



How does Juicy Game Feedback Motivate? Testing Curiosity, Competence, and Effectance

Dominic Kao
kaod@purdue.edu
Purdue University
USA

Nick Ballou
nick.ballou@oii.ox.ac.uk
University of Oxford
UK

Kathrin Gerling
kathrin.gerling@kit.edu
Karlsruhe Institute of Technology
Germany

Heiko Breitsohl
Heiko.Breitsohl@aau.at
University of Klagenfurt
Austria

Sebastian Deterding
s.deterding@imperial.ac.uk
Imperial College London
UK

ABSTRACT

‘Juicy’ or immediate abundant action feedback is widely held to make video games enjoyable and intrinsically motivating. Yet we do not know why it works: Which motives are mediating it? Which features afford it? In a pre-registered ($n=1,699$) online experiment, we tested three motives mapping prior practitioner discourse—effectance, competence, and curiosity—and connected design features. Using a dedicated action RPG and a 2×2 +control design, we varied feedback amplification, success-dependence, and variability and recorded self-reported effectance, competence, curiosity, and enjoyment as well as free-choice playtime. Structural equation models show curiosity as the strongest enjoyment and only playtime predictor and support theorised competence pathways. Success dependence enhanced all motives, while amplification unexpectedly reduced them, possibly because the tested condition unintentionally impeded players’ sense of agency. Our study evidences uncertain success affording curiosity as an underappreciated moment-to-moment engagement driver, directly supports competence-related theory, and suggests that prior juicy game feel guidance ties to legible action-outcome bindings and graded success as preconditions of positive ‘low-level’ user experience.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in interaction design**; • **Applied computing** → **Computer games**.

KEYWORDS

Video games, juicy feedback, juiciness, competence, effectance, curiosity, enjoyment, engagement

ACM Reference Format:

Dominic Kao, Nick Ballou, Kathrin Gerling, Heiko Breitsohl, and Sebastian Deterding. 2024. How does Juicy Game Feedback Motivate? Testing Curiosity, Competence, and Effectance. In *Proceedings of the CHI Conference*



This work is licensed under a Creative Commons Attribution-Share Alike International 4.0 License.

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

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3642656>

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/10.1145/3613904.3642656>

1 INTRODUCTION

Juiciness or immediate “excessive amounts of feedback in relation to user input” [53, p. 139] is widely present in video games. It affects play time, player experience, and motivation [28, 35] and is considered a key part of good game feel [26, 53]. Gamification researchers have therefore called it out as a motivational affordance that could make interactive systems in general more enjoyable and intrinsically motivating [17]. But despite its ubiquity and recognised importance, views on what counts as “juicy” feedback diverge, and we do not know *when and why* immediate excessive feedback is motivating and enjoyable. Answering this question can both help designers better fine-tune feedback and advance HCI understanding of its interaction aesthetics.

Prior work has generated bottom-up characteristics of what developers consider good or juicy game feel [26, 53] and exploratory empirical tests of various forms of amplified feedback and player experience. To advance the field, we need theory-guided, experimental work to systematically identify, isolate, and test precise design features and psychological mechanisms [3]. To this end, we identified three theories and candidate motives that mapped prior practitioner concepts: effectance—the basic positive experience of causing effects [71]; competence—exercising, expanding, and expressing one’s abilities [57]; and curiosity—interest in generating and reducing uncertainty [36]. We conducted a pre-registered, large-scale online experiment ($n=1,699$) with a purpose-designed action role-playing game that allowed us to manipulate design features linked to these theories. In a 2×2 +control study design, we systematically varied amplification (feedback volume), success dependency (whether feedback was triggered by actions or actions succeeding at a challenge), and variability (whether feedback was diverse). We measured effectance, competence, curiosity, and enjoyment with self-report, and tracked voluntary playtime. We used structural equation modelling to analyse relations between design features, experiences, and behaviour.

To our surprise, curiosity emerged as the strongest enjoyment and only playtime predictor, while success dependency drove curiosity, effectance, and competence alike. Contrary to theoretical

predictions, enjoyment did not predict playtime, and amplified feedback *negatively* impacted effectance and competence. We interpret these findings to mean that even at low-level action-feedback loops, motivation and enjoyment arise from reducing uncertainty over action success, which depends on *legible* and *differentiated* amplified feedback. There can therefore be ‘too much of a good thing’ in amplified feedback where it occludes causal action-feedback links or flattens success gradations.

Together, our findings make several contributions to HCI: they evidence curiosity as a major, under-appreciated predictor of enjoyment and engagement in moment-to-moment interaction. They support that the effect of amplified feedback on enjoyment and engagement is almost fully mediated by curiosity and competence, and to a lesser extent, effectance. They directly test and support self-determination theory claims that granular competence feedback drives competence experience and enjoyment. They surface sense of agency as a potential moderator for ‘low’-level positive user experience. And they empirically support and psychologically specify prior design guidance on coherent juicy game feedback [26]: ‘overloading’ amplified feedback likely interferes with competence and effectance by occluding action-feedback links and making differences in task difficulty and success less salient.

2 BACKGROUND

HCI traditionally conceives system feedback as the output part of the input-output interaction loop crossing the “chasm of evaluation” [51]. That is, feedback helps a user form an accurate mental model of the current system state, e.g., whether it has correctly registered user input and shifted into the user-desired state in response. System feedback is thus framed as an informational usability factor, as can be seen in related design heuristics like “visibility of system status” or “offer informative feedback” [32, p. 176].

In contrast, game feedback—particularly juicy feedback—is framed as an active contributor to a positive user experience [17, 26]. This makes juiciness an “interaction aesthetic” [44] or positively beautiful quality of interaction at the “motor level” of bodily, moment-to-moment input-output loops [42]. While juiciness is absent from relevant systematic reviews of interaction aesthetics [42], it is arguably related to Löwgren’s “pliability”, an “almost exaggerated quasi-physicality of the objects that are manipulated” [44].

2.1 Juiciness

The term “juice” or “juiciness” was likely coined by independent game designers Kyle Gray and colleagues [25]:

“‘Juice’ was our wet little term for constant and bountiful user feedback. A juicy game element will bounce and wiggle and squirt and make a little noise when you touch it. A juicy game feels alive and responds to everything you do – tons of cascading action and response for minimal user input. It makes the player feel powerful and in control of the world, and it coaches them through the rules of the game by constantly letting them know on a per-interaction basis how they are doing.”

The term has since been in wide and varied use without a consensus definition [8, 17, 26, 33, 53, 61]. Juiciness is commonly framed as a

part of “game feel”, “the tactile, kinesthetic sense of manipulating a virtual object” afforded by “real-time control of virtual objects in a simulated space, with interactions emphasized by polish” [64, p. 6]. Thus, Hicks and colleagues’ [26] framework of 13 heuristic-style characteristics of game juiciness equates game feel and juiciness. A more recent literature survey by Pichlmair and Johansen [53] analytically separates juiciness as one of three components of game feel: (1) tuning physicality, (2) streamlining support, and (3) juicing amplification, which describes an “intensification of experience” by “adding feedback to emphasise, clarify, and amplify [...] the intended game feel” [53, p. 147].

“Juice” sometimes refers to person-external *feedback features* (e.g., “constant and bountiful feedback” [25]), specifically exaggeration, constancy, and immediacy [26, 53]. Other times, it refers to person-internal *experiential qualities* (e.g., “feels alive” [25], emphasis added). While some propose that juicing refers to amplifying whatever emotional impact a designer wants to evoke [8, 53], others tie it to particular experiences like aliveness, tactility, responsiveness, etc. [8, 24, 26, 53]. Complicating the matter, some suggest that (good) juicy feedback is hard to define and less an isolatable feature than a holistic quality of coherence, consistency, and integration with the game overall [26].

Despite these varied understandings, there are some common denominators: first, “juice” refers to moment-to-moment feedback on performing in-game actions – that is, it is *not* limited to success feedback on attaining difficult game goals [8, 17, 24, 26, 53, 61]. Second, it describes a “large amount” [26, p. 1] of feedback relative to user input (texts variously use terms like abundant, excessive, amplified, or exaggerated) [17, 27, 28, 33, 35, 61, 62]. Third, it “contribute[s] to a positive player experience” [26, p. 1].

To summarise, “juiciness” is presently an ambiguous social category. This necessitates any research to explicate which definition it operates with. Following the common denominators above, we here define and focus on juiciness as the design feature of immediate “excessive amounts of feedback in relation to user input” [53, p. 139] at the level of moment-to-moment interaction. Thus, our present study does not speak to the full range of features or qualities associated with juicy feedback (e.g., tactility or amplifying any designer-intended affect). To avoid confusion, we will use *amplification* or *amplified feedback* in the following whenever we refer specifically to immediate excessive feedback as our chosen definition.

2.1.1 Prior empirical work. Moving to the empirical literature, Juul and Begy’s first between-subjects study (n=46) of an experimental match-3 game found that juicy and non-juicy versions did not significantly differ in ease of use, rated quality, objective performance, or “feeling clever” [34]. In two within-person experiments (n=40 and n=32) across three games, Hicks et al. [27] found a positive effect of visual juiciness on aesthetic appeal across studies, no effect on usability or performance, and an effect on curiosity, competence, and presence/immersion in only one game. Another within-person experiment (n=36) with a VR game comparing a baseline, juicy, gamified and juicy+gamified condition found significantly higher competence, autonomy, relatedness, presence, interest/enjoyment, and preference ranking for the juicy over the baseline condition [28]. Two within-person studies (n=26, 38) that combined low and

high visual juiciness with no, low, and high haptic juiciness found that haptic juice significantly induced player preference as well as enjoyment, appeal, immersion, and meaning, but not mastery [62]. In a within-person study (n=113) of the automatic game design tool Squeezer [31], Johansen and Cook found that players significantly preferred game versions with expert-crafted juicy feedback over a non-juicy baseline, and significantly preferred 2 of 6 juicy feedback versions produced with their automated tool [30]. Finally, in a large between-subjects online study (n=3,018), Kao found evidence for a goldilocks effect: juiciness significantly affected competence, autonomy, presence/immersion, as well as interest/enjoyment, playtime, and performance; however, medium and high juiciness conditions outperformed low and extreme juiciness across all these measures [35].

In summary, the evidence base is characterised by mixed results, high measurement and operationalisation variance, low statistical power, and exploratory effect searching: out of 6 papers, 4 had samples between 26 and 46 per study, 2 used unvalidated single items or preference rankings, and of the 4 papers using validated scales, 3 ran full batteries of 7 up to 20 constructs, while only 2 pre-specified clear hypotheses for at least one individual construct. In tendency, players find games with juicy feedback more appealing, preferable, and enjoyable; for other constructs like mastery, immersion, usability, or objective performance, the evidence is mixed. This could be due to the above methodological variance and issues; the fact that studies used different understandings and operationalisations of juiciness; or because juicy feedback requires subtle and holistic design [26, 53], which different studies were variously (un)successful to ensure.

2.2 Candidate Psychological Mechanisms

So what intrinsically motivating experiences might amplified feedback afford? To answer this, we surveyed practitioner views and prior mediators proposed or tested in the literature, to then map these to existing theories of intrinsic motivation. Gray et al. identify being “powerful and in control of the world” [25] as an engaging feeling arising from juice; Swink [64, p. 10] proposes “the aesthetic sensation of control” and “the pleasure of learning, practicing and mastering a skill”. Schell suggests an interaction feeling “powerful and interesting” [61]. Hicks et al.’s developer survey lists visceral emotion, fantasy fulfilment, mastery, and meaningful actions, but did not link these to existing theories [26]. Juul suggests that juiciness could reduce enjoyment and ease of use (by creating extraneous cognitive load) and improve them [33], specifically by helping players “feeling clever” [34]. Deterding hypothesises that juiciness affords competence via “positive feedback on [...] actions and achievements” and curiosity via “unexpected variety” [17]. Pichlmair and Johansen [53] suggest that juicing “empowers the player”, but do not specify what psychological construct (if any) they refer to.

To map these accounts to existing theory, we assessed (1) which theories of intrinsic motivation have already been suggested and/or empirically supported in the literature, (2) whether these fit the experience labels proposed by practitioners, and (3) whether they provide a convincing rationale why amplified feedback affords enjoyment and motivation. This resulted in three selected constructs:

effectance, competence, and curiosity, detailed below. This selection excludes meaning, autonomy, relatedness, and presence/immersion which had been tested in prior work [27, 28, 35], since we saw no strong theoretical rationale why amplified feedback should evoke these. We also excluded aesthetic appeal, which is conceptualised as a functional positive player experience, not as an intrinsic motive [69].

2.2.1 Effectance. Robert W. White introduced *effectance* to psychology as the proposed single intrinsic motive: in play and other exploratory behavior, we seek the positive satisfaction of a “feeling of efficacy” [71, p. 322]. Building on this, Christoph Klimmt’s *Synergistic Multiprocess Model of Video Game Entertainment* [37, 39] posits effectance as one of three sources of gaming entertainment underlying interactivity, namely the basic, positively valenced sensation of our actions causing change in the world. Klimmt expressly distinguishes effectance from competence [57]: whereas effectance relates to ‘raw’, moment-to-moment causing effects, competence arises from the successful pursuit of difficult goals in longer episodic stretches of gameplay [37, pp. 81-85]. Klimmt suggests that games afford effectance through the temporal immediacy and disproportionality of game output on player input [37, pp. 76-81]. This exactly matches our common denominators of juicy feedback: a positive experience evoked by immediate amplified feedback on moment-to-moment inputs. It maps onto Gray and colleagues’ original idea that juicy feedback “responds to everything you do”, as well as a sense of “control” [25, 64] and being “powerful” [25, 61]; it arguably approximates “meaningful action” [26] and “empowerment” [53].

Manipulating the reliability with which a game registered player inputs as an operational proxy of effectance, Klimmt and colleagues found that players reported significantly lower enjoyment in the low-reliability/effectance condition [38]. Another study found self-reported effectance to be higher in interactive narrative games when players could affect the game state versus watching replays [56]. However, there have been no correlational or mediation analyses of the impact of effectance on enjoyment or behaviour, nor studies testing whether amplified feedback affords more effectance than non-amplified one.

2.2.2 Competence. Self-determination theory (SDT) has become a prevalent theory in games HCI [67], positing that we are intrinsically motivated to play games because it satisfies basic psychological needs for competence, autonomy, and relatedness, while ‘fun’ or interest and enjoyment are experiential signatures of this need satisfaction [59].

Moving to amplified feedback, SDT posits that *positive competence feedback*—environmental events suggesting that we are effective in pursuing challenging activities and goals—is one main causal antecedent of competence satisfaction [57, pp. 121-131], that is, feeling capable in engaging the world as well as growth in our capabilities [57, p. 86]. There is good evidence that positive (verbal) feedback enhances competence satisfaction, which leads to intrinsic motivation and engagement [57, pp. 153-157]. Video games afford immediate, consistent, and dense competence satisfaction via “rich, multi-level, effectance-relevant, positive competence feedback.” [57, p. 514] Restating amplified feedback in their own terms, Rigby and Ryan highlight *granular competence feedback* (see also [57, p. 514]:

In *Guitar Hero*, “there is the immediate feedback that players get on each and every strumming of their virtual guitar; through colourful flashes and sounds, the player sees and hears immediately whether or not they hit the note correctly. [...] these create mastery feedback loops [...] that instantaneously and consistently provide competence satisfactions [...] We call these [...] *granular competence feedback* because they have a one-to-one relationship to each of the player’s individual actions.” [55, p. 23].

Thus, while effectance theory posits that amplified feedback on all player *action* is already enjoyable and engaging, SDT makes competence *success-dependent*: Granular feedback informs players about “whether or not they hit the note correctly”; *if* they do so—if the feedback is *positive* about their abilities to achieve a challenging outcome—, *then* players feel competent [57, p. 152-3].¹ This matches practitioner phrases like “the pleasure of [...] mastering a skill” [64], “mastery” [25], or “feelings of competence and mastery” [26].

Only few SDT studies have directly explored game feedback, with mixed results [9, 48, 52]. One study found that competence satisfaction was higher when the textual result screen was positively rather than negatively phrased. Puzzlingly, positive feedback and competence correlated *negatively* with the likelihood to play the game again [9]. Another study compared competence-supportive and non-supportive game conditions (which included feedback manipulation among others), where the competence-supportive version showed higher competence satisfaction, enjoyment, and motivation for future play [52]. A study examining points as a form of granular competence feedback found no significant effect on competence satisfaction or intrinsic motivation [48]. The above-reviewed studies variously find significantly higher competence under juicy conditions or not. Notably, while Gray and colleagues’ original coinage [25] and effectance theory posit that amplified feedback on just performing actions is enough to make players ‘feel good’, Swink [64] and SDT posit that amplified feedback only leads to a competence experience when provided on an action that displays some player skill. To our knowledge, no prior work has tested these competing claims [27, 28, 35, 62]; juicy conditions either comprise both raw action feedback (e.g., particle effects and animations on player movement) and success-dependent feedback (animation and particle effect on enemy death) (e.g., [27]), or are not documented sufficiently to be able to tell whether both were manipulated (e.g., [35]).

2.2.3 Curiosity. Curiosity is widely studied as the main motivator of play and exploratory behaviour in cognitive science, neuroscience, ethology, and developmental psychology [1, 10, 12, 23, 36], while games research on curiosity is more nascent [22, 65, 66, 72]. Curiosity is commonly theorised as an emotion or intrinsic motive driving information-seeking [36], leading organisms to seek out or generate stimuli with “collative variables” like novelty, surprise, uncertainty, complexity, diversity, or ambiguity [46] that promise information gain. White [71, p. 322] himself stated that environmental responses to our actions need novelty and variance to evoke effectance:

“effectance is aroused by [...] difference-in-sameness. This leads to variability and novelty of response, and interest is best sustained when the resulting action affects the stimulus so as to produce further difference-in-sameness. [...] effectance motivation subsides when a situation has been explored to the point that it no longer presents new possibilities.”

This suggests that what makes juicy feedback engaging is (expected) information gain about causal action-feedback links and the range of possible game behaviour. This matches practitioner labels of juicy feedback like “interesting” [61], “feeling clever” [34] and characteristics like “uncertain outcomes” [26]. As proposed by Deterding in his “lens of juicy feedback”, “unexpected variety stokes curiosity” [17, p. 313]: the novel, surprising, uncertain, and complex ways that different items fall and explode and cascade into chain effects on every swipe keeps match-3 games like *Candy Crush Saga* interesting and pleasantly surprising, eliciting and satisfying curiosity as long as we haven’t learned the possibility space of feedback patterns such that feedback has become fully predictable.

Supporting this logic, game designer Greg Costikyan identifies uncertainty as a core motivational affordance of games, including performance uncertainty and randomness [14]. Developers in Hicks and colleagues’ survey similarly mention “uncertain outcomes” as an important game feel characteristic [26]. A recent qualitative study on moment-to-moment engagement in casual games found that uncertainty over the game’s response to each player’s action induced positively motivating curiosity [40]. In a lottery task study, behavioural and neurological signatures of curiosity tracked the buildup and release of outcome uncertainty [68]. Using an Asteroid-avoidance game, another study found that people waited to see outcome uncertainty-resolving feedback even if doing so did not improve their performance, suggesting that it is disjunct from competence [29]. One above-mentioned study found that juicy feedback drove curiosity [27]. That said, we are not aware of studies testing whether collative variables like variety in amplified feedback drive curiosity, nor severe tests of curiosity mediating enjoyment and voluntary engagement from amplified feedback.

2.3 The present study

To summarise, prior work has generated different conceptualisations of juiciness, descriptive synthesis of practitioner views, and initial, if mixed and mostly exploratory empirical data supporting that *amplified feedback can be enjoyable and motivating*, due to various possible (but largely unspecified) mechanisms and design features. Mapping this work to prior theory, we identified three candidate theories specifying different antecedents and mediators: effectance afforded by *all* amplified feedback; competence afforded by *success-dependent* amplified feedback; and curiosity afforded by *variable* amplified feedback. None of these theoretical predictions have been directly tested to date. Thus, our study aims to *directly test competing theory-derived predictions about how amplified feedback affords enjoyment and intrinsic motivation*, thereby also clarifying conflicting claims in prior practitioner views. Just as multiple motives and emotions can direct our behaviour in parallel [54], amplified feedback could operate via one, two, or all three predicted pathways in parallel. We therefore adopted and

¹Klimmt [37, 39] makes the same argument for competence/control, but again, treats it as a second, separate motive and process.

pre-registered an experimental and analysis strategy that would allow inference to the best explanation under the possibility of multiple parallel effects.²

Because success dependence and variability are qualifications of amplified feedback, we can only manipulate them if feedback is already amplified. This led us to create 2x2 experimental conditions that varied success dependence (yes/no) and variability (yes/no) of amplified feedback, with a dangling control condition with 'standard', non-amplified feedback. We specified two matching confirmatory structural equation models, one for amplification, one for variability and success dependence (see Figure 1).

For each construct, our pre-registration specified two hypotheses about experimental conditions and matching path predictions in our structural equation models, indicated as thick arrows in Figure 1, where the direct paths from amplification, variability, and success dependence to enjoyment tests for the absence of non-mediated effects.

H1 Effectance Based on theory, we expect that *amplified* feedback will generate enjoyment and subsequent voluntary engagement, fully mediated by effectance. This means: **H1a**: Enjoyment will be higher under non-success dependent amplified feedback than standard non-success dependent feedback. **H1b**: Voluntary engagement will be higher under non-success dependent amplified feedback than standard non-success dependent feedback. As causal paths, our model tested:

- There will be a significant positive correlation between amplification and effectance.
- There will be a significant positive correlation between effectance and enjoyment.

H2 Competence Based on theory, we expect that amplified feedback will cause competence experience *iff* it is success-dependent, which will lead to enjoyment and subsequent voluntary engagement. This means that: **H2a**: Enjoyment will be higher under success-dependent than non-success-dependent amplified feedback. **H2b**: Voluntary engagement will be higher under success-dependent than non-success-dependent juicy feedback. As causal paths, our model tested:

- There will be a significant positive correlation between success dependence and competence.
- There will be a significant positive correlation between competence and enjoyment.

H3 Curiosity Based on theory, we expect that varied amplified feedback will lead to enjoyment and subsequent voluntary engagement, fully mediated by curiosity. This means that: **H3a**: Enjoyment will be higher under variable than non-variable amplified feedback. **H3b**: Voluntary engagement will be higher under variable than non-variable amplified feedback. As causal paths, our model tested:

- There will be a significant positive correlation between feedback variability and curiosity.
- There will be a significant positive correlation between curiosity and enjoyment.

Voluntary engagement As all theories suggest that enjoyment leads to voluntary engagement, our model tested: There will be a significant correlation between enjoyment and voluntary engagement.

2.4 Modifications to preregistered analysis

We added (1) potential covariances between curiosity, effectance, and competence and (2) potential direct effects from design features (juiciness, success dependency, and variability) to voluntary engagement as exploratory analyses to our pre-registered models. These explore potential pathways to voluntary engagement not mediated by enjoyment since our enjoyment measure, the Intrinsic Motivation Inventory (IMI; 58), is somewhat SDT-specific, and not all intrinsic motivation theories predict that enjoyment mediates [50]. Figure 4 shows the final models. See SEM_Analysis.R for the pre-registered and main_analysis.R for the final analysis script on the OSF repository.³ The pre-registration and connected materials use the term “juicy” throughout to refer to immediate excessive feedback. The present manuscript replaces “juicy” with “amplified” throughout to signal the specific definition of juiciness we refer to.

3 METHODS

Our preregistration² as well as data and analysis scripts³ are accessible on the Open Science Framework (OSF). Detailed documentation of study conditions (including gameplay video) can be found on our online documentation.⁴

3.1 Study Game Platform

3.1.1 Final Game Platform. The game is a PC fantasy action role-playing game (RPG) controlled with keyboard and mouse, similar to, e.g., *Diablo*, where the player controls a knight. The game begins with a skippable ~1 minute tutorial teaching movement and combat controls. Players can move (using W, A, S, D), jump (spacebar), and sword attack (left-click). The game features five areas with different monsters (e.g., goblins, slimes, spiders) and quests (e.g., defeat the hobgoblin). Earning experience points from defeating monsters and completing quests levels up the player’s character, increasing their health and damage dealt. If a player dies, they respawn at the beginning of their current area. Monsters move and attack differently, and some have special abilities, e.g., giant spider attacks immobilise the player. The environment further features grass and butterflies that respond if the player character moves through them.

3.1.2 Development and Validation. Over the course of 20 months, we developed and tested the game using Rapid Iterative Testing and Evaluation (RITE) with university students recruited through a university’s newsletter as test users. We conducted four rounds of testing with 3-4 participants each. Participants played all study conditions in randomised order, then were interviewed using semi-structured interviews (see OSF repository for interview script³), and were compensated with a \$15 gift card. Feedback from each round was synthesised and translated into changes to the game for the next round until no additional issues were raised. This was interwoven

²Pre-registration: <https://osf.io/yvu3c/>

³OSF repository: <https://osf.io/sveb2/>

⁴Study game platform documentation: <https://arpgdocs.readthedocs.io/>

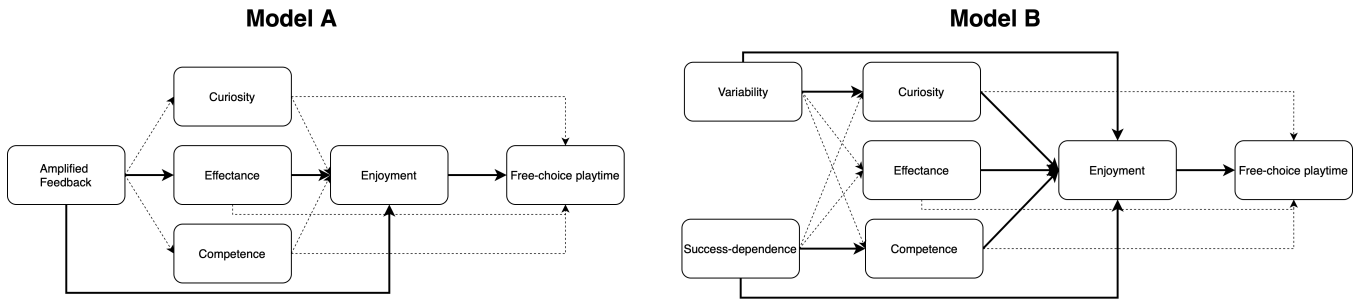


Figure 1: Structure of the two confirmatory structural equation models.



Figure 2: Screenshots of conditions: mid-swing (top row), hitting an enemy (middle row), and hitting with no enemy present (bottom row). Success-Dependent (+SD) impact effects are triggered only when hitting an enemy. Non-Success-Dependent (-SD) impact effects appear with every swing, regardless of a hit. Varied (+V) effects are selected randomly each time they are triggered. Non-Varied (-V) effects have a single effect.

with three additional rounds of feedback and iteration with the authors and external experts to ensure the game’s suitability for the study.

Absent well-specified consensus definitions or operationalisations of juiciness, we considered but discarded expert member checking with game developers as a construct validity test: since game developers have diverging and differently well-developed views on what constitutes juiciness [8, 24, 26], they may not have aligned with each other nor our tested conceptualisation of juiciness. Instead, we directly involved creators of the empirically grounded game juiciness framework [26] as an external validity check that our amplified conditions fit their characteristics of game juiciness, while our standard, non-amplified condition did not. This led, e.g., to the addition of rising butterflies and moving grass animations to match their characteristic B4, “ambient feedback” [26]. Similarly, we ended up using commercial Unity animation assets (e.g., sword swing animations from the “Magic Slashes FX” package) that fit

the fantasy setting of the game, in line with characteristic A2, “thematic coherence.” The involved creators were satisfied that our final version fit their characteristics of game juiciness.

For construct validity, we mainly aimed to ensure that our conditions operationalised our definition of amplified feedback and theoretically relevant design features (success dependence, variability), as our main aim was theory-testing. Matching our definition, all manipulated feedback is directly caused by and immediately follows moment-to-moment player actions: moving and sword attack. Feedback comprised direct movement and sword swinging sound and animation, and indirect grass movement, butterfly movement, enemy hit, enemy kill, potion pickup, and experience point gain/levelling up depending on when and where the player moved or attacked. During RITE, playtesters played all conditions and were then openly interviewed to describe differences they noticed between conditions, and we checked that their descriptions matched our desired differences and no condition was unclear.

For *amplification*, we further checked that playtesters described amplified conditions as having stronger or more feedback than the

standard condition, by asking them to identify differences between conditions. At the same time, we expressly prompted playtesters to report whether the standard condition provided sufficient feedback such that they still understood gameplay and the consequences of their actions. For *success dependence*, our manipulation needed to operationalise SDT predictions that competence arises from at least marginally difficult and skill-requiring tasks. We chose to manipulate sword attacking as a basic, thematically fitting game action that (a) produces plausible immediate feedback on actuation (swooshing sound and animation) for amplified non-success-dependent feedback, and (b) also involves a non-trivial challenge with plausible amplified success-dependent feedback, here: sound and animations on actually hitting a moving enemy, as well as enemy death. To ensure the sword attack involved challenge, we balanced the game during RITE such that players would sometimes miss enemies on attack. To operationalise *variability*, varied conditions used randomised different sounds and animations from the commercial Unity asset packages, while the non-varied conditions used a single fixed sound and animation from the same package, again aiming for thematic coherence. We asked particular interview questions to check that the audio-visual effects for varied feedback were all similar in perceived strength for each action as well as comparable to the non-varied effect for each action.

To avoid game difficulty and learnability confounds, we checked whether the game was easy to learn and playtesters performed roughly equally well across conditions. To control for variability in performance, we first ran a benchmarking procedure with 80 online participants recruited from Prolific. The benchmark tested the most performance-intensive actions in the game for each condition and recorded the frames per second (FPS). Based on these results, we (a) optimised code to improve performance (focused on collision detection computations), (b) created a set of minimum machine requirements for participants, and (c) devised a short benchmarking test that ran at startup in the main study. This test, performed invisibly to the user behind a loading screen, simulated the most performance-intensive situation: attacking an enemy consecutively with background juicy environmental features present. We performed this test for all five conditions and then took the lowest average FPS across conditions and set this as the maximum frame rate for that participant. This controlled for variance in frame rate across conditions and participants—a participant’s FPS will always be limited by the most performance-intensive condition, regardless of their actual assigned study condition. If that lowest FPS was less than 30, we informed the user that their machine doesn’t meet requirements and excluded them from completing the study.

3.2 Conditions

The study used a between-subjects, 2 x 2 design with amplified (A) and variously success-dependent (+/-SD) and varied (+/-V) conditions plus an additional dangling non-amplified control group (STND), leading to 5 total groups that participants were randomly assigned to (Figure 2 and Table 1; see the online documentation⁴ for condition details including video, audio, and source code).

3.3 Measures

3.3.1 Dependent variables. We measured **effectance** using Klimmt et al.’s effectance in games scale [38], specifically, the adapted version that is recommended in Ballou et al.’s validation study [4] (i.e., items 5, 7, 9, and 11, measured on a 7-pt Likert scale from 1 (“not at all true”) to 7 (“very true”). We use the adapted version because the original scale was found not to be unidimensional, suggesting it is not construct-valid [4]. We verified using confirmatory factor analysis that the adapted 4-item measure functioned well in our sample (strong model fit and all item loadings > .69; see OSF repository³). We measured **curiosity** and **competence** using the curiosity and mastery subscales of the Player Experience Inventory/PXI [69], measured on 7-pt Likert scales from -3 (“strongly disagree”) to 3 (“strongly agree”). We measured **enjoyment** and **self-reported intrinsic motivation** using the interest/enjoyment subscale of the Intrinsic Motivation Inventory/IMI [58], measured on a 7-pt Likert scale from 1 (“not at all true”) to 7 (“very true”). We omitted all other subscales from the PXI and IMI since they were not of theoretical relevance to our study and the PXI and IMI subscales are all fully disjunct. Both PXI and IMI are well-validated scales widely used in games HCI and prior juiciness research [27, 28, 35], and showed very good model fit with high loadings in our sample (all items loaded > .74 with the exception of one IMI item at .59; see OSF repository³). Finally, we operationalised **voluntary engagement** or the behavioural expression of intrinsic motivation as voluntary time on task, in line with common practice [58]. After 10 minutes of mandatory playtime, players filled in a survey on the measures above and were informed that they could now end the experiment or continue playing for as long as they want. We measured voluntary engagement as minutes of play after this point.

3.3.2 Other measures. To describe our sample and check for potential confounds, we additionally measured *prior play experience* for video games and action role-playing games using the questions “How would you rate your prior experience playing video games?” and “How would you rate your prior experience playing action role-playing video games?” both on a 7-pt Likert scale from 1 (“minimal”) to 7 (“extensive”). We also asked “How many hours of video games do you play approximately per week on average?” To test that *game difficulty* did not accidentally co-vary across conditions, we measured challenge using the PXI challenge subscale [69] on a 7-pt Likert scale from -3 (“strongly disagree”) to 3 (“strongly agree”). We also included one attention check item.

3.4 Sample Size Determination

Although we use validated questionnaires, these have not yet been used in studies to determine effect size benchmarks (e.g., using anchor-based methods [2]), and prior literature was insufficient to estimate expected effect sizes or specify a smallest effect size of interest, precluding an *a priori* power analysis. We followed recommended practice in these circumstances [41] to instead base our sample size on the largest number of people that could be collected with the monetary resources available to the project. This maximally affordable sample recruited was N=1,700. Based on past experience, we expected to remove up to 200 participants who failed our attention check, meaning we would have a sample of N=1,500 valid participants.

Table 1: Brief descriptions of conditions.

	Description
STND	<i>Standard Non-Varied Feedback</i> without amplification. Sound effects on swinging, hitting, and enemy death. No enemy death animation (the enemy disappears).
A-SD-V	<i>Amplified Non-Success-Dependent Non-Varied Feedback</i> . Exaggerated audiovisual impact effects occur even without hitting an enemy, but there is only one effect for each feature, e.g., one swing effect, one impact effect, one enemy death sound effect, etc.
A-SD+V	<i>Amplified Non-Success-Dependent Varied Feedback</i> . Exaggerated audiovisual impact effects occur even without hitting an enemy, and there are many possible effects, e.g., many swing effects, many impact effects, many enemy death sound effects, etc.
A+SD-V	<i>Amplified Success-Dependent Non-Varied Feedback</i> . Impact effects will occur only when the player successfully hits an enemy, but there exists only one effect, e.g., one swing effect, one impact effect, etc.
A+SD+V	<i>Amplified Success-Dependent Varied Feedback</i> . Impact effects will occur only when the player successfully hits an enemy and are varied, so there exist many possible effects, e.g., many swing effects, many impact effects, etc.

To determine the effect size sensitivity of this sample size, prior to collecting data, we ran a Monte Carlo simulation in R using lavaan version 0.6-11. Simulation results indicated that with 1500 participants, we have 89–91% power to detect standardised effects of $\beta = .1$ on the pathways from curiosity, effectance, and competence to enjoyment; and 77-80% power to detect standardised effects of $\beta = .1$ on the pathways from juiciness, variability, and success dependence to curiosity, effectance, and competence. By common benchmarks, our study thus had adequate power to detect small ($\beta < .2$) effects [19]. Further details and results of our simulations are available in the script SEM_Analysis.R found in the preregistration.

3.5 Participants

We recruited 1,706 participants via the online recruitment platform Prolific. Out of these, only 6 failed the benchmarking test and only 1 was removed for failing the attention check question. This left us with $n=1,699$ valid participants for an average 339.8 participants per condition ($SD=22.1$) (see Table 1 for n per condition). We used simple randomisation, which can naturally lead to some variation in group sizes. Our group sizes deviated at most 2% from the average 20%, which is considered acceptable especially at large sample sizes, and does not affect the robustness of our results. 60.8% of participants identified as men, 36.4% as women, 2.1% as gender variant/non-conforming, and 0.7% as transgender. Participants had an average age of 27.1 ($SD=8.0$). Using Prolific’s pre-screening criteria, the task was available only to participants who were at least 18 years old, had English fluency, and had a desktop computer with working audio. Participants were told before accepting the task that they must own a Windows or Mac machine that meets our minimum requirements which are specified in the preregistration. There was no limitation on geographic location. Participants were from South Africa (18.1%), Poland (17.5%), Portugal (15.8%), Italy (8.3%), Mexico (8.1%), United Kingdom (6.6%), Greece (4.7%), Spain (4.4%), Hungary (2.6%), Chile (1.8%), Netherlands (1.6%), Czech Republic (1.5%), France (1.4%), Germany (1.1%), Canada (0.8%), and a combined 5.9% from 15 other countries. The Institutional Review Board (IRB) at the first author’s institution approved the study. All participants provided informed consent, and were paid an average US\$11.29 per hour, including excluded participants.

3.6 Procedure

Participants are tasked to fill out a consent form and then randomly assigned to a condition. They then undergo an audio check, during which they are required to type a spoken English word. Next, they

download and play their condition version of the game that is compatible with their operating system (Windows or MacOS). Upon launching the game for the first time, the benchmarking test is automatically conducted, as described in Section 3.1.2. If a participant fails the benchmarking, they are instructed to return the task on Prolific.

Participants passing the benchmark are instructed to play the game for a minimum of 10 minutes. At the 10-minute mark, participants are automatically prompted with an in-game survey containing measures for effectance, curiosity, competence, and enjoyment. The order of survey measures and items within each measure are randomised for each participant. Upon completion of the survey, participants are informed that they may exit the game whenever they wish. Voluntary engagement time is automatically logged thereafter. Upon exiting the game, participants complete a post-survey that includes measures for prior play experience and challenge, followed by a demographics questionnaire.

3.7 Analysis

We fit two structural equation models to analyse our results, specified in Figure 1. Models are fit using the `sem()` function of lavaan version 0.6-11 using the robust MLR estimator. Precise model syntax and model fitting code are available in the `main_analysis.R` script on OSF.³

In Model A, we assess the effect of amplification on effectance, competence, and curiosity, and ultimately enjoyment and free-choice playtime. We include only the standard feedback condition (STND) and the amplified non-success-dependent, non-varied feedback condition (A-SD-V); this allows us to isolate the effect of amplification itself. In Model B, we assess the effects of success-dependence and variability on effectance, competence, and curiosity, and ultimately enjoyment and free-choice playtime, using all amplified conditions.

For both Model A and Model B, we conduct model comparisons with alternative models where the mediated pathways (i.e., the pathways from the design feature \rightarrow effectance, competence, and curiosity; and from effectance, competence, and curiosity \rightarrow enjoyment) are fixed to 0. We compare model fit between the restricted and unrestricted models using ΔCFI , $\Delta RMSEA$, and $\Delta \chi^2$ statistics.

We evaluate the fit of each SEM model using standard model fit indices (CFI, RMSEA, SRMR, χ^2). We do not specify cut-off criteria in advance, given evidence that these are overly simplistic [11]. Instead, we decide whether the models fit appropriately using all these indices in tandem. Where the models do not fit sufficiently

well, we use residual covariances, modification indices, and domain expertise to make adjustments that improve the model fit, and transparently report all decisions and alterations made. If we achieve a well-fitting model, we infer the presence of a significant effect when the p-value of a particular pathway is $< .05$.

4 RESULTS

4.1 Descriptive Measures

Descriptive results can be seen in Figure 3 and Table 2. All effect sizes are presented in Figure 4 as unstandardised coefficients, which we believe are more informative than standardised ones, including levels of significance.

4.1.1 Prior Play Experience. Participants reported playing an average of 11.9 (SD=13.6) hours of video games per week, rated their experience playing video games $M=5.2$ (SD=1.6) on a 7-pt Likert scale, and rated their experience playing action-roleplaying video games $M=4.7$ (SD=1.8) on a 7-pt Likert scale. ANOVAs found no significant differences between conditions on hours of video games per week ($F[4, 1694]=0.820, p=0.513, \eta_p^2=.002$), experience playing video games ($F[4, 1694]=1.249, p=0.288, \eta_p^2=.003$), and experience playing action RPGs ($F[4, 1694]=1.006, p=0.403, \eta_p^2=.002$), suggesting these did not confound results.

4.1.2 Challenge. Participants reported an average challenge of 0.76 (SD=1.3) on a 7-pt Likert scale from -3 to 3. An ANOVA found no significant difference between conditions ($F[4, 1694]=2.321, p=0.242, \eta_p^2=.003$), suggesting that difficulty did not confound results.

4.2 Model A: Amplification and Effectance

Model A ($n = 678$, Figure 4, top) fit the data well (CFI = .976, RMSEA = .05290%CI [.045, .059], SRMR = .038). Model fit was drastically improved compared to a reduced model in which the mediation pathways were constrained to 0 ($\Delta\chi^2(6) = 515.74, p < .001$), supporting the mediating role of our candidate mechanisms.

Amplified feedback unexpectedly led to significantly *lower effectance* ($-.19, p < .05$) and *competence* ($-.43, p < .001$) compared to the standard condition. We also did not find the predicted association between enjoyment and voluntary engagement. The only predicted relation supported by the data is a significant (if small) positive association between effectance and enjoyment ($.13, p < .05$). We thus reject **H1a** and **H1b**.

Exploratory analyses similarly found no direct effect of amplification on enjoyment or free-choice playtime. Effectance significantly covaried with curiosity ($.37, p < .001$) and competence ($.58, p < .001$), and competence with curiosity ($.57, p < .001$). Curiosity had the strongest and only significant association with enjoyment ($.75, p < .001$), and the only significant association with voluntary engagement ($.99, p < .001$). Competence was significantly associated with enjoyment ($.27, p < .001$).

4.3 Model B: Success-dependence and Competence, Variability and Curiosity

Model B ($n = 1699$, Figure 4, bottom) fit the data well (CFI = .977, RMSEA = .048 90% CI [.044, .052], SRMR = .038). Model fit was drastically improved compared to a reduced model in which the

mediation pathways were constrained to 0 ($\Delta\chi^2(9) = 1365.6, p < .001$).

As predicted, success-dependence led to significantly greater **competence** ($.45, p < .001$), and competence was again significantly positively associated with enjoyment ($.26, p < .001$). Competence and voluntary engagement were higher in success-dependent than non-dependent juicy conditions (Table 3); we found no significant direct effects of success-dependence on enjoyment or voluntary engagement. This supports **H2a** and **H2b** and our broader SDT-based hypothesis that success-dependent amplified feedback drives enjoyment mediated by competence. Contrary to prediction, enjoyment again was not significantly associated with voluntary engagement, which we return to below.

Also counter to our predictions, variability showed no significant correlation with **curiosity**, while as predicted, curiosity strongly correlated with enjoyment ($.76, p < .001$). Varied conditions showed slightly higher enjoyment than their non-varied counterparts, mirrored in a small but significant correlation between variability and enjoyment ($.11, p < .05$).

Further exploratory analyses showed that variability had no significant associations with competence and effectance, while success-dependence had significant positive effects on effectance ($.26, p < .001$) and curiosity ($.29, p < .001$). Curiosity had again the strongest and only significant correlation with voluntary engagement ($.86, p < .001$). Again, all three mediators significantly covaried (all $p < .001$): effectance with curiosity ($.36$) and competence ($.53$), competence with curiosity ($.53$). This leads us to accept **H3a** (enjoyment is higher under variable feedback) and reject **H3b** (voluntary engagement is higher under variable feedback).

5 DISCUSSION

We will discuss ramifications of our results for each theory in turn, followed by general reflection on enjoyment and voluntary engagement and games HCI, prior work on juiciness, and implications for design.

5.1 Amplification, Effectance and Klimmt's Multi-Process Model

Surprisingly, **amplification had a significant negative effect on effectance and competence**. Is this a case of juicy feedback frustrating users via extraneous cognitive load, as Juul [34] proposes? We think not, or rather, we think this explanation is a case of imprecise theory use by Juul. In cognitive load theory [63], extraneous load refers to information unnecessary for *learning* something that interferes with limited working memory; swapping gems (as studied by Juul [34]) or in our case, actuating movement or sword attacks, arguably involve no such information memorisation.

We propose that our results can instead be explained by the amplified condition (A-SD-V) tested in Model A unintentionally impeding so-called *outcome binding*—attributing an observed event to one's prior intentional action [43]. Outcome binding is a constituent subprocess of *sense of agency* or “the feeling of control over actions and their consequences” [49]. Both may sound the same as effectance, but describe a ‘neutral’, low-level cognitive process, while effectance describes a (resultant) positively valenced affective-motivational state.

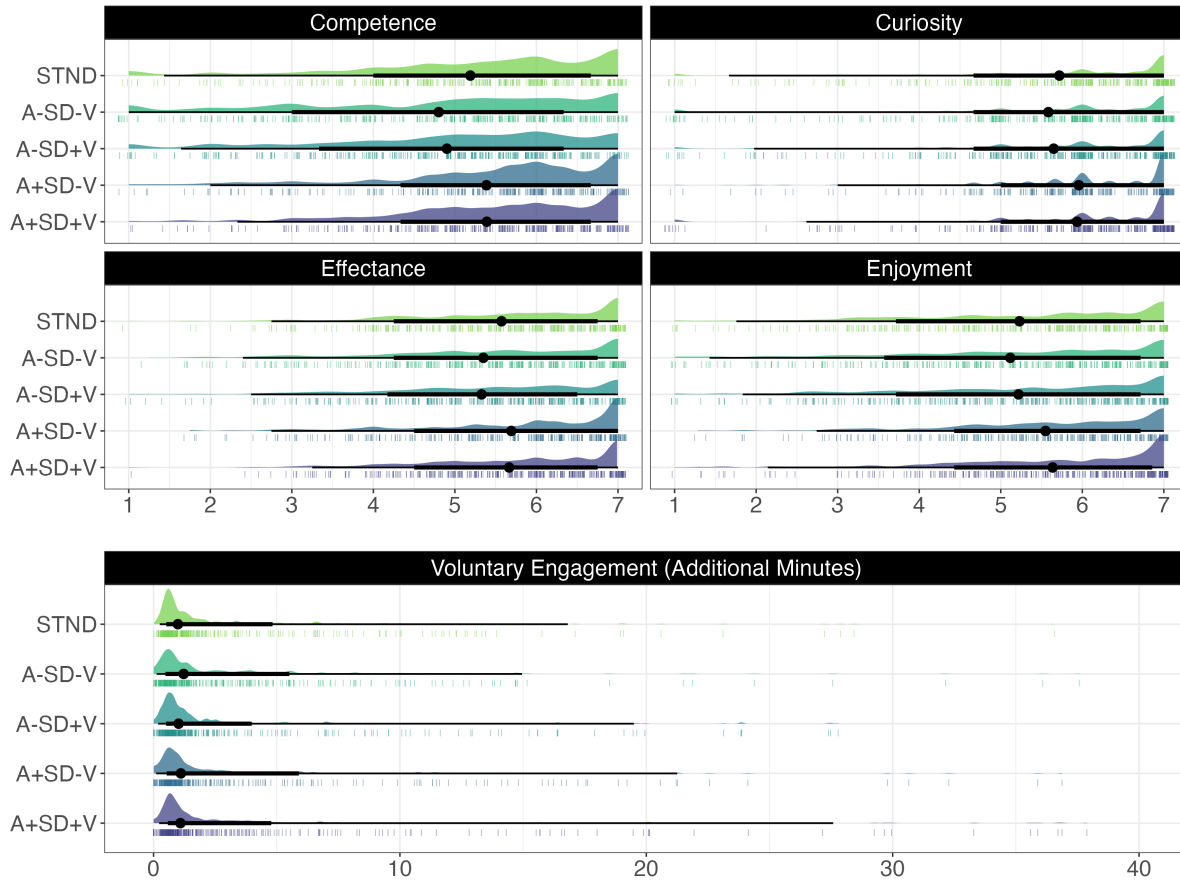


Figure 3: Descriptive statistics for player experience and voluntary engagement across conditions. STND = standard non-varied (control); A-SD-V = amplified, non-success-dependent, non-varied; A-SD+V = amplified, non-success-dependent, varied; A+SD-V = amplified, success-dependent, non-varied; A+SD+V = amplified, success-dependent, varied. 20 data points (1.2%) with more than 40 min voluntary playtime are omitted for legibility.

Table 2: Descriptive statistics for player experience constructs and playtime across each condition (numeric).

	N	Effectance		Competence		Curiosity		Enjoyment		Playtime	
		M	SD	M	SD	M	SD	M	SD	M	SD
STND	333	5.57	1.18	5.19	1.43	5.72	1.43	5.23	1.45	13.31	6.17
A-SD-V	345	5.35	1.33	4.80	1.63	5.58	1.48	5.12	1.54	13.65	7.58
A-SD+V	358	5.33	1.29	4.90	1.49	5.65	1.37	5.22	1.48	13.47	7.27
A+SD-V	305	5.69	1.24	5.39	1.34	5.96	1.08	5.55	1.23	14.76	10.18
A+SD+V	358	5.66	1.11	5.39	1.26	5.94	1.24	5.64	1.27	14.19	8.00
ALL CONDITIONS	1699	5.52	1.24	5.13	1.45	5.77	1.34	5.35	1.42	13.86	7.90

How did condition A-SD-V impede outcome binding? Inspecting gameplay video of both the control (STND) and A-SD-V (Figure 5) shows that in the latter, the large light glow ball of the Weapon Swing animation often visually engulfs and occludes subsequent triggers of the same animation as well as separate animation effects for hitting or killing an opponent. Thus, players may not have sensed that they hit repeatedly or caused enemy hits or deaths. Our logic here is simple: To feel positively efficacious or competent over

causing an observed event, we must first sense *that* our actions caused it – and that may have been impeded by sword attack feedback occluding success feedback. In line with this logic, prior HCI work on sense of agency finds it impeded by delayed (i.e., laggy or nonresponsive) as well as incongruous or unreliable feedback [43]. This maps Klimmt’s proposal that effectance depends on the temporal contiguity [39], and his finding that unreliable feedback is less enjoyable [38].

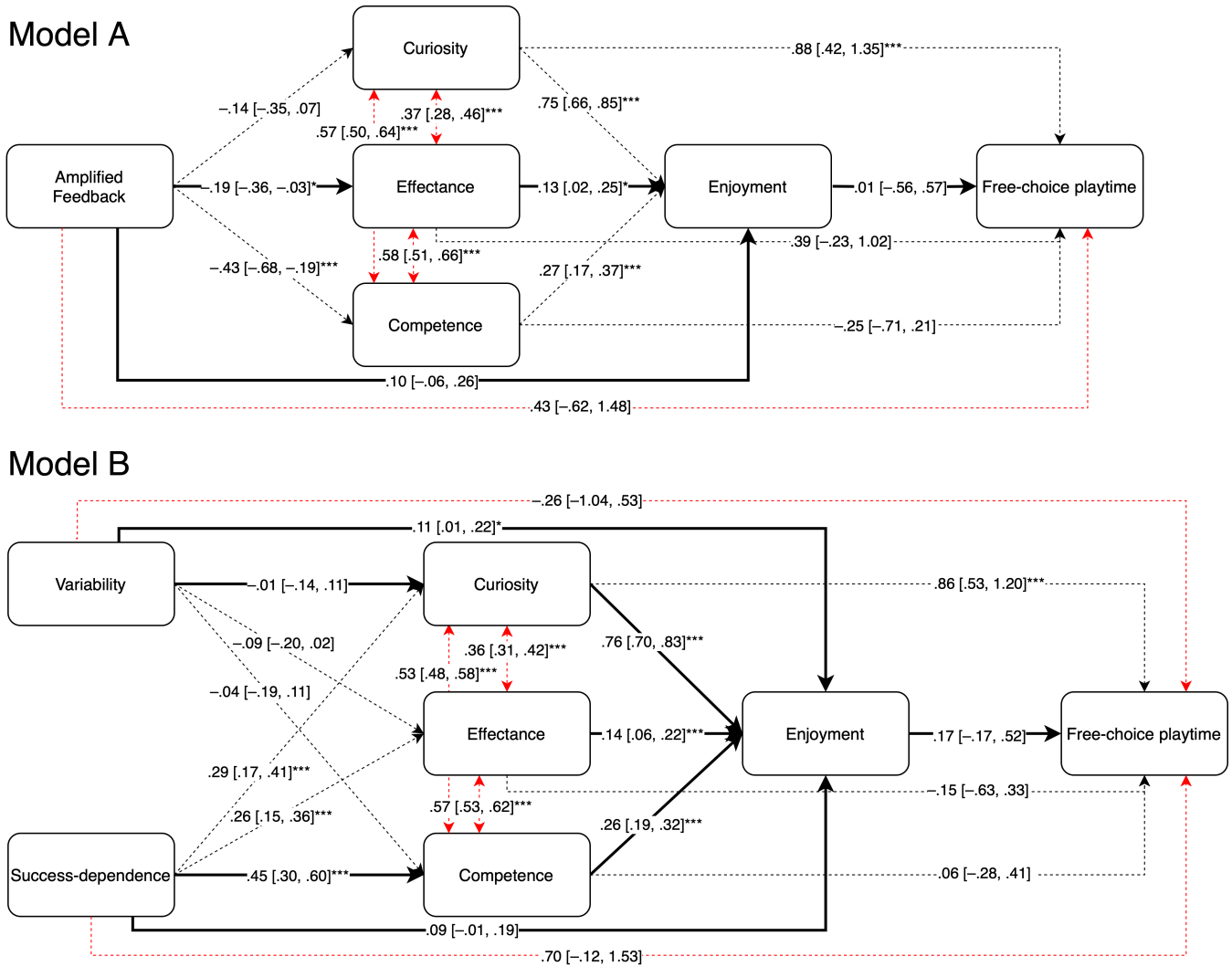


Figure 4: Results of the two confirmatory structural equation models. Solid black paths reference primary hypothesis tests, with dashed lines for all other relations. All coefficients and 95% CIs (in brackets) are unstandardised effects: effects on curiosity, effectance, competence, and enjoyment represent points on a 7-pt Likert scale, and effects on free-choice playtime refer to minutes. Red paths were not originally included in the preregistered analysis plan.

* $p < .05$; ** $p < .01$; *** $p < .001$.



Figure 5: Still frames depicting three consecutive attacks in sequence from left to right. The top row (STND) and bottom row (A-SD-V) highlight the contrast in visual occlusion between the two conditions.

Next, we turn to the major disagreement between Klimmt's effectance-based Multi-Process Model [39] and SDT [57]. SDT (like Swink [64]) argues that enjoyable juicy feedback must be success-dependent, because positive (competence) experience requires observing attaining intended and at least marginally difficult tasks, while Gray and colleagues [25] and Klimmt argue that merely observing change caused by actions is already a positive (effectance) experience. If the latter were the case, success-dependence should not affect effectance. Yet we found a small-to-medium positive association between the two (.26, $p < .001$). A plausible counter-argument would be that our success-dependent conditions gave players more *different* action-consequence links to observe: swinging *and* hitting *and* killing. Also, we found significant (if small) independent associations between effectance and enjoyment. Thus, we do not see sufficient warrant to accept or reject Klimmt's proposal that effectance is a fully separate mediator. Further research with more careful manipulations is warranted here.

5.2 Success Dependence, Competence, and SDT

Our study directly tested and found support for SDT claims on granular positive competence feedback [55, 57]: **success-dependent amplified feedback caused competence satisfaction with a moderately large effect size, which was significantly associated with enjoyment**, while success-dependent amplified feedback had no significant direct effect on enjoyment, supporting mediation.

The observed negative effect of amplification on competence can be explained in at least two SDT-congruent ways: (1) if amplified feedback effects are trivial to accomplish (just by actuating a sword swing), they do not engender competence over doing well at something challenging, but may also comparatively 'cheapen' the attainment of more difficult tasks (like hitting or killing enemies) triggering the same intensity of competence feedback. Hicks and colleagues [27] offer a similar explanation for why juiciness increased competence in only one of three games in their study. (2) As suggested above, our amplified feedback condition may have impeded action-feedback outcome binding, which is a logical precondition for feeling competent over observed feedback. Interestingly, SDT has long proposed that competence satisfaction from a challenging task only occurs under perceived self-determination or having intentionally engaged in the task [15, 16]. This has been later reframed into competence satisfaction only occurring under parallel autonomy satisfaction [57]. Our findings suggest that at a moment-to-moment level, outcome binding ("my action caused this effect") as an aspect of agency could be the actual precondition of competence, or a further one in addition to intentional binding ("I caused this action"). A recent review has highlighted a similar ambiguity of autonomy and agency concepts in HCI, variously referring to different aspects of causality and identity [6]. Thus, our findings point to plausible 'microscopic' feedback design features (outcome binding and success gradation) as preconditions not discussed or tested in prior work.

5.3 Variability and Curiosity

Curiosity was the strongest predictor of enjoyment by far and the *only* construct significantly correlated with voluntary

engagement. Every point increase in curiosity was associated with a .75 (Model A)/.76 (Model B) increase in enjoyment and .88/.86 additional minutes of gameplay, or a 10.8% gain of the average playtime of 7.9 minutes.

However, contrary to our hypothesis, **varied amplified feedback showed no significant association with curiosity, while success dependence did**. One possible explanation comes from recent accounts of curiosity as expected uncertainty or prediction error reduction [13, 18, 21]: In varied state spaces with irreducible aleatoric uncertainty (that is, true stochastic randomness like dice rolls or our varied feedback), once possible states are known, there is no further satisfying uncertainty reduction to be had: we know which sides of the die can show or which animation loops can be triggered. In contrast, in state spaces where variance follows a reliable generating rule, even after the possible states are known, we can continue to satisfyingly reduce epistemic uncertainty over learning the rule when and why each state is likely to occur [60]. This would also explain why success dependence increased curiosity: as Kumari and colleagues [40] found, even at moment-to-moment levels of input-output loops, people are curious to reduce "outcome uncertainty" over whether they succeed at each action. Our success-dependent non-varied conditions could have made action-outcome links with uncertain success (i.e., enemy hits and kills) more salient than their counterparts: something players perceived that they could learn to reliably predict and cause better and better.

In summary, our findings support that curiosity mediates the effect of amplified feedback on enjoyment and engagement, but is afforded by reducible outcome uncertainty in success-dependent feedback, not by uniform amplified nor (truly) randomly varied amplified feedback. While nascent work has suggested curiosity as a potential driver of gaming enjoyment and motivation [45, 65, 66], our study provides (to our knowledge) the first quantitative empirical support in games HCI that curiosity strongly drives enjoyment and voluntary engagement. Prior games HCI work has suggested gameplay forms like exploration afforded by open world games or social simulations afford curiosity [22]. Our data supports that curiosity already operates at the level of game feel. Our findings also suggest that aleatoric uncertainty or stochastic randomness may not uniformly add to curiosity (as Costikyan [14] seems to imply).

5.4 Enjoyment and Voluntary Engagement

SDT as the main theory of intrinsic motivation proposes that proximate experiences like competence need satisfaction lead to felt intrinsic motivation (here: enjoyment/interest, measured with the IMI), which leads to intrinsically motivated behaviour, here measured as voluntary engagement [57]. In line with this, our data showed expected motive-enjoyment correlations for effectance, curiosity, and competence. Although voluntary playtime substantially varied between conditions (lowest 6.17 min in STND, highest 10.18 min in A+SD-V), **we found no significant enjoyment-playtime association**, as predicted by SDT. Meanwhile, curiosity was strongly directly associated with voluntary playtime in both models. Put differently, players played more when they were curious in the game, but not when they enjoyed it.

We see one potential explanation in well-evidenced neurological differential mechanisms for ‘liking’ and ‘wanting’ [7]. Liking refers to positive hedonic experience upon consummatory behaviour (e.g., enjoying licking ice-cream), while ‘wanting’ or incentive salience refers to neural systems driving both anticipatory approach and consummatory behaviour even in the absence of liking (e.g., going to the ice cream store in order to buy and eat an ice cream). Enjoyment measures like the IMI or player experience measures like ‘mastery’ in the PXI chiefly capture post-hoc hedonic ‘liking’ experiences, while curiosity-motivated behaviour has recently been linked to ‘wanting’ [20]. More specifically, Murayama’s reinforcement learning model of curiosity suggests that a ‘wanting’ state (manifesting as a sense of interest or curiosity) energises behaviour, while satisfying this want generates enjoyment or ‘liking’ experiences of surprise, insight, learning, or uncertainty reduction, reinforcing incentive salience [50]. Put differently, our players played voluntarily if they *wanted* to satisfy their curiosity about a novel game or whether they can beat it. Satisfying this curiosity then generated positive *liking* experiences of enjoyment, reinforcing their *wanting* to play. This would fit that curiosity correlated with enjoyment, but also, *and independently*, with voluntary engagement, and the lack of enjoyment-engagement links. We think this points games HCI toward curiosity and expectations as potential under-appreciated mediators of actual play behaviour.

Summarising across theories, we can say that the impact of amplified feedback on enjoyment is mediated by curiosity and competence (and potentially, effectance), but not fully shaped by the design features predicted in each theory: success-dependent feedback affords competence as predicted, but also curiosity and effectance, while random feedback variety has no impact on curiosity.

5.5 Implications for Juiciness and Games HCI

We highlighted throughout that the current ambiguity of the term “juice” urges clarity and care in definitions, operationalisations, and delimitations. We here specifically tested and speak to variants of juiciness understood as *amplified* feedback in immediate response to moment-to-moment player input as a common denominator of much prior work [17, 26, 28, 33, 35, 61, 62]. Here, our findings support prior observations [27, 34, 35] that amplified feedback on its own does not necessarily lead to enjoyment or intrinsic motivation, and can even detract from them. Further, our findings add psychological specification to prior design guidance. First, by sword swinging feedback engulfing enemy hit and death feedback, our A-SD-V condition arguably violated characteristic C3 of Hicks and colleagues’ game juiciness framework [26]: “Unambiguous: Can information be connected to actions and only interpreted in one way?” As unpacked in detail above, our findings suggest an explanatory mechanism for this characteristic not previously discussed in the literature: when amplified feedback occludes or becomes undifferentiable from other action feedback, it may hamper *action binding* [49] as a sub-component of sense of agency, which is a logical precondition for both effectance and competence.

Second, to test competing claims of effectance theory and SDT, our non-success-dependent conditions ensured equally strong amplified feedback on both acting (sword swinging) and succeeding

(hitting and killing an enemy). This may have violated characteristics A4 (“Feedback Coherence: Does feedback reflect the importance of the event?”), B3 (“Highlighting: Are feedback elements that highlight information in harmony with other systems?”), and C4.A (“Relevant: Is feedback giving in response to game critical events or is feedback received on minor player actions that require no further action.”) [26]. All three principles (A4, B3, C4.A) imply a kind of coherent information hierarchy, which Hicks and colleagues have tied to an informational function [26, p. 6]. By inadvertently manipulating this hierarchy, our conditions may have violated good “game juiciness” as specified in [26]. Yet in doing so, they also provide empirical evidence *that* coherent differentiated feedback has a direct motivational impact, not just an informational one, and *that* this impact is mediated by curiosity, effectance, and competence. We derive three new testable mechanisms for why this may be for future research: For competence, we suggested that the more difficult to obtain an outcome (killing enemy > hitting enemy > swinging a sword), the more amplified feedback should be to signal the extent of the player’s displayed competence. This seems intuitive but to our knowledge, has not been suggested nor tested in the literature [26, 53, 55]. For effectance, we propose that differentiated feedback allows the observation and learning of multiple, differentiated action-outcome links. For curiosity, we suggest (in line with [40]) that differentiated success feedback makes reducible outcome uncertainty more salient.

For games user research and games HCI more widely, we see three upshots from our study. First, it provides robust evidence that curiosity is an important enjoyment driver for moment-to-moment game control, also outside ‘typical’ gameplay forms and genres like open-world exploration. Second, curiosity *not* enjoyment predicts moment-to-moment play continuation, counter to SDT-informed games HCI [67] and wider received wisdom [47, 70]. This suggests that expectational wanting may matter more and more directly for play behaviour than consummatory liking [7], in line with emergent work on the role of expectations in player (dis)engagement [5]. This invites future games HCI to explore curiosity as a potential alternative mediator to well-established concepts (like flow or need satisfaction), and separately study expected and realised positive player experiences (where common practice is to rely on self-report measures of the latter, such as the PXI). A third upshot is that *outcome binding* as part of sense of agency could be an important moderator of *positive* user experiences. It also provides an alternative, more low-level mechanism for SDT-informed games HCI why competence and autonomy satisfaction interact [57] with concrete and different interaction design implications. And it could help theoretically clarify ambiguity around the concept of autonomy in HCI itself [6].

5.6 Implications for Design

The unexpectedly negative effects of our A-SD-V condition underlines practitioner wisdom that fine-tuning juicy feedback is a delicate holistic design task [24, 26]. Our findings add some-in parts speculative-design guidance to existing work [26, 53]. First, we have good evidence that *success-dependent* amplified feedback enhances competence, curiosity, effectance, and enjoyment. Including such

granular success feedback is already a design recommendation derived from SDT [55], but hadn't been empirically validated before. Second, amplified feedback on mere action should be structured and proportioned to not diminish the relative value of success on challenging tasks. This aligns with existing guideline C4.A by Hicks and colleagues [26], but specifies it. Third, random variety in feedback is unlikely to enhance player experience and thus not recommended – this specifies Deterding's [17] design guidance for unexpected variety in juicy feedback, which did not preclude randomised variety. Fourth, when assessing the effectiveness of juicy feedback, game designers and games user researchers should not exclusively rely on self-reported enjoyment (e.g. via the IMI), but also assess actual voluntary playtime (and/or potentially, curiosity).

6 LIMITATIONS AND FUTURE WORK

If our occluding non-success-dependent feedback impeded agency, this could have suppressed true effectance effects. Future work should repeat our design, but with expert game design practitioners helping to ensure unambiguous and easily differentiable action-feedback links. Positively, we see rich future work in testing the potential moderating role of outcome binding and sense of agency on positive low-level user experience, including competence and effectance. Similarly, we identified further candidate detail design features worthy of further study: differentiable gradations of success feedback and truly random versus learnably generated feedback variation. Finally, on a theoretical level, we proposed sense of agency/outcome binding as a potential explanatory mechanism for SDT's proposed autonomy-competence coupling, as well as wanting and learned expectations as potential engagement drivers besides enjoyment. Both are exciting areas of future work.

In terms of generalisability, our study does not speak to all understandings of juiciness (such as enhancing any designer-intended game feel, tactility, aliveness, ...), nor claim external validity toward 'industry-level' polish. It was conducted in an action RPG and over a relatively short play duration, suggesting future work to test longer time scales and other genres, particularly turn-based and/or puzzle games where juicy feedback is less likely to directly interfere with real-time action. Our between-subjects design is more ecologically valid (few people play design variants of the same game in direct comparison), but may be less sensitive to real effects than within-person studies prevalent in prior work.

7 CONCLUSION

Juicy feedback is a widely used yet poorly specified game design concept for positive user experience afforded by immediate amplified feedback on player action. To systematically unpack when and why amplified feedback is engaging, we tested three candidate theories with linked mediators and design features: effectance and amplified feedback; competence and success-dependent feedback; and curiosity and varied feedback. For games HCI, our findings contribute strong direct support for SDT-based claims that success-dependent amplified feedback drives enjoyment, mediated by competence. They also suggest that sense of agency, specifically outcome binding ("My action caused this effect") may be an important moderator of competence, effectance, and positive user experience in moment-to-moment interaction. Outcome binding

may explain why good game feel is associated with unambiguous, coherent, highlighting and relevant feedback, and when and why amplified feedback becomes 'too much of a good thing' (as opposed to cognitive load, proposed in prior work). Players need to be able to learn to attribute screen events to their actions, which requires reliable (nonrandom) and non-mutually occluding action feedback. And they may feel more competence if feedback kind and volume communicates the difficulty of an attained game state. Our findings further show curiosity to be an under-appreciated factor of enjoyment and engagement even for low-level interaction, likely driven by reducible outcome uncertainty, and propose that not all forms of stochastic randomness produce such engaging uncertainty. They also suggest curious and/or expectation-driven wanting as a potential overlooked engagement driver in games next to enjoyment and proximate positive experiences like need satisfaction. Overall, we hope to have contributed to moving games HCI research on juicy feedback from exploratory effect searching to careful and systematic theory-testing.

ACKNOWLEDGMENTS

We thank Kieran Hicks for his involvement in testing and providing feedback on our study game platform. Additionally, we thank Lukas Marinovic for helping execute RITE testing. Sebastian Deterding declares that the present work has no relation with his current employment at Amazon UK Services Ltd.

REFERENCES

- [1] Marc Malmdorf Andersen, Julian Kiverstein, Mark Miller, and Andreas Roepstorff. 2023. Play in predictive minds: A cognitive theory of play. *Psychological Review* 130, 2 (March 2023), 462–479. <https://doi.org/10.1037/rev0000369>
- [2] Farid Anvari and Daniël Lakens. 2021. Using Anchor-Based Methods to Determine the Smallest Effect Size of Interest. *Journal of Experimental Social Psychology* 96 (Sept. 2021), 104159. <https://doi.org/10.1016/j.jesp.2021.104159>
- [3] Nick Ballou. 2023. A Manifesto for More Productive Psychological Games Research. *Games: Research and Practice* 1, 1 (2023), 1–26. <https://doi.org/10.1145/3582929>
- [4] Nick Ballou, Heiko Breitsohl, Dominic Kao, Kathrin Gerling, and Sebastian Deterding. 2021. Not Very Effective: Validity Issues of the Effectance in Games Scale. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play*. ACM, Virtual Event Austria, 55–60. <https://doi.org/10.1145/3450337.3483492>
- [5] Nick Ballou and Sebastian Deterding. 2023. 'I Just Wanted to Get it Over and Done With': A Grounded Theory of Psychological Need Frustration in Video Games. *Proceedings of the ACM on Human-Computer Interaction* 7, CHI PLAY (2023), 217–236.
- [6] Dan Bennett, Oussama Metatla, Anne Roudaut, and Elisa D Mekler. 2023. How does HCI Understand Human Agency and Autonomy?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [7] Kent C Berridge, Terry E Robinson, and J Wayne Aldridge. 2009. Dissecting components of reward: 'liking', 'wanting', and learning. *Current opinion in pharmacology* 9, 1 (2009), 65–73.
- [8] Brown, Lisa. 2016. The Nuance of Juice. <https://www.youtube.com/watch?v=qtgWBUIOjK4>
- [9] Christian Burgers, Allison Eden, Mélisande D. Van Engelenburg, and Sander Buningh. 2015. How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior* 48 (July 2015), 94–103. <https://doi.org/10.1016/j.chb.2015.01.038>
- [10] Richard W Byrne. 2013. Animal curiosity. *Current Biology* 23, 11 (2013), R469–R470.
- [11] Feinian Chen, Patrick J Curran, Kenneth A Bollen, James Kirby, and Pamela Paxton. 2008. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological methods & research* 36, 4 (2008), 462–494.
- [12] Junyi Chu and Laura E. Schulz. 2020. Play, Curiosity, and Cognition. *Annual Review of Developmental Psychology* 2, 1 (Dec. 2020), 317–343. <https://doi.org/10.1146/annurev-devpsych-070120-014806>

- [13] Andy Clark. 2018. A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences* 17, 3 (2018), 521–534.
- [14] Greg Costikyan. 2013. *Uncertainty in Games*. MIT Press, Cambridge, MA, London.
- [15] Edward L. Deci. 1980. *The psychology of self-determination*. Lexington Books.
- [16] Edward L. Deci and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press, New York.
- [17] Sebastian Deterding. 2015. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction* 30, 3-4 (2015), 294–335. <https://doi.org/10.1080/07370024.2014.993471>
- [18] Sebastian Deterding, Marc Malmdorf Andersen, Julian Kiverstein, and Mark Miller. 2022. Mastering uncertainty: A predictive processing account of enjoying uncertain success in video game play. *Frontiers in psychology* 13 (2022), 924953.
- [19] Carl F Fey, Tianyou Hu, and Andrew Delios. 2023. The Measurement and communication of effect sizes in management research. *Management and Organization Review* 19, 1 (2023), 176–197.
- [20] Lily FitzGibbon, Johnny King L Lau, and Kou Murayama. 2020. The seductive lure of curiosity: Information as a motivationally salient reward. *Current Opinion in Behavioral Sciences* 35 (2020), 21–27.
- [21] Karl J Friston, Marco Lin, Christopher D Frith, Giovanni Pezzulo, J Allan Hobson, and Sasha Ondobaka. 2017. Active inference, curiosity and insight. *Neural computation* 29, 10 (2017), 2633–2683.
- [22] Marcello A Gómez-Maureira and Isabelle Kniestedt. 2019. Exploring video games that invoke curiosity. *Entertainment Computing* 32 (2019), 100320.
- [23] Jacqueline Gottlieb and Pierre-yves Oudeyer. 2018. Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience* 19 (2018), 758–770. <https://doi.org/10.1038/s41583-018-0078-0> Publisher: Springer US.
- [24] grapefrukt. 2012. Juice it or lose it - a talk by Martin Jonasson & Petri Purho. <https://www.youtube.com/watch?v=Fy0aCDmgngx>
- [25] Kyle Gray, Kyle Gabler, Shalin Shodhan, and Matt Kucic. 2005. How to Prototype a Game in Under 7 Days. <https://www.gamedeveloper.com/game-platforms/how-to-prototype-a-game-in-under-7-days>
- [26] Kieran Hicks, Patrick Dickinson, Juicy Holopainen, and Kathrin Gerling. 2018. Good Game Feel: An Empirically Grounded Framework for Juicy Design. In *Proceedings of DiGRA 2018*. DiGRA.
- [27] Kieran Hicks, Kathrin Gerling, Patrick Dickinson, and Vero Vanden Abeele. 2019. Juicy Game Design: Understanding the Impact of Visual Embellishments on Player Experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, Barcelona Spain, 185–197. <https://doi.org/10.1145/3311350.3347171>
- [28] Kieran Hicks, Kathrin Gerling, Graham Richardson, Tom Pike, Oliver Burman, and Patrick Dickinson. 2019. Understanding the effects of gamification and juiciness on players. *IEEE Conference on Computational Intelligence and Games, CIG 2019-Augus* (2019). <https://doi.org/10.1109/CIG.2019.8848105>
- [29] Linus Holm, Gustaf Wadenholt, and Paul Schrater. 2019. Episodic curiosity for avoiding asteroids: Per-trial information gain for choice outcomes drive information seeking. *Scientific Reports* 9, 1 (Aug. 2019), 11265. <https://doi.org/10.1038/s41598-019-47671-x>
- [30] Mads Johansen and Michael Cook. 2021. Challenges in Generating Juice Effects for Automatically Designed Games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 17. 42–49. <https://doi.org/10.1609/aiide.v17i1.18889>
- [31] Mads Johansen, Martin Pichlmair, and Sebastian Risi. 2020. Squeezer - A Tool for Designing Juicy Effects. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. ACM, Virtual Event Canada, 282–286. <https://doi.org/10.1145/3383668.3419862>
- [32] Jeff Johnson. 2010. *Designing with the mind in mind: simple guide to understanding user interface design guidelines*. Morgan Kaufmann.
- [33] Jesper Juul. 2010. *A casual revolution: Reinventing video games and their players*. MIT press.
- [34] Jesper Juul and Jason Scott Begy. 2016. Good Feedback for bad Players? A preliminary Study of 'juicy' Interface feedback. In *Proceedings of first joint FDG/DiGRA Conference*. Dundee.
- [35] Dominic Kao. 2020. The effects of juiciness in an action RPG. *Entertainment Computing* 34 (May 2020), 100359. <https://doi.org/10.1016/j.entcom.2020.100359>
- [36] Celeste Kidd and Benjamin Y. Hayden. 2015. The Psychology and Neuroscience of Curiosity. *Neuron* 88, 3 (2015), 449–460. <https://doi.org/10.1016/j.neuron.2015.09.010> Publisher: Elsevier Inc. ISBN: 0066-4308 (Print).
- [37] Christoph Klimmt. 2006. *Computerspielen als Handlung: Dimensionen und Determinanten des Erlebens interaktiver Unterhaltungsangebote*. Vol. 2. Herbert von Halem Verlag.
- [38] Christoph Klimmt, Tilo Hartmann, and Andreas Frey. 2007. Effectance and control as determinants of video game enjoyment. *Cyberpsychology and Behavior* 10, 6 (2007), 845–847. <https://doi.org/10.1089/cpb.2007.9942>
- [39] Christoph Klimmt and Daniel Possler. 2021. A Synergistic Multiprocess Model of Video Game Entertainment. In *The Oxford Handbook of Entertainment Theory*, Peter Vorderer and Christoph Klimmt (Eds.). Oxford University Press, Oxford, 623–646.
- [40] Shringi Kumari, Sebastian Deterding, and Jonathan Freeman. 2019. The Role of Uncertainty in Moment-to-Moment Player Motivation: A Grounded Theory. In *CHI PLAY '19 Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM Press, New York. <https://doi.org/10.1145/3311350.3347148>
- [41] Daniël Lakens. 2022. Sample Size Justification. *Collabra: Psychology* 8, 1 (March 2022), 33267. <https://doi.org/10.1525/collabra.33267>
- [42] Eva Lenz, Sarah Diefenbach, and Marc Hassenzahl. 2014. Aesthetics of interaction: a literature synthesis. In *NordiCHI'14*. ACM Press, New York, 628–637. <http://dl.acm.org/citation.cfm?id=2639198>
- [43] Hannah Limerick, David Coyle, and James W Moore. 2014. The experience of agency in human-computer interactions: a review. *Frontiers in human neuroscience* 8 (2014), 643.
- [44] Jonas Löwgren. 2009. Toward an articulation of interaction aesthetics. *The New Review of Hypermedia and Multimedia* 15, 2 (2009), 129–146.
- [45] A G Marcello. 2018. Games that Make Curious: An Exploratory Survey into Digital Games that Invoke Curiosity, Vol. 11112. 76–89. <https://doi.org/10.1007/978-3-319-99426-0>
- [46] Manuela M. Marin. 2022. The Role of Collative Variables in Aesthetic Experiences. In *The Oxford Handbook of Empirical Aesthetics*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198824350.013.20>
- [47] Elisa D. Mekler, Julia Ayumi Bopp, Alexandre N. Tuch, and Klaus Opwis. 2014. A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. ACM Press, New York, New York, USA, 927–936. <https://doi.org/10.1145/2556288.2557078>
- [48] Elisa D. Mekler, Florian Brühlmann, Alexandre N. Tuch, and Klaus Opwis. 2017. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior* 71 (2017), 525–534. <https://doi.org/10.1016/j.chb.2015.08.048> Publisher: Elsevier Ltd ISBN: 0747-5632(Print).
- [49] James W Moore. 2016. What is the sense of agency and why does it matter? *Frontiers in psychology* 7 (2016), 1272.
- [50] Kou Murayama, Lily FitzGibbon, and Michiko Sakaki. 2019. Process account of curiosity and interest: A reward-learning perspective. *Educational Psychology Review* 31 (2019), 875–895.
- [51] Don Norman. 1988. *The design of everyday things*. Basic books.
- [52] Wei Peng, Jih-Hsuan Lin, Karin a. Pfeiffer, and Brian Winn. 2012. Need Satisfaction Supportive Game Features as Motivational Determinants: An Experimental Study of a Self-Determination Theory Guided Exergame. *Media Psychology* 15, 2 (2012), 175–196. <https://doi.org/10.1080/15213269.2012.673850> ISBN: 1521-3269.
- [53] Martin Pichlmair and Mads Johansen. 2022. Designing Game Feel: A Survey. *IEEE Transactions on Games* 14, 2 (June 2022), 138–152. <https://doi.org/10.1109/TG.2021.3072241>
- [54] Johnmarshall Reeve. 2014. *Understanding Motivation and Emotion* (6th ed.). John Wiley and Sons, Hoboken, NJ.
- [55] Scott Rigby and Richard M. Ryan. 2011. *Glued to Games: How Video Games Draw Us in and Hold Us Spellbound*. Praeger, Santa Barbara, Denver, Oxford.
- [56] Christian Roth, Christoph Klimmt, Ivar E Vermeulen, and Peter Vorderer. 2011. The experience of interactive storytelling: comparing "Fahrenheit" with "Façade". In *Entertainment Computing-ICEC 2011: 10th International Conference, ICEC 2011, Vancouver, Canada, October 5-8, 2011. Proceedings 10*. Springer, 13–21.
- [57] Richard M Ryan and Edward L Deci. 2017. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Press, New York.
- [58] Richard M Ryan, Valerie Mims, and Richard Koestner. 1983. Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of personality and Social Psychology* 45, 4 (1983), 736.
- [59] Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (2006), 344–360. <https://doi.org/10.1007/s11031-006-9051-8>
- [60] Cansu Sancaktar, Sebastian Blaes, and Georg Martius. 2022. Curious exploration via structured world models yields zero-shot object manipulation. *Advances in Neural Information Processing Systems* 35 (2022), 24170–24183.
- [61] Jesse Schell. 2008. *The Art of Game Design: A book of lenses*. CRC Press.
- [62] Tanay Singhal and Oliver Schneider. 2021. Juicy Haptic Design: Vibrotactile Embellishments Can Improve Player Experience in Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–11. <https://doi.org/10.1145/3411764.3445463>
- [63] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. *Cognitive Load Theory*. Springer New York, New York, NY. <https://doi.org/10.1007/978-1-4419-8126-4>
- [64] Steve Swink. 2009. *Game Feel: A Game Designer's Guide to Virtual Sensation*. Morgan Kaufman, Amsterdam et al.
- [65] Alexandra To, Safinah Ali, Kaufman Geoff, and Jessica Hammer. 2016. Integrating Curiosity and Uncertainty in Game Design. In *Proceedings of 1st International Joint Conference of DiGRA and FDG*. DiGRA, Dundee, 1–16.

- [66] Alexandra To, Jarrek Holmes, Elaine Fath, Eda Zhang, Geoff Kaufman, and Jessica Hammer. 2018. Modeling and Designing for Key Elements of Curiosity: Risking Failure, Valuing Questions. *Transactions of the Digital Games Research Association* 4, 2 (2018). <https://doi.org/10.26503/todigra.v4i2.92>
- [67] April Tyack and Elisa D Mekler. 2020. Self-determination theory in HCI games research: current uses and open questions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu, 1–22. <https://doi.org/10.1145/3313831.3376723>
- [68] Lieke L.F. Van Lieshout, Annelinde R.E. Vandenbroucke, Nils C.J. Müller, Roshan Cools, and Floris P. De Lange. 2018. Induction and Relief of Curiosity Elicit Parietal and Frontal Activity. *The Journal of Neuroscience* 38, 10 (March 2018), 2579–2588. <https://doi.org/10.1523/JNEUROSCI.2816-17.2018>
- [69] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and Validation of the Player Experience Inventory: A Scale to Measure Player Experiences at the Level of Functional and Psychosocial Consequences. *International Journal of Human-Computer Studies* 135 (March 2020), 102370. <https://doi.org/10.1016/j.ijhcs.2019.102370>
- [70] Peter Vorderer, Christoph Klimmt, and Ute Ritterfeld. 2004. Enjoyment: At the heart of media entertainment. *Communication theory* 14, 4 (2004), 388–408.
- [71] Robert W White. 1959. Motivation reconsidered: The concept of competence. *Psychological Review* 66, 5 (1959), 297–333.
- [72] Georgios N. Yannakakis and John Hallam. 2007. Modeling and Augmenting Game Entertainment Through Challenge and Curiosity. *International Journal on Artificial Intelligence Tools* 16 (2007), 981–999. <https://doi.org/10.1142/S0218213007003667> ISBN: 0218-2130.