engaging and inclusive experiences that welcome players with disparate skills, we need a comprehensive understanding and exploration of this design space.

In this work, we make a two-fold contribution. First, we propose a design space for player balancing mechanics, capturing aspects that can be manipulated to create different approaches. The proposed design space can serve as a tool to ideate, fine-tune, and report future designs of player balancing, as well as to deconstruct and analyze existing ones. It can be particularly important to better understand how specific implementations shape player experience and other perceptions (e.g., fairness). The design space comprises aspects that characterize balancing mechanics in terms of their 1) Determination (i.e. who and what determines its Existence, Activation, and Configuration), 2) Timing (i.e. when and for how long it happens), 3) TARGETING (i.e. which players it affects), 4) EFFECT (i.e. how intense it is and how much it depends on variable factors), 5) FEEDBACK (i.e. how visible it is), and 6) type of INFORMATION used for system decision.

Second, we report on a mixed-methods study, focused on the effect of two subcategories of our design space: TARGETING DIRECTION (i.e. whether the balancing assists low-performing players or hinders high-performing ones) and EFFECT DEPENDENCY ON SKILL (i.e. how much the effect depends on players' skill). This study is incorporated in the paper as an exercise of the proposed design space, while contributing to the current understanding of how players perceive balancing mechanics (and specific implementations of them). To achieve this, we developed a competitive game prototype with seven in-game balancing mechanics that manipulate our selected subcategories. We then conducted a mixed-methods study with eight pairs of participants—in each session, pairs engaged with the mechanics followed by a questionnaire and a group semi-structured interview. We were focused on understanding perceptions of enjoyment, effectiveness, and fairness toward different balancing approaches.

The results of a thematic analysis (over interview transcripts and observation notes), supported by questionnaire results, reveal implications to inform future designs: in short, 1) DEPENDENCY ON SKILL is favored for preserving a sense of achievement, yet skill-independent mechanics can still be enjoyed, when potentiating a comeback without guaranteeing victory; 2) reducing player agency by directly manipulating control and taking DETERMINATION away from the player can be detrimental to player experience; and 3) the way the balancing effect is fine-tuned (e.g., INTENSITY) and contextualized within the gameplay is essential, as an obtrusive effect can break engagement. We contribute a nuanced view of balancing in multiplayer gaming, extending past work that conceptualizes its design space and operationalizes specific dimensions to identify their associated trade-offs.

2 RELATED WORK

We start by discussing skill level in the context of gaming experiences, how it can affect player experience, and how the issue has been tackled in single-player experiences. We then detail how player balancing has been pictured in multiplayer experiences, focusing on previous works that inspect player perspectives toward balancing. Finally, we mention works that, similarly to ours, formalize dimensions of game design, including previous design spaces.

2.1 Skill and challenge in gaming

Digital games offer challenges that require a variety of skills to be surpassed. While different views exist, it is generally accepted that gaming often requires a combination of cognitive (e.g., problem-solving), physical (e.g., rapid reactions), emotional (e.g., dealing with tension), and social (e.g., coordinating with others) skills [8, 10, 19, 22, 60]. The difficulty level of a game depends on how demanding it is in these various aspects. As with any activity, players may become comfortable with higher levels of difficulty as they practice their skills [38], while some skills are innate, residing in given traits and abilities [28].

Being challenged is one of the main reasons for playing games [19, 41, 60], and some find enjoyment in facing extreme difficulty [50]. Still, established theory [16, 18, 54, 55] indicates that, generally, an optimal experience lies in the balance between difficulty level and players' skill. In a state of flow [18], players feel challenged but capable, experiencing a continuous sense of progress, which keeps them motivated to continue playing [16]. When a game is too easy, players may experience boredom, while too difficult they may get frustrated or anxious—both can lead to disengagement [16, 18, 39]. Similarly, the sense of competence while and after playing a game is shown to be an essential need for player engagement [55].

Many single-player games provide ways to adjust the challenge presented based on the player's performance or preferences [56]. This is typically achieved by offering multiple difficulty levels (e.g., easy/normal/hard mode) [9, 28]. In these cases, players can choose the difficulty that suits their level of skill. Some employ dynamic difficulty adjustment, which continuously evaluates players' performance during gameplay and adjusts the challenge accordingly [9, 28, 56, 64]—artificial intelligence techniques can also be leveraged [40, 51, 53]. Difficulty adjustment in single-player gaming has been explored by research in different contexts, including aim assistance in first-person shooters [61], level generation in platform games [49], and adjustment based on physiological signals [43].

2.2 Balancing in multiplayer experiences

Difficulty adjustment is a more complex challenge in multiplayer contexts as the difficulty comes not only from what is designed into the gameplay but often from co-players' performance. Mismatched skills and expertise in these contexts have a significant impact. Players may end up frustrated for not being able to keep up with other players or, in turn, feeling unchallenged when outperforming [15]. While this mismatch and its consequences are mainly investigated in competitive scenarios [5, 14, 20, 28, 35, 36, 52], they are also relevant in collaborative ones—where, for instance, collaboration can degrade into one player making all the decisions [63].

Past approaches aim at tackling this issue, exploring mechanisms referred to as balancing [14, 28, 35, 36, 52], multiplayer dynamic difficulty adjustment [1–4], skill assistance [20], and performance/skill accommodation [5, 29, 37][1]. Past work [6] shows there is no comprehensive definition of "balancing" in game design. Generally, the term is used to encapsulate the fine-tuning of rules and difficulty to keep a game winnable and replayable, while fair and challenging for all players [6]. In this work, we specifically focus on player balancing, which we define as any decision or change applied to the gameplay that (potentially) are beneficial to low-performing players' performance and/or detrimental to high-performing players' performance, (potentially) bringing them closer in performance (i.e. where performance is the degree of success in achieving a determined in-game goal). Terms like catch-up and comeback mechanics are also used (common in tabletop game design [24]).

Many games use matchmaking systems to group players of equivalent skill levels. However, these require a measure of player skill, which can be difficult to accurately determine (especially for players new to the game) [5] and may not reflect the current state and

[1]While all terms mentioned would apply, we opt to use "player balancing", as the adjustment might not be necessarily dynamic, assistive, or accommodating.

performance level of a player (e.g., a player might have "*off days*") [4]). Also, while matchmaking can be a strategy to prevent unbalanced matches among strangers playing online, it does not so for friends and family who want to play together [4].

Some titles are designed with mechanisms that balance player skill within the gameplay. One of the most popular in-game balancing mechanisms is "*rubber banding*" in racing games [4, 28, 62], where players who are behind are faster or receive better power-ups, giving them a chance to catch up. Such design options prevent any one player from completely dominating the game or falling too far behind. As recognized by previous works [5, 14, 20, 35], asymmetric game design can also be a balancing strategy—for instance, some games allow a second player to join with a simpler role (e.g., playing as Mario's hat in Super Mario Odyssey [48]). Other types of balancing have been researched, such as aim assistance in shooting games [5, 20], speed and steering adjustment in racing games [14], and manipulation of both physical and digital elements in hybrid and exertional games [28, 36, 52].

A question that gets center stage when implementing balancing mechanics is how players feel about it (e.g., if it is fair [56]). Previous studies compare different balancing techniques in terms of effectiveness and in the perceptions and reactions they elicit from players [5, 14, 28, 35]. Many [3, 20, 28, 35, 36] have focused on visibility of the balancing, studying how the experience differs when players are made aware of the adjustment and when they are not. Some of these works suggest that explicit feedback may not have a negative impact and can actually contribute to a sense of fairness as players are made aware and can adapt to balancing [5, 20, 35, 36]. However, for low-performing players knowing about the balancing can reduce self-esteem [28], achievement [52], and motivation [36]. Past work also suggests that balancing that still depends on player input is preferred to direct performance manipulation [52].

To understand comprehensively the impact of design choices when it comes to balancing, it is important to develop a shared understanding of its design space [4]. Two works have proposed frameworks to formalize balancing in digital games [2, 29]. The first one, developed by Baldwin et al. [2], is based on the analysis of commercial titles and identifies seven high-level design components: Determination, Automation, Recipient, Skill Dependency, User Action, Duration, and Visibility. The second [29] specifically focus on action-level assistance and systematizes how the balancing modifies the mapping of player input to output. Supported by the first framework, Baldwin et al. [1] have subsequently explored how specific dimensions can influence player perspectives, through a survey and interviews. Centered on the hypothetical manipulation of each component (with examples), participants highlighted the importance of player control, personal benefit, and awareness. Further work is needed to understand the effect of these and other design dimensions in the experience.

2.3 Structuring digital games

Formalizing aspects of game design may be essential to project and fine-tune the experiences provided by specific designs [34]. Past work provides frameworks, models, and theories that help researchers and practitioners in processes of ideation, design iteration, analysis of empirical evidence, and simply as an efficient lens to present new designs [23, 34, 56, 57]. Past efforts [23, 56, 57] propose definitions and structural elements of games, distinguishing between formal (e.g., rules), experiential and dramatic (e.g., story), and cultural aspects. Specific concepts of game design have also been formalized before, such as asymmetry [33] and, as aforementioned, player balancing [2, 29].

A design space is a type of framework focused on capturing the range of possible approaches when designing a specific type of artifact [44]. It outlines the key parameters and variables that designers can manipulate or consider when creating a design. By testing different combinations of design elements, one can determine which ones are most effective in achieving the desired outcomes and identify trade-offs [13, 30, 46]. Related to game research, design spaces have been proposed to conceptualize dynamic components and interactions [7] and auditory representation forms in virtual environments [30].

Summary

Past research has explored different approaches to balancing, by manipulating certain design aspects, such as visibility [3, 20, 28, 35, 36], timing [14] and dynamism of the effect [5]. Other works have identified key aspects when designing balancing mechanisms [2, 29]. However, it is not fully understood what aspects can be manipulated to achieve different implementations and how players may react to these. To achieve effective solutions to mitigate the impact of disparate skills, we need research toward a complete and nuanced understanding of its design space and associated trade-offs.

3 THE DESIGN SPACE OF PLAYER BALANCING MECHANICS

Past research shows that player experience and perceptions toward balancing greatly vary according to how it is implemented. As a driving motivation, we aimed to add to this knowledge. Yet, we realized that past work formalizing the design space [2] did not capture dimensions that were found to be relevant in past work, some of which we intended to explore. As such, we first advance the understanding of potential design possibilities, by proposing a design space capturing high-level variables that shape player balancing mechanics. We first describe the procedure we followed to derive this contribution, then detail the design space and its (sub)categories, and finally provide examples of its application.

3.1 Procedure

To build the design space, we relied on scrutinizing related research, balancing mechanics in commercial games, and on our own experiences as gamers. Four authors (DG, DB, PP, and AR were involved in this process, remotely meeting weekly or bi-weekly during two months (from January to March 2022), with the time in between used to reflect, find new examples or counterpoints, and question our beliefs. We leveraged the framework by Baldwin et al. [2] as a starting point[2]. We identified relevant concepts to the design space by reviewing related research (e.g., timing aspects previously explored Cechanowicz et al. [14] that were not formalized) and by

[2]We did not build on the framework proposed by Goll Rossau et al. [29], given the work focuses on a subset of player balancing mechanics (i.e. assistance through input manipulation) and captures specific aspects limited to that space.

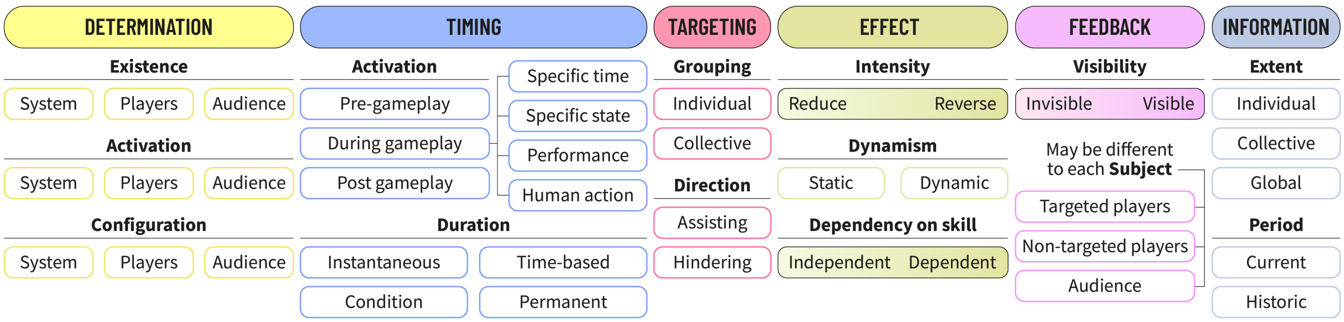

Figure 1: Graphical representation of the design space of player balancing mechanics.

discussing balancing examples we knew of and/or searched and familiarized with before the meetings (e.g., the differences between the hindering effects of a green and blue shell in Mario Kart). This allowed us to identify gaps in the existing framework [2], which triggered discussions on how to integrate the identified concepts.

In our meetings, we gradually developed the structure of the design space, by creating a sheets document that contained the name of the categories, subcategories, their definition, and possible values for these, along with examples. The final structure was agreed upon after multiple iterations, with each version improving in the clarity of definition and distinction between concepts, as well as the relationship/hierarchy between them. We continued to review related material, applying our categories to existing balancing mechanics, re-organizing and refining when needed.

All the authors involved in the process regularly play digital and tabletop games (daily or weekly). They have played digital games since childhood and are familiar with a variety of games, including competitive and collaborative, single and multiplayer. All authors have designed/developed games, mostly in the context of their research, with some public releases in store fronts. DG, DB, and PP have between one to four years of experience in research, while AR has over ten years.

Below we present the consolidated version of the design space after the insights from the study presented in section 4. We detail all that was added and changed from the pre-study design space for each category.

3.2 Description of categories and values

The design space comprises six main categories: 1) **Determination**, 2) **Timing**, 3) **Targeting**, 4) **Effect**, 5) **Feedback**, and 6) type of **Information** used for system decision. Next, we present the definition of each category, subcategory and value to be applied, along with representative examples. The design space is available online[3], with further details and examples. A graphical summary of the design space is displayed in Figure 1.

3.2.1 Determination. Refers to who/what makes the decisions that dictate if and how the mechanic functions.

This category encompasses decision-making in terms of Existence (i.e. whether the mechanic is enabled or available to use), Activation (i.e. when the mechanic takes effect), and Configuration (i.e. the parameters that shape the effect of the mechanic e.g., duration). Each of these aspects can be determined by the *system*, by the *players*, and/or by other partakers (namely, the *audience*). It is possible that a mechanic is determined by a mix of agents and, as such, multiple values can apply to one single mechanic.

Example: In *Ultimate Chicken Horse* [25], the score balancer allows the adjustment of points gained by each player (e.g., extra 10% of points when scoring a successful run). Existence and Configuration is *determined by the players* who adjust the balancing intensity (percentage value). Activation is *determined by the system*, as players do not decide when it takes effect.

Process: The Automation component from Baldwin et al. [2] framework captured whether the application of the balancing was decided by the system (automated) or by the players (manual). This component was extended in order to capture decision-making at different levels (e.g., Configuration), and agents external to the gameplay (audience). The User Action component (which dictates whether the player determines when the mechanic activates) [2] was integrated in this category as Determination of Activation. The subcategories were influenced by past work [36] that discusses determination at different levels (in enabling and in configuring the mechanic). Determination by the audience is influenced by research exploring audience participation in gaming [42, 45, 58].

3.2.2 Timing. Refers to the circumstances that trigger the effect of the mechanic, its duration, and termination. It encompasses Timing of Activation and Duration.

Activation can happen *pre-gameplay* (or at the start), *during gameplay*, or *post gameplay*. Additionally, if it takes effect during gameplay, it may trigger at a *specific time* (e.g., 20 seconds after the match begins), at a *specific state* (e.g., checkpoint), based on players' *performance* (e.g., when one player has ten more points than the other), or based on *human action* (e.g., when the losing player presses a specific key). A mechanic can have multiple activations during gameplay. Additionally, the Activation of a balancing mechanic can happen at multiple stages (see example below). Note that Activation may be determined by players and still be independent of human action (e.g., before gameplay, players determine that Activation should happen at a specific time).

The Duration of the mechanic can be *instantaneous* (i.e. completed within a single moment), *time-based* (i.e. taking effect for a determined period of time), based on a *condition* (e.g., until the

[3]Detailed description of the design space. http://tinyurl.com/bdd7c4r7

losing player reaches ten points), or *permanent* (e.g., until the competition is resolved).

Example: In *Mario Kart* [47], players lagging behind can receive a recovery item named "Bullet Bill". This item significantly increases the speed and automatically steers the car. In terms of TIMING, ACTIVATION happens *during gameplay* at two stages: the moment the players receive the item (*specific state*) and the moment they use it by pressing a specific input (*human action*). The DURATION of its effect is *time-based*, lasting a few seconds.

Process: In Baldwin et al. [2] framework, the Determination component indicates whether the mechanic is activated pre-gameplay or during gameplay. This was extended to specifically capture what triggers the activation during gameplay (e.g., a specific state) and post gameplay activation. The Duration component [2], which captures instantaneous and time-based duration, was extended to capture additional timings (condition and permanent). The added timings were influenced by past work [14] that focus on manipulating when and for how long the mechanic is activated.

3.2.3 TARGETING.

Refers to the subject(s) whose gameplay is affected by the balancing mechanic. It captures the GROUPING and DIRECTION of the balancing.

The GROUPING can be *individual*, if it targets one single player or *collective*, if it targets a group of players (e.g., members of the same team in a team-based game).

DIRECTION captures which side (low- or high-performing) is targeted (i.e. if their gameplay is somehow affected). The balancing mechanic can *assist* low-performing players and/or *hinder* high-performing players. In games where players interact directly (in particular, competitive games), it can happen that the mechanic simultaneously affects low-performing and high-performing players (see example below). When the mechanic simultaneously assists one single player and hinders another single player, GROUPING should be considered collective.

Example: In Paladins [26], the team currently losing the match captures the objective at a faster rate than the team winning. The TARGETING GROUPING is collective, as it affects multiple players. In terms of DIRECTION, it certainly assists low-performing players, but it simultaneously hinders high-performing players as it becomes harder for the winning team to defend and capture objectives.

Process: The Recipient component from Baldwin et al. [2] framework indicates if the mechanic affects an individual player or a team (this value was renamed to collective, to cover situations where multiple players are affected beyond team-based games e.g., collaborative games). We added a new subcategory to capture the Direction of the balancing, as the distinction between hindering and assisting mechanics was identified and discussed before [5, 14, 52].

3.2.4 EFFECT.

Refers to the changes[4] made to the gameplay and the objective impact they have on players' performance. It captures INTENSITY, DYNAMISM, and DEPENDENCY ON SKILL.

INTENSITY indicates the objective impact that the changes made have on players' performance. INTENSITY is a spectrum that goes from an intended (but may not always result in[5]) *reducing* of disparity to an intended *reversing* of disparity.

DYNAMISM indicates if the Effect changes based on variable elements of gameplay, which usually consist of the real-time disparity between players (e.g., the more a player lags behind in performance, the more they is assisted). The balancing can be *dynamic* or *static*.

DEPENDENCY ON SKILL indicates how much the Effect is dependent on the skill of the player being targeted by the balancing mechanic. This is a spectrum, ranging from *independent* to *dependent*. To be independent, the balancing must have a fixed direct increase in the performance of the player(s), without being dependent on how they play (e.g., increasing their score).

Example: In *League of Legends* [27], players who kill other champions get an increasing bounty (*dynamic* balancing with each kill increasing it), while players that die see their bounty decrease (i.e. making them a less appealing target). This mechanic is designed to give an opportunity to the losing team to focus on the high-performing opponents to collect the bounty and discourage focusing on low-performing players to *reduce the disparity*. This extra gold can help these players catch up, but its INTENSITY still largely depends on if they can kill the player and on how they spend the bounty reward (high DEPENDENCY ON SKILL).

Process: The Skill Dependency component from Baldwin et al. [2] framework was directly translated to this design space. Intensity and dynamism were considered before but only added to the design space after the user study reported in section 4, due to many discussions being centered around them. Past works have mentioned balancing "*aggressiveness*" [35] and "*intensity*" [28], which helped us define EFFECT INTENSITY. Similarly, the EFFECT DYNAMISM was influenced by past work that compares between "static" and "dynamic" techniques [5, 14]. These subcategories were not formalized in Baldwin et al. [2] framework.

3.2.5 FEEDBACK.

Refers to the way the mechanic is conveyed in the gameplay (visuals, audio, and other modalities).

VISIBILITY indicates how much feedback is given to the players, and consists of a spectrum ranging from *visible* to *invisible*. Arguably, no mechanic is totally invisible, since the very effect of the mechanic, however subtle it may be, somehow alters the gameplay. A mechanic may be more visible to some players than others (more experienced players will better understand changes in gameplay). By design, FEEDBACK can be different for different partakers. As such, this category also captures VISIBILITY specific to a SUBJECT, including *targeted players* (i.e. players whose gameplay is affected), *non-targeted players*, and other partakers (namely, the *audience*).

Example: There is some debate around player balancing in sports games such as *FIFA* [21] (e.g., higher-performing players failing a bigger percentage of shots) [17]. If balancing exists, it is almost totally *invisible* as it remains unclear among the community if it is actually implemented. While in this case the game provides the same FEEDBACK to everyone, we could imagine a feature that would give additional visual cues to a livestream audience, making the balancing *visible* to a specific SUBJECT.

Process: The Visibility component from Baldwin et al. [2] framework indicates whether and to which player(s) feedback is given

[4]The type of changes applied by a balancing mechanic can vary immensely, depending on the game environment and aspects that can be manipulated within it. As such, we do not propose a typology of specific effects (e.g., speed boost, aim assistance) but rather identify the higher-level categories that characterize them. Recent work proposes a typology of balancing effects, focused on input manipulation [29].

[5]Intensity may not be fixed or constant for one single mechanic as its Effect may be dynamic or dependent on skill.

regarding the presence of balancing. This component was extended to capture visibility to audience and the range of visibility based on the feedback given by design. We did not include the Awareness component of Baldwin et al. [1] framework's updated version, as this does not consist of a dimension controlled by design but rather a product of the experience. Visibility, noticeability, and awareness were aspects explored before in this context [3, 20, 28, 35, 36].

3.2.6 INFORMATION. Refers to the type of data that is used to dictate if and how the mechanic should be applied, when there is some DETERMINATION from the system (e.g., how the system evaluates which player should be benefited). This data may come from in-game activity, but also activity external to the game (e.g., number of trophies unlocked in console account). The category captures the EXTENT and PERIOD of the information.

The EXTENT can be *individual*, when pertaining to one single player, *collective*, when pertaining to a group of players, or even *global*, when pertaining to the entire player base.

The PERIOD can be *current*, if pertaining to data gathered from the current gaming session, or *historic*, if pertaining to data gathered from past sessions.

Example: In *Super Smash Bros. Brawl* [59], players who are lagging behind on points can gain a powerful ability ("Pity Final Smash") after being knocked out. The Information used to configure the mechanic is of *collective* EXTENT (as it compares multiple players' performance captured by the number of points) and pertains to *current* PERIOD (collected during that specific match).

Process: The Determination component from Baldwin et al. [2] framework mentioned the temporality of the data used to determine if the balancing existed (current or past performance). We made this a distinct subcategory and extended it to also capture the EXTENT of information.

3.3 Application

The proposed model encapsulates the design space for balancing mechanics within multiplayer games[6]. Like past design spaces and other types of frameworks [2, 7, 13, 29, 30, 44, 46], it can be leveraged by researchers and designers in multiple ways, from analysis to development and communication. First, it offers a tool for **analysis**, which can help dissect and systematize current approaches, while identifying prevalent and underexplored practices. Second, it can serve as a catalyst for **ideation**—by manipulating the subcategories of the design space, designers can envision novel balancing mechanics, tailoring them to unique player experiences. From a research perspective, it can drive the **design of studies** that explore specific permutations and their impact on player experience. Lastly, the design space provides a platform for effectively presenting new and past designs, fostering **clearer discourse and knowledge exchange** within the research and gaming community.

The complexity of balancing mechanics can vary widely—more complex ones may pose a challenge when applying the design space (e.g., multiple stages of activation). Additionally, the granularity of what qualifies as a balancing mechanic warrants consideration. To give an example: does the rubberbanding in Mario Kart consist of the asymmetric distribution of power-ups, or rather the specific power-ups (e.g., Bullet Bill) given to players behind, that directly affects performance? In some cases, it might even be unclear if a mechanic is considered balancing. We advocate for researchers and practitioners to tailor their definition and level of scrutiny based on a well-defined objective, and apply the design space accordingly.

[6]The design space draws from existing implementations found in research and the gaming industry. While these references consist exclusively of competitive games, we believe the design space is also applicable to collaborative games.

3.4 Limitations

While the presented design space stems from a careful review of related research supported by examples from industry, it does not include a systematic search, collection, and review of publications/games. Its usefulness is also not validated comprehensively (e.g., testing with researchers and game designers). Intentionally, all (sub)categories encapsulate high-level concepts related to player balancing mechanics—further research is needed to formalize more specific aspects of player balancing (similar to how Goll Rossau et al. [29] categorize action-level assistance through input manipulation). We do not claim the design space created to be final, but rather a next step in how we formalize player balancing.

4 UNDERSTANDING THE EFFECTS OF TARGETING DIRECTION AND DEPENDENCY ON SKILL

As a second step of this work, we aimed to operationalize our design space to explore specific permutations and better understand their effects on player perceptions (e.g., fairness). Specifically, we sought to gather nuanced perspectives about a selection of mechanics where we manipulated two subcategories captured by our design space: TARGETING DIRECTION and EFFECT DEPENDENCY ON SKILL. This selection was supported by our expertise in the field and past research [1, 2, 5, 14, 52] that distinguished these subcategories but did not focus on their effect on the player experience (e.g., Rogers et al. [52] discussing "boosts" and "handicaps" as different forms of balancing). We anticipated that varying these subcategories would impact player experience, given they dictate who is affected by the effect and how it relates to player skill.

To achieve this, we developed a competitive racing game with seven in-game balancing mechanics [Figure 2; Table 1] that result from manipulating these subcategories into different combinations (while maintaining all other subcategories fixed). We then conducted a user study with eight pairs, who played the developed game, experienced the seven balancing mechanics, and shared their perceptions through individual questionnaires and a group interview. We relied on the design space to define the study (e.g., select and isolate specific variables) and to align our analysis and discussion (e.g., identify and present how specific variables affected the experience). We had the following research questions:

- **RQ1:** How do players currently view and embrace (or not) player balancing mechanics in multiplayer games?
- **RQ2:** What is the impact of TARGETING DIRECTION and EFFECT DEPENDENCY ON SKILL in players' perceptions of enjoyment, fairness, and effectiveness in a competitive multiplayer experience?

Table 1: Details about the balancing mechanics, including their given name, Targeting Direction (TD) which can be Assisting low-performing players (A) or Hindering high-performing players (H), Effect Dependency on skill (DS) which can be Low (L) or High (H) dependency, and Description.

Name	TD	DS	Description
Boost with invincibility	A	L	For a few seconds, the targeted player gets a speed boost and a shield that allows them to ignore obstacles. Blue particles and the shield appear before activation.
Boost without invincibility	A	H	The targeted player gets a speed boost for a few seconds (the player still has to dodge obstacles). Blue particles appear behind the character before activation.
Obstacle reduction	A	H	For a few seconds, it removes a number of obstacles from the targeted player's track (they disappear into a smoke cloud which stays visible for a couple of seconds).
Obstacle addition	H	H	For a few seconds, it adds a number of obstacles to the targeted player's track (a beacon of light for each added obstacle indicates their position).
Wall with tunnel	H	H	It spawns a wall across the track that blocks visibility and makes it difficult for the player to pass (bound to pass through a narrow entrance).
Timed gate	H	L	It spawns a gate across the track that entirely blocks the passage, until a timer runs out (the player has to stop and wait until the traffic light turns green and the gate opens).
Forced handbrake	H	L	The targeted player gets a sudden decrease of speed they cannot avoid. Red particles appear behind the character when activated.

Below, we provide details on the game prototype and the process behind its design, as well as the entire procedure followed in the user study.

4.1 Game prototype

The game is played on the computer and supports two-player local multiplayer (split screen). It is a racing game, where each player controls a penguin sitting on a floater going down an ice track and tries to reach the finish line first.

Each player is on a separate track and cannot directly interact with each other. The tracks consist of multiple segments connected by either a 90-degree curve to the left, to the right, or a U-shaped one. The tracks from each player are exactly the same but mirrored (left turns for the first player are right turns for the second). The tracks have various elements in between the starting and the finish line, namely 1) curves, 2) obstacles, 3) ice patches, and 4) ramps.

Players have to steer the floater to dodge various obstacles and turn on the curves. Each player starts the game with three hit points and when players hit an obstacle, they lose one hit point, lose some speed, and gain temporary invincibility. When they lose all their hit points, the floater immediately stops and the penguin is launched in the air—having to wait four seconds to reset the floater and return to the track.

Players press assigned keys to steer the floater and slow down (brake). Players do not need to accelerate, as the speed is gained automatically until a maximum speed is reached. When hitting an obstacle, the floater loses speed. There are ice patches scattered on the track—if a player goes over one of them they gain a speed boost that momentarily breaks through the speed maximum.

4.1.1 Balancing mechanics. The seven balancing mechanics are designed based on combinations of values for the two selected subcategories, Targeting Direction (assisting or hindering) and Dependency on skill (low or high dependency). These are described in Table 1. The game, at its core, promotes symmetric competition, but the balancing mechanics are only applicable to one player at all moments (either the low- or high-performing player). We defined fixed values for the non-manipulated subcategories, ensuring they were viable to implement and equivalent across mechanics, in order to specifically focus on differences associated with the subcategories we explored. The fixed values for the non-manipulated subcategories were:

- Determination of Existence, Activation, and Parametrization (all fully determined by the system).
- Timing of Activation (during gameplay) and Duration (time-based).
- Targeting Extent (individual).
- Effect Intensity (intended to even up disparity) and Dynamism (static).
- Feedback (visible to both players with visual/auditory feedback before and during the effect).
- Information (individual Extent and current Period).

Screenshots for every mechanic are presented in Figure 2 and clips are available[7].

4.1.2 Development. To develop our prototype, we extended an existing game, available online and open source[8]. We used existing free-to-use sound effects and models (and built some custom models). Most design decisions and changes to the existing game revolved around the design and fine-tuning of the balancing mechanics. The game prototype is made available online, with options to experience individually each balancing mechanic[9].

4.2 Participants.

We mainly relied on word of mouth to reach our participants. We also disseminated an online call for participation through our research institute. Our sample is constituted by 16 participants (14M, 2F), aged 22-29 (M=23.75, SD=2.65), mostly university students,

[7]Balancing mechanics demonstration video. http://tinyurl.com/32ax7cx7
[8]Sled Racing 3D. https://sethmakesgames.itch.io/sled-racing-3d
[9]Game prototype. http://tinyurl.com/3pfv8ezv

Figure 2: Screenshots demonstrating each balancing mechanic. A – Boost with invincibility; B – Boost without invincibility; C – Obstacle addition; D – Obstacle reduction; E – Wall with tunnel; F – Timed gate; G – Forced handbrake.

with varying levels of gaming expertise [Table 2]. The participants enrolled in the study in pairs, having an established relationship before the study (all friends). We will refer to these participants by a letter to identify the pair and a number to distinguish between players of the same pair (e.g., D2).

4.3 Study sessions.

The sessions were conducted in a room within our university. At least two researchers were present in all sessions. One led the study, while the other/s took notes and observations.

All pairs played a series of races, side-by-side on the same laptop (split-screen). In the first one they raced with no balancing (baseline). Subsequent races allowed participants to experience each balancing mechanic separately: first in a short **practice run** (15 seconds with no curves, where the mechanic would activate for both simultaneously), followed by a **racing match** (approximately 2 minutes, where the mechanic would activate based on performance). We also gave a short description of the mechanic before every practice run. All races had different pre-generated tracks (the same across pairs). The difficulty of each track was carefully built to be equivalent and the exposure to the balancing mechanics was counterbalanced between sessions.

After experiencing all balancing variants, participants were asked to fill a questionnaire individually (not allowed to see or discuss each others' responses). Then, they took part in a group interview. In every session, we used a tablet to accompany the presentation of the study with slides and, after playing the game, to show clips of each mechanic, so players had a visual aid when talking about and comparing between mechanics. Each session lasted between 1h and 1h30 (playing would last for approximately 30m).

4.3.1 Questionnaire. The questionnaire[10] administered to participants had three sections: 1) demographics and gaming habits; 2) miniPXI [31][11], a validated scale to assess player experience, based on eleven items (each to assess a specific construct e.g., Mastery), measured via 7-point Likert items; 3) player preferences, where participants rated each of the seven balancing mechanics in three aspects: i) how much they enjoyed the mechanic, ii) how effective in balancing performance it is, and iii) how fair it is.

4.3.2 Group interview. All interviews were recorded and transcribed. The script to conduct[12] the interview was divided into three sections. In the first section, questions were centered on previous gaming experiences where participants were significantly outperforming or underperforming. The second section revolved around opinions on the concept of player balancing in competitive games and discussion of its use in existing games. The third section delved into their perceptions regarding the balancing existent in our game, each specific mechanic and comparisons between them.

4.4 Analysis

Our analysis is centered on qualitative data collected through the group interviews supported by our quantitative data (miniPXI and mechanic ratings). To analyze qualitative data, we followed the method described as codebook thematic analysis—combining a structured coding procedure and reflective interpretation [11, 12]. DB started by re-familiarizing with the data, generating initial codes and describing them in a shared document. This author reached a first codebook from a subset of interviews and met with DG to discuss his interpretation of the data, refine, merge, and substantiate code definitions. DB proceeded to code all the interviews, adding codes when needed, and, again, met with the team to discuss and agree on a final version of the codebook[13]. Our process includes a mix of inductive and deductive, semantic and latent coding. DG and DB (see positionality statement in section 3.1) worked together to identify common patterns and shared meaning across the data. They worked together on a document with an outline of themes, meeting to discuss their individual views and reach a common understanding. Finally, the whole team met to discuss this document, agreeing on a final outline of themes[14].

We did not conduct null-hypothesis statistical testing over the quantitative results, due to the small sample. These results do not serve as generalizable insights, but rather as an indication of participants' preferences, which we interpret and illustrate with the analysis over the qualitative data.

4.5 Study limitations

While all players experienced all balancing mechanics, the first pair did not complete two matches due to an unexpected bug that was

[10]Questionnaire. http://tinyurl.com/ykcdk8ry
[11]As with any single-item scale, its validation reveals limited reliability to assess specific constructs, but good enough to screen for overall player experience.
[12]Interview script. http://tinyurl.com/3erxdjed
[13]Final codebook. http://tinyurl.com/3w7ujkya
[14]Themes outline. http://tinyurl.com/mtcufde7

Table 2: Details about the participants, including identifier (ID), Age, Gender (Gd), gaming frequency in any platform (monthly/weekly/daily or almost), self-reported gaming expertise (experienced/very experienced), self-reported level of competitiveness when playing digital games (little/somewhat/very/extremely competitive), and pair relationship (friends who regularly/occasionally/do not play together).

ID	Age	Gd	Freq.	Expertise	Compet.	ID	Age	Gd	Freq.	Expertise	Compet.	Relationship
A1	22	M	Weekly	Experienced	Little	**A2**	22	M	Weekly	Very exp.	Somewhat	Friends (regularly)
B1	22	F	Weekly	Experienced	Somewhat	**B2**	22	M	Daily	Very exp.	Extremely	Friends (occasionally)
C1	29	M	Daily	Very exp.	Extremely	**C2**	28	M	Weekly	Very exp.	Very	Friends (do not)
D1	26	F	Monthly	Experienced	Very	**D2**	28	M	Daily	Experienced	Extremely	Friends (do not)
E1	22	M	Daily	Very exp.	Little	**E2**	22	M	Daily	Experienced	Very	Friends (regularly)
F1	22	M	Weekly	Experienced	Somewhat	**F2**	22	M	Weekly	Very exp.	Very	Friends (regularly)
G1	22	M	Daily	Very exp.	Somewhat	**G2**	22	M	Daily	Very exp.	Somewhat	Friends (do not)
H1	23	M	Weekly	Experienced	Very	**H2**	26	M	Daily	Experienced	Little	Friends (do not)

resolved for all other pairs. We included the results since they still gave insightful perspectives regarding both the non-affected and affected matches.

The user guide for applying miniPXI "strongly promote[s]" [32] using a scoring from -3 to +3 with a neutral 0. By mistake, the questionnaire was presented to participants as a 7-point likert scale from 1 to 7 and reported as such. Since these results were only used to assess if the players had an overall positive experience—and not as comparison between mechanics or to derive implications—we believe they are still valuable to include.

Not all players experienced the balancing mechanics as both low-performing and high-performing, as such their perspectives are biased by their side of the experience. However, we believe that by not forcing players to be necessarily exposed to both sides of the balancing, we ensure players have a (although biased) more accurate representation of what would happen in ecological valid settings where skill disparity forces most balancing to take place in one direction.

We implemented seven mechanics, not out of the necessity to add variety or consider other variables at play, but to have mechanics that were directly comparable in terms of gameplay—e.g., handbrake vs boosts, obstacle reduction vs addition. This resulted in an imbalance between variations (only one mechanic for the assisting/low permutation), but allowed us to elicit these comparisons in the interview.

Our sample was solely comprised of young individuals between 22 and 29 years old who regularly play games. Our recruiting efforts were purposefully aligned to gather a somewhat homogeneous sample and reduce the number of affecting variables in the study. Unintentionally, our sample is also skewed toward male players. Our results highlight implications that do not necessarily generalize across populations and contexts and further research is needed to comprehend how these factors affect players' perspectives.

Finally, we note that we did not focus on effectiveness of balancing, but rather considered self-perceived effectiveness as a measure of player preference.

4.6 Findings

We now present the results of our analysis. We start by providing a brief description of the gameplay experiences, centered on the game logs and results from the miniPXI questionnaire. We then present in detail the six themes resulting from our analysis, with representative quotes.

Every mechanic was activated at least once across all racing matches (on average, 6 times per match), with 20 matches (36%) where it was activated only for one player. In most matches, players finished close to each other (except for the aforementioned matches where a bug happened), averaging a difference of 5.2 seconds between them (SD=4.06)—maximum with *wall with tunnel* (M=14.0 seconds, SD=17.76) and minimum with *forced handbrake* (M=2.0 seconds, SD=2.29).

Ratings resulting from the administration of miniPXI averaged M=5.86 (SD=1.35), which indicates a positive overall player experience. The components with the highest mean value were Clarity of Goals (M=6.88, SD=0.34) and Ease of control (M=6.63, SD=0.72), while the one with the lowest mean value was Mastery (M=4.81, SD=1.80).

We calculated mean values for the ratings of enjoyment, effectiveness, and fairness given to each mechanic. Note these values only serve to complement qualitative data. In terms of enjoyment, *obstacle reduction* (M=4.19, SD=1.22) and the *boost with invincibility* (M=4.19, SD=0.98) had the highest mean value, while the *boost without invincibility* (M=2.69, SD=1.45) had the lowest. In terms of effectiveness, the *boost with invincibility* (M=4.25, SD=0.77) had the highest mean value, while the *boost without invincibility* (M=2.63, SD=1.2) had the lowest. In terms of fairness, *obstacle reduction* (M=4.06, SD=1.06) had the highest mean value, while the *forced handbrake* (M=2.56, SD=1.46) had the lowest. A summary of enjoyment, effectiveness, and fairness ratings are presented in Table 3. Quantitative results (miniPXI and ratings) are available in full[15].

4.6.1 Experiences of skill disparity. Participants recounted past experiences of skill disparity when gaming, often under a negative tone. Some (B2, D1, D2, H1, H2) described the **lack of motivation to improve** and to play seriously when unchallenged by their opponents. Others (A2, B1, G1, H2) highlighted the frustration in having **lower-performing players as their teammates**, resulting in anticipated failure and waste of time. As low-performing players, participants also mentioned feeling **unproductive**, **impotent**, and **not like actually playing**: "*I'm just there filling a slot*" (F2).

[15]Questionnaire responses. http://tinyurl.com/3r2v22ss

Table 3: Average (and standard deviation) for ratings of enjoyment, effectiveness, and fairness for each balancing mechanic.

Measure	Boost w/ Inv.	Boost w/o Inv.	Obst. reduction	Obst. addition	Wall w/ tunnel	Timed gate	Forced brake
Enjoyment	4.19 (0.98)	2.69 (1.45)	4.19 (1.22)	3.88 (1.41)	3.56 (1.09)	3.50 (1.63)	2.88 (1.50)
Effectiveness	4.25 (0.77)	2.63 (1.2)	4.13 (1.31)	4.25 (0.77)	3.81 (0.91)	3.94 (1.39)	3.75 (1.06)
Fairness	3.69 (1.25)	3.13 (1.5)	4.06 (1.06)	3.75 (1.13)	3.44 (1.31)	3.00 (1.55)	2.56 (1.46)

Some (A1, A2, B1, H1, H2) explained that negative emotions typically arise due to **toxic attitudes**: "*If [the person] is like, just showing off, it's a bit unpleasant*" (A2). When playing among friends and family, some participants adopt an **unconcerned mindset** and disregard the outcome: "*It is more tolerable [with friends]. I would say there is also a frustration, a bit hidden*" (G2). Others, when outperforming would **adjust their performance** and start to play "*soft*" (E1), and "*toy around*" (E2). Yet, low-performing players may not appreciate this attitude: "*I hate it when he [her friend] lets me win. I hate it, I freak out*" (B2). Situations like these may lead to players **giving up** on playing a game or on playing with friends and family: "*The children I babysit are much better than me, so they exclude me from the game, they [say] it becomes boring*" (D2).

Beyond these, participants (B1, B2, C2, D1, G2, H1, H2) also highlighted positive factors such as the **drive to improve themselves and learn from others** as a lower-performing player: "*It gives me motivation to try to understand how the person is playing, to try to adapt my way of playing and try to win*" (C2). Additionally, participants (C2, E1, F1, H2) mentioned positive experiences when **facilitating the entry and enjoyment** of someone who has less expertise with a game.

4.6.2 Acceptance of balancing dependent on context. Participants (A1, B2, E1) highlighted how matchmaking is the typical form of balancing encountered in commercial titles, which is seen favorably by most. There were split opinions regarding the application of balancing during the gameplay. For some participants (A2, E1, F1, F2, G1), in-game balancing is **acceptable in casual contexts** (where the competitive aspect is reduced), and when the priority is social: "*Racing games of this style, which we play with friends, I think there's not [a problem], nobody takes it too seriously. [...] In a more competitive, like [Counter Strike] or Siege, people would get more upset*" (F2). B1 detailed that she would embrace balancing in a context where the **disparity between players is actually significant**: "*I think it's fair if there is a huge discrepancy between the two players. [...] I have less experience in games than [B2], but I'm not a 10 year old child who doesn't know how to use the keyboard, I don't think it's necessary*".

Some participants (C1 and B2) were not accepting of in-game balancing at all, explaining that it may **devalue competition**. C1 (who only considers matchmaking as a valid balancing mechanism) explains that, by applying balancing, a game might even lose its competitive nature: "*I don't consider Mario Kart competitive, because it changes the rules of the game amid. It rewards people who are behind, and handicaps are given to people who are ahead. [...] To be competitive, the game has to be equal for both*". A1 highlighted the fact that balancing may **devalue practice**, as assisted players might not get motivated to improve their skills: "*The person wouldn't really want to learn to play better*". Additionally, participants (D2, E1, G2) highlighted that, with balancing, the victory would feel less deserved and devalue achievement: "*Eventually I'm going to win just because he's dealing with twenty thousand obstacles*" (G2).

4.6.3 Targeting Direction: The trade-offs of assisting versus hindering. In participants' answers there was no indication of preference for assisting or hindering mechanics. However, some participants (C1, E1, F1, F2, G2) were vocal about their perceived unfairness of hindering mechanics. These participants justified their perspective by highlighting that it was "*counter-intuitive*" (G2) and "*unfair*" (E1) to **penalize a good performance**. Some argued that it affects negatively the experience of high-performing players: "*I think that buffs to someone who is behind is preferable to de-buffs. [...] [De-buffs] are annoying for the player in front*" (F1). Such negative reaction may come from the fact that players do not feel rewarded when performing well, and instead feel punished: "*Those who have skill do not feel that they should be better, because they will end up having some kind of handicap*" (C1).

Other perspectives (B1, B2, C2, G1, G2) leaned to a preference toward hindering mechanics. Participants argued that the balancing should center on upping the difficulty level for high-performing players. Taking the perspective of a low-performing player, B2 highlighted that he would **feel bad if he knew he was playing on "*easy mode*"**. As the assisted player, B1 also expressed their discontent with the changes applied by the balancing to her individual gameplay, as she was trying to focus and improve her performance: "*I preferred not to be touched [laughs]*". From the perspective of high-performing players, hindering can also be seen positively, as it **avoids high-performing players from getting bored**: "*[Hindering] is even more fun, because I felt a little more adrenaline, like he's right there following me*" (G2).

4.6.4 Dependency on skill: Balancing as a tool instead of direct effect. Most participants (except B1 and E2 who were neutral) expressed their preference toward skill dependent mechanics in the interview. This was especially justified for hindering mechanics that, according to some, should provide **additional challenges** but still provide the **opportunity to deal with them** (either mitigate their negative effect or completely avoid it): "*I don't like this type of balancing [talking about the timed gate], because you're blocking the person and they cannot get out of there, regardless of the skill she has*" (C2).

The *forced handbrake* had low ratings in terms of enjoyment (M=2.88, SD=1.50) and fairness (M=2.56, SD=1.46). In the interview, almost all participants (except for C2, E2, F1, and F2) talked negatively about the *forced handbrake*, describing it as "*irritating*" (C1), "*frustrating*" (B1), and "*nonsensical*" (F2). As its effect directly affects performance, participants felt it was more punishing: "*One thing is making my life difficult so he can catch me, another thing is you slowing me down, because he's not doing well*" (F2).

The same was highlighted for the *timed gate*. Participants explained that, in comparison with the *handbrake*, the *gate* did not "*take away control*" (D2) as much from the player and actually added something that players have to deal with: "*It adds the factor that we need to brake, and not just dodge*" (A1). Some also recognized it ends up involving some skill as players can gain some milliseconds by timing it right: "*You have to be as close as possible to the gate and start accelerating as soon as it turns green*" (A2).

Similarly, participants argued that assisting mechanics are fairer when, instead of a direct effect in low-performing players' performance, provide an opportunity for them to catch up, depending on how well they leverage the benefit: "*If the player has to grasp how to use the advantages, I think it ends up being fair, as they have an easier time, but they also have to learn to deal with it*" (A1). Participants gave examples in commercial titles, where exclusive benefits are given to those who are behind that do not directly affect performance (e.g., balancing kill rewards in Counter Strike).

The **sense of control** surfaced as an important factor, especially when participants were explaining why they did not enjoy the *boost without invincibility*, which had low ratings in terms of enjoyment (M=2.69, SD=1.45), and effectiveness (M=2.63, SD=1.20). Given the **risk** associated with its effect (gaining speed made avoiding obstacles more difficult), most participants (except D2, F1, and F2) felt that it did not assist performance: "*It didn't help me much. I guess. Because I got the boost, I hit a tree. [laughs] [...] So I think when I was helped by the game, it often felt like it was a double-edged sword*" (B1). The lack of control in the Activation of this mechanic was determinant, with participants disliking the fact that it was triggered irrespective of their will: "*If I have a tree in front of me, and I get the boost at that moment and I don't have time to dodge, it even gives me a disadvantage*" (A2).

Its skill-independent equivalent (*boost with invincibility*) had high ratings in terms of enjoyment (M=4.19, SD=0.98) and effectiveness (M=4.25, SD=0.77). While its effect is not dependent on player's skill, participants recognized that it succeeded in opening the way for a comeback while not guaranteeing it: "*I got the boost, but that didn't mean I would get ahead of him*" (A2). Some participants (A1, B1, B2, C2, H1, H2) also recognized it still had some skill involved, as it was difficult to get back again on dodging obstacles, after its effect. Still, some (A2, G1) argued that it was not the most fair mechanic, as it temporarily takes the challenge away: "*The shielded one is a cheat*" (G1).

In terms of effectiveness, some participants (D1, D2, E2, and F1) recognized that, given the balancing effect of skill-independent mechanics did not depend on anyone's actions, it did a better job in actually balancing performances: "*The gate is more effective, right, because you're at a standstill and there's no way to escape*" (D2).

4.6.5 Naturalness, prominence, and transparency. Some participants (B1, B2, C1, C2) pointed out how the *obstacle addition* and *removal* felt more **integral to the game**, blending with the overall gameplay: "*It feels like a core mechanic of the game*" (B2). C2 echoed this perception, expressing how obstacle manipulation did not give a sense of "*acquiring an upgrade*" but instead focused on subtly manipulating difficulty: "*We end up just dodging more obstacles*". Conversely, there was some concern with mechanics that felt overly obvious in their attempt to level the playing field (e.g., timed gate), leading to a **sense of interference in the gameplay**: "*It becomes too blatant that you are trying to balance the game, and you are balancing by stopping the other player*" (C2). As many emphasized (A1, A2, C1, C2, D2, G1, G2, H1, H2), Intensity of the effect can be key so balancing does not feel extreme and affect engagement: "*I think all [the mechanics] are too extreme in terms of the balancing. All of them. I think it should be more subtle things*" (G1). This factor was shown to be relevant in participants' acceptance of balancing: "*[When asked if in-game balancing could be envisioned in an esports context] Maybe if they're not as intense or as visible*" (A2).

Amidst these discussions, some participants started discussing to what extent should the balancing effect be noticeable. While this was not the focus of our study and questions, we had **contradicting perspectives about Visibility**. G1 was of the opinion that balancing should be imperceptible, allowing players to fully engage without being explicitly aware of the adjustments taking place: "*I think it should be subtle enough for no one to notice*" (G1). According to this participant, the competitive dynamic that emerges when aware of the balancing may feel strange: "*If you're aware there's a balance, it'll go like 'Damn, I'm ahead and he's already getting the damn boost!', and the other is like: 'Shush, just let me pass'*" (G1). Others (D2 and G2) underscored the importance of players having a clear understanding of how balancing mechanics operate: "*The players should be aware of it. Hiding it from players, I think that's bad*" (D2). Some explained that, by ensuring full transparency, players are able to adapt to the balancing and have a more informed approach to gameplay. G2 illustrated his view with an example occurred during the study: "*I knew that, if I stayed behind, I wouldn't be targeted by the brake. I slowed down a bit and ended up winning in the end. It was part of my strategy*" (G2). Similarly, while playing, A1 joked: "*The trick is to stay behind! [laughs]*".

4.6.6 Fine-tuning and dynamic application of balancing. Participants highlighted that the Effect Intensity of certain mechanics was not adequate at all times. B2, for example, mentioned that hindering mechanics were usually not enough for B1 to "*get ahead*". Some (A1, A2, C1, E1, F1, F2, H2) felt that *obstacle addition* would add too many obstacles: "*It seemed that instead of being hard mode, it was extreme [laughs]*" (H2). The opposite was highlighted (A1, A2, E1) for *obstacle reduction*: "*It took everything out of my way, which also took a bit of the fun*" (D1). Participants suggested that the effect of such mechanics should have been **dynamic instead of static**: "*I think it may also depend on how far apart players are*" (A1).

Dynamism was also envisioned in the integration and application of the different balancing mechanics: "*If I was very distant or something like that, I think it would be cool [to apply the handbrake]. But if the distance was too small, I would prefer a more coherent balance, not so abrupt*" (H2). Participants also emphasized how fine-tuning of Timing aspects can have a significant impact on how players perceive the mechanic. For instance, D1 and G2 mentioned that balancing becomes less acceptable and increases the sense of unfairness when applied at the **last moments of competition**. Similarly, G2 pointed out that Intensity and Timing must be adjusted so the effect alone does not guarantee that the low-performing player overtakes: "*Depending on the other player's position, you can't brake so much if the opponent is close, otherwise that simply means that he's guaranteed to overtake, and I don't think that makes sense*" (G2).

5 DISCUSSION

Perspectives regarding in-game balancing may be contradictory and depend on factors including context (e.g., playing with known others or strangers), the type of game being played (e.g., playing Mario Kart or Counter Strike), and specificities of the balancing approach (e.g., Dependency on skill). In this section, we discuss common trends found in our study and relate them with prior work. We start by discussing balancing under the light of three central topics: 1) merit, 2) player agency, and 3) obtrusiveness. We conclude by enumerating design implications, organized according to our design space.

5.1 Merit: deserving of a comeback

Competitive gaming often pits players against one another in an environment that rewards skill and dedication. As such, the implementation of certain balancing mechanics raised concerns that the core value of competition was not preserved. For some, this perspective did not vary depending on specific implementations, but rather encapsulated their general view toward balancing in gaming. When the outcome is determined by assistance detached from players' performance, it can lead to a disconnect between effort and reward. Consequently, players (both low- and high-performing players) may experience unfulfillment and a lack of motivation to invest time and effort in improving their skills. In our study, this was accentuated by the **preference for skill-dependent mechanics** (expressed by almost all participants in the interview), which increase the possibility of a comeback without guaranteeing victory outright. This is in line with previous work[1, 52], which also calls to the attention the potential negative effect of nullifying the contribution of skill to match outcome.

For assisting mechanics, participants pictured an ideal scenario where a player could **receive an advantage in relation to other players, while still putting their skills to the test**—accordingly, the *obstacle reduction* was considered to be the most enjoyable and fairer mechanic. Participants mentioned that balancing mechanics themselves may need to be learned and efficiently used to have a benefiting effect. Yet, skill-dependent assistance needs to be carefully crafted, given the negative perspectives and low ratings of enjoyment, effectiveness, and fairness given to the *boost without invincibility*. The additional difficulty stemming from having to dodge obstacles at an higher speed made participants feel the mechanic was not assisting but rather hindering their performance. Participants' perspectives suggest that the effect of assisting mechanics, **while preferably dependent on skill, should not heighten the skill required from the gameplay or potentially harm performance**.

In contrast, the *boost with invincibility* was rated as one of the most enjoyed mechanics, even though it is skill-independent. This approach, as recognized by the participants, offered a window of opportunity for the low-performing player, yet placing the onus on their subsequent performance to capitalize on the temporary advantage. This suggests that skill-independent assisting can still be positively received, if the mechanic **empowers players temporarily, without diminishing the value of their achievements**. Fine-tuning aspects of Timing and Effect Intensity (explored in past work [14]) may be essential to achieve this balance.

For hindering mechanics, participants argued that additional challenges should be presented with opportunities for affected players to conquer them and merit their victories. The advice here is to **avoid the feeling that proficient players are being penalized for their skill, but rather rewarded with more demanding challenges**. We can even envision hindering mechanics that present additional challenges, accompanied by additional rewards, such as unlockable content or in-game achievements.

The way players attribute their successes and failures plays a role in their perception of merit. Past work [20, 28] highlights that, when player balancing exists, players might start attributing performance to a source outside their control rather than their own skills. Attribution biases exist, where, for instance, players usually take credit for their successes, but attribute failures externally [20]. Understanding how attribution is shaped by specific design decisions and contexts, taking into account potential biases, should be a concern of future research.

Finally, even though skill-dependent mechanics were preferred to preserve a sense of merit, participants recognized a limitation—the more their effect becomes dependent on players' skill, the less effective they may turn out. There is a **delicate balance between ensuring a balanced environment, while simultaneously fostering a sense of merit associated with victory**.

5.2 Agency: retaining player control

The activation of all mechanics was triggered automatically (based on player's distance). For certain mechanics—in particular, the *boosts* and *forced handbrake*—, this aspect was deemed as a negative factor. Participants expressed dissatisfaction with the way these mechanics **directly manipulated their speed without their consent**. This made participants feel like the system was taking the control away from them, which was seen under a negative light. This was especially aggravated for the *boost without invincibility* as, in many cases, it would activate when participants did not want it (in turn, they wished to have a strategic decision by activating it themselves). Rogers et al. [52] have found that user-triggered balancing (e.g., pushing a button) is preferred to system-triggered, even when its effect is not dependent on skill (e.g., direct score manipulation). Player control has been identified as a relevant aspect for balancing by other works as well [1, 36], but particularly for high-performing players [1]. In our results, we found that **player control was relevant for both low- and high-performing players**, when targeted by the balancing mechanic.

It is possible that other mechanics did not elicit this negative perspective from players, as they altered something in the environment (e.g., adding obstacles) but did not alter player control. Participants' comparisons between the *forced handbrake* and the *timed gate* illustrate this possibility. While these mechanisms hinder performance in an equivalent manner (players are forced to stop or slow down for some seconds), they yielded differing perceptions of agency. *Forced handbrake*, which takes speed away abruptly, results in a diminished sense of agency. Conversely, the *timed gate*, where players themselves have to brake (or choose to not brake and collide with the gate), allowed players to retain their sense of agency despite the imposed constraint. This differentiation underscores the importance of carefully considering the balancing impact on player agency.

As past work highlights [36], affording players the option to choose if and how they want to engage with balancing mechanisms adds a layer of agency in itself. When provided that option, players themselves can evaluate if the context is appropriate (e.g., playing with close friends, co-located). Also, if able to customize the effect (e.g., its DURATION), players can fine-tune the balancing to the intended experience. Yet, Baldwin et al. [1] point out that players might feel socially embarrassed when admitting the need for player balancing, and may prefer the system to decide automatically if it should be enabled or not.

Several participants expressed a positive outlook on some skill disparity experiences, where they found opportunities to learn from high-performing players and teach less experienced friends and family. To enhance this learning process, games could implement **features that facilitate the acquisition of skills and strategies from others**. By empowering players to highlight exemplary moves or strategies, the potential for mutual learning is augmented, and the players can take an active role in shaping the learning curve. Having high-performing players equipped with such tools and possibly enabling assisting features (before or during gameplay) to ease the learning process further reinforces player agency and social engagement.

A negative emotion prominent in participants' past experiences as low-performing players is feeling impotent and useless, merely "*filling a slot*". This feeling and the sensation of being a passive participant may be exacerbated by mechanics that decrease player agency, such as directly manipulating input, simplifying challenges, or bypassing obstacles. To preserve a sense of agency, balancing mechanisms should be carefully designed to **facilitate player-determined balance and skill acquisition**.

5.3 Obtrusiveness: integrated into the gameplay

Prior research [3, 20, 28, 35, 36] has explored VISIBILITY, centered on comparing the experience when players are aware of the balancing and when they are not. A few participants discussed this aspect, with one mentioning he would prefer to be unaware, as balancing can create a strange dynamic between players, where the **balancing disrupts the flow of competition**. Others highlighted that they would not like the game to have balancing mechanics they are not aware of, which is in line with past work that recommends transparency [20, 36].

Regardless of these opposing perspectives, most mechanics were considered to be too **obtrusive** (not integrated in the core gameplay) and/or **excessive** (the effect was exaggerated). The exception were *obstacle addition* and *reduction* which felt "*natural*" to some. While VISIBILITY and awareness were mostly explored before as a binary perception (i.e. players either know or not that balancing is being applied), future research should acknowledge it as a spectrum to better understand their effect on player experience. The emphasis shifts from whether players are aware of the existence of these mechanics to whether they are **sufficiently visible to overtly affect the core gameplay**.

The way the implementation is fine-tuned (e.g., in terms of EFFECT INTENSITY and TIMING) is pivotal in shaping perceptions of balancing obtrusiveness. As highlighted by participants, subtle adjustments that gradually influence the gameplay as the match unfolds can create a more cohesive experience. **Dynamic balancing might help in achieving subtle balancing**—participants suggested that INTENSITY (e.g., speed increase) and the use of different balancing approaches (e.g., using the wall with the tunnel or the handbrake) may vary depending on how close players' performances are. Abrupt interventions when players are in close proximity can be glaring and disrupt the flow of competition, as such dynamism can be an important asset. Aspects other than implementation might also be relevant to ensure non-obtrusive balancing (e.g., thematic integration).

For some, contention surrounding player balancing stems from concerns that mid-game rule alterations disrupt the integrity of the competition. However, if the balancing is transparent and embedded as a natural extension of the gameplay, players may perceive it as an intrinsic part of the game, and adapt their gameplay accordingly. Yet, it is important to recognize that full transparency may lead to **strategic exploitation**, as players may leverage balancing deliberately to their advantage (e.g., staying behind to avoid the opponent getting assistance). While this can enhance strategic depth, it can also mitigate the balancing effect.

These distinctions and the sense of integration expressed by participants might have been overemphasized in this specific context, given that they first tried the game with no balancing and the mechanics were conveyed as modular additions. In an ecological valid environment, players could have become accustomed to the mechanics and perceive them organically embedded within the game. Still, future designs should strive to present balancing as a **dynamic feature that defines the game**, instead of a post hoc adjustment. This reframing encourages players to view the interplay of mechanics and competition as an inseparable whole rather than disparate elements.

5.4 Implications for the design of player balancing mechanics

Our results revealed implications for the design space of player balancing mechanics. We focus on implications found for the subcategories we manipulated, TARGETING DIRECTION and EFFECT DEPENDENCY ON SKILL (RQ2), but also present indications given for the remaining subcategories (which, note, are not empirically supported in our study by direct comparison between mechanics).

5.4.1 TARGETING DIRECTION.

- Assisting mechanics might make low-performing players feel like their experience is being diminished. Mechanics that take away or substitute challenges (e.g., removing obstacles) may be especially aggravating of this perception.
- Hindering mechanics might make players feel like the game is penalizing good performance. Mechanics can be reframed as offering new challenges (and possibly rewards) to engage high-performing players.

5.4.2 EFFECT DEPENDENCY ON SKILL.

- DEPENDENCY ON SKILL is preferred for preserving a sense of achievement and merit.

- The effectiveness of assisting mechanics can (and, for some players and contexts should) depend on player skill, but should not increase the skill required or potentially have a detrimental effect in performance.
- Skill-independent mechanics can still be appropriate, when providing a temporary and subtle adjustment that potentiate a comeback but does not guarantee it.

5.4.3 Other indications.

- Activation triggered by human action may enhance the sense of agency, as players can strategically time when they want the balancing to activate. This might be especially relevant when player control is directly manipulated.
- Activation in the final moments of a match might increase the sense of outcome detached from player performance, and induce sentiments of unfairness. The same can happen if activated when players are close to each other in performance.
- The type and Intensity of the effect must be carefully adjusted to avoid the perception of obtrusiveness within the gameplay—the employment of a dynamic Effect (e.g., proportional to the performance disparity) is recommended.
- Feedback should be balanced, as a prominent and/or persistent visible effect may break engagement and lead to a strange dynamic between players (where they are overly conscious of the effect on performance and match outcome).

5.5 Limitations of scope & Outlook

We note that our observations are based on the manipulation of two design dimensions (with all other dimensions fixed to a chosen value), with subjacent choices regarding game genre, aesthetics, among others—our observations might not generalize to every context and type of game. We focused on a competitive racing game, aligned with past work that has a similar focus [14]. Also, our results relate to two-player (head-to-head) competition, which is not necessarily the type of competition found in racing games. The study itself was focused on the specific context of co-located casual shared play between friends. Covering all possible design permutations and contexts would be impossible in one work and further research is needed to understand the effects of different balancing approaches. In particular, future work should address **balancing in collaborative scenarios and within teams**, which is shown to also be an issue in our study.

Participants were aware the goal of the study was to study balancing mechanics. However, many competitive games have embedded balancing mechanics that players do not necessarily construed as a balancing mechanic but rather as the game is. We believe we can provide insights into how players perceive balancing mechanics once they are aware, but not necessarily how they react to them when faced in an unprompted environment as it would be in first play sessions with a new game. We used a game with low skill variation, which does not correspond to the type of game that would most need balancing. It is important that future research on balancing takes these aspects into account and **strive to conduct studies in ecologically valid settings**. Further, although the design space is used as a tool fundamental to design, conduct, and report our study, we acknowledge its usefulness still needs to be validated.

6 CONCLUSION

In this work, we first present a design space grounded in related research and examples from the industry, with potential applications in analysis, ideation, study design, and clearer discourse. In particular, the design space enables future research to 1) explore specific design variables while controlling covariates and 2) center and/or organize their analyses and reflections on specific subcategories e.g., design implications organized by subcategory.

We then present the results of a mixed-methods study focused on exploring how Targeting Direction and Effect Dependency on skill (two subcategories of our design space) affect self-perceived enjoyment, fairness, and balancing effectiveness. Our observations emphasize that player balancing is not accepted/enjoyed by everyone and/or in every circumstance (RQ1). Participants favored these mechanics in specific contexts (e.g., casual gaming, teaching a game to a newcomer), and highlighted factors that facilitated engagement (RQ1), namely 1) sense of merit, 2) sense of agency, and 3) natural integration in the game.

We discuss how the manipulated subcategories shaped these perceptions (RQ2)—for instance, how 1) some hindering mechanics were perceived as penalizing good performance and considered unfair, 2) mechanics that depend on skill were preferred for preserving a sense of merit and fairness, but considered less effective.

This work constitutes one more step in how we formalize player balancing and contributes with a better understanding on what makes players accept and enjoy these mechanics in games. Further research (e.g., addressing other design variables and other contexts) is needed toward more inclusive multiplayer gaming where skill disparity is considered and embraced by design.

ACKNOWLEDGMENTS

We thank the participants in our user study and the anonymous reviewers for their valuable feedback. This work was supported by FCT through project “Plug n’ Play: Exploring Asymmetry and Modularity for Inclusive Game Design" ref. 2022.08895.PTDC (http://doi.org/10.54499/2022.08895.PTDC), the Institutional CEEC (https://doi.org/10.54499/CEECINST/00032/2018/CP1523/CT0003), scholarships ref. UI/BD/151178/2021 (https://doi.org/10.54499/UI/BD/151178/2021) and ref. 2022.12448.BD, and the LASIGE Research Unit, ref. UIDB/00408/2020 (https://doi.org/10.54499/UIDB/00408/2020) and ref. UIDP/00408/2020 (https://doi.org/10.54499/UIDP/00408/2020).

REFERENCES

[1] Alexander Baldwin, Daniel Johnson, and Peta Wyeth. 2016. Crowd-Pleaser: Player Perspectives of Multiplayer Dynamic Difficulty Adjustment in Video Games. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 326–337. https://doi.org/10.1145/2967934.2968100

[2] Alexander Baldwin, Daniel Johnson, Peta Wyeth, and Penny Sweetser. 2013. A framework of Dynamic Difficulty Adjustment in competitive multiplayer video games. In *2013 IEEE International Games Innovation Conference (IGIC)*. 16–19. https://doi.org/10.1109/IGIC.2013.6659150

[3] Alexander Baldwin, Daniel Johnson, and Peta A. Wyeth. 2014. The Effect of Multiplayer Dynamic Difficulty Adjustment on the Player Experience of Video Games. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 1489–1494. https://doi.org/10.1145/2559206.2581285

[4] Alexander J. Baldwin. 2017. *Balancing act : the effect of dynamic difficulty adjustment in competitive multiplayer video games.* Ph. D. Dissertation. Queensland University of Technology. https://doi.org/10.5204/thesis.eprints.102669

[5] Scott Bateman, Regan L. Mandryk, Tadeusz Stach, and Carl Gutwin. 2011. Target Assistance for Subtly Balancing Competitive Play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2355–2364. https://doi.org/10.1145/1978942.1979287

[6] Alexander Becker and Daniel Görlich. 2020. What is Game Balancing? - An Examination of Concepts. *ParadigmPlus* 1, 1 (Apr. 2020), 22–41. https://doi.org/10.55969/paradigmplus.v1n1a2

[7] Kristopher J. Blom and Steffi Beckhaus. 2014. The design space of dynamic interactive virtual environments. *Virtual Reality* 18, 2 (01 Jun 2014), 101–116. https://doi.org/10.1007/s10055-013-0232-y

[8] Julia Ayumi Bopp, Klaus Opwis, and Elisa D. Mekler. 2018. "An Odd Kind of Pleasure": Differentiating Emotional Challenge in Digital Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173615

[9] Barbaros Bostan and Sertaç Öğüt. 2009. Game challenges and difficulty levels: lessons learned From RPGs. In *International simulation and gaming association conference*. 1–11.

[10] Nicholas Bowman, Joseph Wasserman, and Jaime Banks. 2018. *Development of the Video Game Demand Scale*. 208–233. https://doi.org/10.4324/9781351235266-13

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa

[12] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806 arXiv:https://doi.org/10.1080/2159676X.2019.1628806

[13] Stuart K. Card, Jock D. Mackinlay, and George G. Robertson. 1990. The Design Space of Input Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '90)*. Association for Computing Machinery, New York, NY, USA, 117–124. https://doi.org/10.1145/97243.97263

[14] Jared E. Cechanowicz, Carl Gutwin, Scott Bateman, Regan Mandryk, and Ian Stavness. 2014. Improving Player Balancing in Racing Games. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (Toronto, Ontario, Canada) *(CHI PLAY '14)*. Association for Computing Machinery, New York, NY, USA, 47–56. https://doi.org/10.1145/2658537.2658701

[15] Delwin Clarke and P. Robert Duimering. 2006. How Computer Gamers Experience the Game Situation: A Behavioral Study. 4, 3 (jul 2006), 6–es. https://doi.org/10.1145/1146816.1146827

[16] Benjamin Cowley, Darryl Charles, Michaela Black, and Ray Hickey. 2008. Toward an understanding of flow in video games. *Comput. Entertain.* 6 (07 2008), 1–27. https://doi.org/10.1145/1371216.1371223

[17] Joe Craven. 2023. What is DDA in gaming and why is it so controversial? FIFA 23, EA Patent & more. https://www.dexerto.com/fifa/what-is-dda-in-gaming-fifa-22-ea-1766885/

[18] Mihaly Csikszentmihalyi. 1991. *Flow: The Psychology of Optimal Experience*. Harper Perennial, New York, NY. http://www.amazon.com/gp/product/0060920432/ref=si3_rdr_bb_product/104-4616565-4570345

[19] Alena Denisova, Paul Cairns, Christian Guckelsberger, and David Zendle. 2020. Measuring perceived challenge in digital games: Development & validation of the challenge originating from recent gameplay interaction scale (CORGIS). *International Journal of Human-Computer Studies* 137 (2020), 102383. https://doi.org/10.1016/j.ijhcs.2019.102383

[20] Ansgar E. Depping, Regan L. Mandryk, Chengzhao Li, Carl Gutwin, and Rodrigo Vicencio-Moreira. 2016. How Disclosing Skill Assistance Affects Play Experience in a Multiplayer First-Person Shooter Game. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3462–3472. https://doi.org/10.1145/2858036.2858156

[21] EA. 2022. EA SPORTS™ FIFA 23. Digital game [Microsoft Windows, PlayStation 4, PlayStation 5, Xbox One, Xbox Series X/S, Nintendo Switch]. www.ea.com/games/fifa/fifa-23

[22] Alex Flint, Alena Denisova, and Nick Bowman. 2023. Comparing Measures of Perceived Challenge and Demand in Video Games: Exploring the Conceptual Dimensions of CORGIS and VGDS. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 571, 19 pages. https://doi.org/10.1145/3544548.3581409

[23] Tracy Fullerton. 2008. *Game Design Workshop. A Playcentric Approach to Creating Innovative Games*. https://doi.org/10.1201/b22309

[24] The Thoughtful Gamer. 2017. Catch-up mechanisms - the thoughtful gamer. https://thethoughtfulgamer.com/2017/03/28/catch-up-mechanisms/

[25] Clever Endeavour Games. 2016. Ultimate Chicken Horse. Digital game [Microsoft Windows, Linux, Mac, Playstation 4, Xbox One, Nintendo Switch]. https://www.cleverendeavourgames.com/ultimate-chicken-horse

[26] Evil Mojo Games. 2018. Paladins. Digital game [Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch]. https://www.paladins.com

[27] Riot Games. 2009. League of Legends. Digital game [Microsoft Windows, Mac]. https://www.leagueoflegends.com/

[28] Kathrin Maria Gerling, Matthew Miller, Regan L. Mandryk, Max Valentin Birk, and Jan David Smeddinck. 2014. Effects of Balancing for Physical Abilities on Player Performance, Experience and Self-Esteem in Exergames. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2201–2210. https://doi.org/10.1145/2556288.2556963

[29] Ingeborg Goll Rossau, Rasmus Bugge Skammelsen, Jundefineddrzej Jacek Czapla, Bastian Ilsø Hougaard, Hendrik Knoche, and Mads Jochumsen. 2021. How Can We Help? Towards a Design Framework for Performance-Accommodation Mechanisms for Users Struggling with Input. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Austria) *(CHI PLAY '21)*. Association for Computing Machinery, New York, NY, USA, 10–16. https://doi.org/10.1145/3450337.3483497

[30] João Guerreiro, Yujin Kim, Rodrigo Nogueira, SeungA Chung, André Rodrigues, and Uran Oh. 2023. The Design Space of the Auditory Representation of Objects and Their Behaviours in Virtual Reality for Blind People. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2763–2773. https://doi.org/10.1109/TVCG.2023.3247094

[31] Aqeel Haider, Casper Harteveld, Daniel Johnson, Max Birk, Regan Mandryk, Magy El-Nasr, Lennart Nacke, Kathrin Gerling, and Vero Vanden Abeele. 2022. miniPXI: Development and Validation of an Eleven-Item Measure of the Player Experience Inventory. *Proceedings of the ACM on Human-Computer Interaction* 6 (11 2022), 1–26. https://doi.org/10.1145/3549507

[32] Aqeel Haider, Casper Harteveld, Daniel Johnson, Max Birk, Regan Mandryk, Magy Seif El Nasr, Lennart E. Nacke, Kathrin Gerling, and Vero Vanden Abeele. 2023. https://playerexperienceinventory.org/docs

[33] John Harris, Mark Hancock, and Stacey D. Scott. 2016. Leveraging Asymmetries in Multiplayer Games: Investigating Design Elements of Interdependent Play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 350–361. https://doi.org/10.1145/2967934.2968113

[34] Robin Hunicke, Marc Leblanc, and Robert Zubek. 2004. MDA: A Formal Approach to Game Design and Game Research. *AAAI Workshop - Technical Report* 1 (01 2004).

[35] Susan Hwang, Adrian L. Jessup Schneider, Daniel Clarke, Alexander Macintosh, Lauren Switzer, Darcy Fehlings, and T.C. Nicholas Graham. 2017. How Game Balancing Affects Play: Player Adaptation in an Exergame for Children with Cerebral Palsy. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 699–710. https://doi.org/10.1145/3064663.3064664

[36] Mads Møller Jensen and Kaj Grønbæk. 2016. Design Strategies for Balancing Exertion Games: A Study of Three Approaches. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) *(DIS '16)*. Association for Computing Machinery, New York, NY, USA, 936–946. https://doi.org/10.1145/2901790.2901843

[37] Mads Jochumsen, Bastian Ilsø Hougaard, Mathias Sand Kristensen, and Hendrik Knoche. 2022. Implementing Performance Accommodation Mechanisms in Online BCI for Stroke Rehabilitation: A Study on Perceived Control and Frustration. *Sensors* 22, 23 (2022). https://doi.org/10.3390/s22239051

[38] Jesper Juul. 2013. *The Art of Failure: An Essay on the Pain of Playing Video Games*. The MIT Press.

[39] Madison Klarkowski, Daniel Johnson, Peta Wyeth, Mitchell McEwan, Cody Phillips, and Simon Smith. 2016. Operationalising and Evaluating Sub-Optimal and Optimal Play Experiences through Challenge-Skill Manipulation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5583–5594. https://doi.org/10.1145/2858036.2858563

[40] Raúl Lara-Cabrera, Carlos Cotta, and Antonio J. Fernández-Leiva. 2014. On balance and dynamism in procedural content generation with self-adaptive evolutionary algorithms. *Natural Computing* 13, 2 (01 Jun 2014), 157–168. https://doi.org/10.1007/s11047-014-9418-9

[41] Nicole Lazzaro. 2012. *Why We Play: Affect and the Fun of Games—Designing Emotions for Games, Entertainment Interfaces, and Interactive Products*. Taylor & Francis.

[42] Pascal Lessel, Maximilian Altmeyer, Julian Sahner, and Antonio Krüger. 2022. Streamer's Hell - Investigating Audience Influence in Live-Streams Beyond the Game. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 252 (oct 2022), 27 pages. https://doi.org/10.1145/3549515

[43] Changchun Liu, Pramila Agrawal, Nilanjan Sarkar, and Shuo Chen. 2009. Dynamic Difficulty Adjustment in Computer Games Through Real-Time Anxiety-Based Affective Feedback. *International Journal of Human–Computer Interaction* 25, 6 (2009), 506–529. https://doi.org/10.1080/10447310902963944 arXiv:https://doi.org/10.1080/10447310902963944

[44] Allan Maclean, Richard Young, Victoria Bellotti, and Thomas Moran. 2004. Design Space Analysis: Bridging from Theory to Practice via Design Rationale. *Proceedings of Esprit '91* (10 2004).

[45] Bernhard Maurer, Ilhan Aslan, Martin Wuchse, Katja Neureiter, and Manfred Tscheligi. 2015. Gaze-Based Onlooker Integration: Exploring the In-Between of Active Player and Passive Spectator in Co-Located Gaming. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (<conf-loc>, <city>London</city>, <country>United Kingdom</country>, </conf-loc>) *(CHI PLAY '15)*. Association for Computing Machinery, New York, NY, USA, 163–173. https://doi.org/10.1145/2793107.2793126

[46] Hugo Nicolau, André Rodrigues, André Santos, Tiago Guerreiro, Kyle Montague, and João Guerreiro. 2019. The Design Space of Nonvisual Word Completion. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 249–261. https://doi.org/10.1145/3308561.3353786

[47] Nintendo. 2014. Mario Kart™ 8. Digital game [Wii U, Nintendo Switch]. https://mariokart8.nintendo.com

[48] Nintendo. 2017. Super Mario Odyssey. Digital game [Nintendo Switch]. https://www.nintendo.com/games/detail/super-mario-odyssey-switch/

[49] Christopher Pedersen, Julian Togelius, and Georgios N. Yannakakis. 2010. Modeling player experience for content creation. *IEEE Transactions on Computational Intelligence and AI in Games* 2, 1 (March 2010), 54–67. https://doi.org/10.1109/TCIAIG.2010.2043950 Funding Information: Manuscript received October 02, 2009; manuscript revised December 17, 2009. Date of manuscript acceptance February 17, 2010; date of publication February 25, 2010; date of current version March 17, 2010. This work was supported in part by the Danish Research Agency, Ministry of Science, Technology and Innovation under Project AGameComIn number 274-09-0083..

[50] Serge Petralito, Florian Brühlmann, Glena Iten, Elisa D. Mekler, and Klaus Opwis. 2017. A Good Reason to Die: How Avatar Death and High Challenges Enable Positive Experiences. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5087–5097. https://doi.org/10.1145/3025453.3026047

[51] Johannes Pfau, Jan David Smeddinck, and Rainer Malaka. 2020. Enemy Within: Long-Term Motivation Effects of Deep Player Behavior Models for Dynamic Difficulty Adjustment. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376423

[52] Katja Rogers, Mark Colley, David Lehr, Julian Frommel, Marcel Walch, Lennart E. Nacke, and Michael Weber. 2018. KickAR: Exploring Game Balancing Through Boosts and Handicaps in Augmented Reality Table Football. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173740

[53] Florian Rupp, Manuel Eberhardinger, and Kai Eckert. 2023. Balancing of competitive two-player Game Levels with Reinforcement Learning. (2023). arXiv:2306.04429 [cs.LG]

[54] Richard Ryan and Edward Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *The American psychologist* 55 (02 2000), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

[55] Richard Ryan, C. Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30 (12 2006), 344–360. https://doi.org/10.1007/s11031-006-9051-8

[56] Katie Salen and Eric Zimmerman. 2003. *Rules of Play: Game Design Fundamentals*. The MIT Press.

[57] Jesse Schell. [2008]. *The art of game design : a book of lenses*. Amsterdam ; Boston : Elsevier/Morgan Kaufmann, [2008] ©2008. https://search.library.wisc.edu/catalog/9910060875802121 Includes bibliographical references (pages 477-479) and index..

[58] Joseph Seering, Saiph Savage, Michael Eagle, Joshua Churchin, Rachel Moeller, Jeffrey P. Bigham, and Jessica Hammer. 2017. Audience Participation Games: Blurring the Line Between Player and Spectator. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 429–440. https://doi.org/10.1145/3064663.3064732

[59] Game Arts Sora Ltd. 2008. Super Smash Bros. Brawl. Digital game [Wii]. https://www.smashbros.com/wii/en_us/index.html

[60] Jukka Vahlo and Veli-Matti Karhulahti. 2020. Challenge types in gaming validation of video game challenge inventory (CHA). *International Journal of Human-Computer Studies* 143 (2020), 102473. https://doi.org/10.1016/j.ijhcs.2020.102473

[61] Rodrigo Vicencio-Moreira, Regan L. Mandryk, Carl Gutwin, and Scott Bateman. 2014. The Effectiveness (or Lack Thereof) of Aim-Assist Techniques in First-Person Shooter Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 937–946. https://doi.org/10.1145/2556288.2557308

[62] Anthony Youssef and Stephen Cossell. 2009. Thoughts on Adjusting Perceived Difficulty in Games. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (Sydney, Australia) *(IE '09)*. Association for Computing Machinery, New York, NY, USA, Article 14, 4 pages. https://doi.org/10.1145/1746050.1746064

[63] José P. Zagal, Jochen Rick, and Idris Hsi. 2006. Collaborative games: Lessons learned from board games. *Simulation & Gaming* 37, 1 (2006), 24–40. https://doi.org/10.1177/1046878105282279 arXiv:https://doi.org/10.1177/1046878105282279

[64] Mohammad Zohaib. 2018. Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Advances in Human-Computer Interaction* 2018 (01 Nov 2018), 5681652. https://doi.org/10.1155/2018/5681652