

[> [jesper juul: text](#)]

---

# Fear of Failing? The Many Meanings of Difficulty in Video Games

Jesper Juul  
[www.jesperjuul.net](http://www.jesperjuul.net)

From Mark J. P. Wolf & Bernard Perron (eds.): *The Video Game Theory Reader 2*. New York: Routledge 2009. pp. 237-252.

## Winning isn't everything

It is quite simple: When you play a game, you want to win. Winning makes you happy, losing makes you unhappy. If this seems self-evident, there is nonetheless a contradictory viewpoint, according to which games should be "neither too easy nor too hard", implying that players also want *not* to win, at least part of the time. This is a contradiction I will try resolve in what follows. The question is:

- Question 1: What is the role of failure in video games?

The simplest theory of failure states that failing serves as a contrast to winning, that failure thereby makes winning all the more enjoyable. There is, however, much more to failure. The study of players discussed in this essay indicates that failure serves the deeper function of making players readjust their perception of a game. In effect, *failure adds content* by making the player see new nuances in a game. The study shows that players have quite elaborate theories of failure as a source of enjoyment in games.

Even so, given the negative connotations of failing, would a game be better received if players did not feel responsible for failing, but rather blamed failures on the game or on bad luck? This is the second question:

- Question 2: Do players *prefer* games where they do *not* feel responsible for failing?

This study strongly indicates that this is *not* the case. Players clearly prefer feeling responsible for failing in a game; not feeling responsible is tied to a negative perception of a game.

In effect, this sharpens the contradiction between players as wanting to win and players wanting games to be challenging: failing, and feeling responsible for failing, makes players enjoy a game *more*, not less. Closer examination reveals that the apparent contradiction originates from two separate perspectives on games: a *goal-oriented* perspective wherein players want to win, and an *aesthetic perspective* wherein players prefer games with the right amount of challenge and variation. Nevertheless, these two perspectives still present opposing considerations – the goal-oriented perspective suggests that games should be as easy as possible; the aesthetic perspective suggests that games should not be too easy.

To examine this, I will look take a closer look at the role of failure and punishment. I am writing here about single-player games.[1](#)

## Failure and Punishment

Failure can be described as being unsuccessful at some task in a game, and punishment is what happens to the player as a result. We can distinguish between different types of punishment for player failure:[2](#)

- **Energy punishment:** Loss of energy, bringing the player closer to life punishment.

- **Life punishment:** Loss of a *life* (or “retry”), bringing the player closer to game termination.
- Game termination punishment: Game over.
- **Setback punishment:** Having to replay part of the game; losing abilities.

Losing energy brings the player closer to losing a life, and losing a life often leads to some type of setback. In this perspective, all failures eventually translate into setbacks, and the player's use of time and energy is the most fundamental currency of games.

Whereas early video games in the arcade, on the home console, or for personal computers, tended to force the player to replay the entire game after failing, many home games from the mid-1980s and on became much more lenient by dispersing save points, allowing the player to save the game at will, or letting the player restart at the latest level played even after game over. As a recent example of this design principle, after reaching game over in *Super Mario Galaxy* (Nintendo EAD Tokyo, 2007), the player loses coins and collectables, but not overall progress in the game.

In the new area of downloadable casual games,<sup>3</sup> there is a movement from life punishment to energy punishment, with many games featuring energy bars, timers, or other types of soft evaluations of player performance as with the timer in *Big City Adventure: San Francisco* (Jolly Bear Games 2007) (see Figure 1).



Figure 1 : *Big City Adventure: San Francisco* - a timer gradually runs out. (Jolly Bear Games 2007)

The psychological *attribution theory* provides a framework for examining different types of failure and punishment in games. According to attribution theory, for any event, people tend to attribute that event to certain causes. Harold K. Kelley distinguishes between three types of attributions that people can make in an event involving a person and an entity:

- **Person:** The event was caused by personal traits, such as skill or disposition.
- **Entity:** The event was caused by characteristics of the entity.
- **Circumstances:** The event was based on transient causes such as luck, chance, or an extraordinary effort from the person. F. Försterling, *Attribution: An Introduction to Theories, Research and Applications* (London: Psychology Press, 2001), 46-47, hereafter cited in the text as *Försterling*.

In the case of receiving a low grade for a school test, a person may decide that this was due to

(a) person – personal disposition such as lack of skill, (b) entity – an unfair test, or (c) circumstance – having slept badly, having not studied enough. This maps quite well to many common exclamations in video gaming: a player who loses a game can claim that “I am terrible at video games”, “This is an unfair game”, or “I will win next time”.

During the research for this essay, I developed the hypothesis that energy punishment is being more widely used because it makes the cause of failure less obvious: If the game is over due to a single, identifiable mistake, it is straightforward for the player to attribute failure to his or her own performance or skill (circumstance or person), but if the game is over due to an accumulation of small mistakes, the player is less likely to feel responsible for failing, and the player should be less likely to experience failing as an emotionally negative event. This is the second question mentioned in the introduction: do players *prefer* feeling less responsible for failing?

## Video Game Theory Through Game Prototypes

To elaborate this discussion, a game prototype study was conducted. This is not without precedent. In a study made 25 years ago, Thomas W. Malone explored the question “Why are computer games so captivating?” by creating a number of game prototypes with the same core game, but with different features (music, scorekeeping, fantasy, types of feedback) (Malone 1982). In order to explore the attraction of the variations of the game, he let some children play these prototypes and examined how long each prototype was able to keep the attention of young players. From this, he deduced a number of guidelines for developing games and interfaces.

Following Malone, the questions in this essay can be approached as empirical questions – *What do players prefer?* They can, however, also be approached as aesthetic questions – *What is a good game?* These are two historically separate approaches that I nevertheless believe can inform each other in the following.

In collaboration with the game company Gamelab, I developed a game prototype specifically designed to gather data on how players perceive failure. The custom game could be described as a combination of *Pac-Man* (Namco 1980) and *Snake* (Originally Gremlin 1977): using the mouse, the player controls a snake that grows as the player collects pills; the player must avoid opponents; and a special power pill allows the player to attack opponents for a short while (see Figure 2).

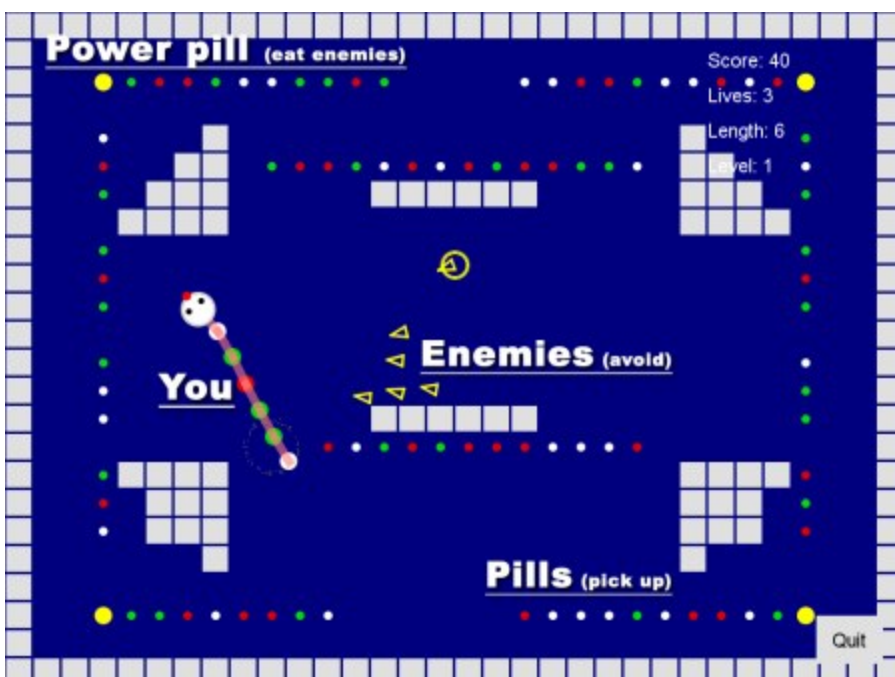


Figure 2 : Game prototype for the test.

The game was designed with two game modes, an *energy punishment* mode where the player would lose a tail part when hit by opponents, and a *life punishment* mode where the player could make only a single mistake before losing a life. In both games, the player has three lives, and the game consists of four levels. We attempted to balance the two games so that they were equally hard (as measured in the number of levels that players would complete). Another reason for developing a new game was that this would give insight to the players' initial experience of learning a new game, and be less a reflection of their previous experience with that game.

## First Test, Offline

A preliminary test was conducted offline. Five males and four females from Gamelab's tester base participated. All participants had some experience with and interest in games, and came to the Gamelab offices (see Appendix 1 at the end of this essay for a description of the test procedure). Players were asked how they would rate the game, had they found it on the Internet. The rating scale went from 1 to 10, with 10 being the best rating. Additionally, players were asked open questions about their views on failure in games.

Contrary to expectations, this small sample gave no indication that players preferred the energy punishment version of the game. On the other hand, there were indications that the players' ratings were closely tied to their performance in the game, such that a player performing badly would dislike the game, a player performing fairly well would like the game, but a player performing very well would *also dislike the game*. Given the interesting implications of this result, it was decided to focus on only one version of the game (energy punishment), and run a new test online with a bigger sample.

## Second Test, Online

85 players were recruited online<sup>4</sup> and asked to play the game and answer a questionnaire (see Appendix 2 for a description of the test procedure). The players recruited were overwhelmingly male (73 out of 85), and the majority had a game console in their home (also 73 out of 85). Players were generally avid game players (see Figure 3).

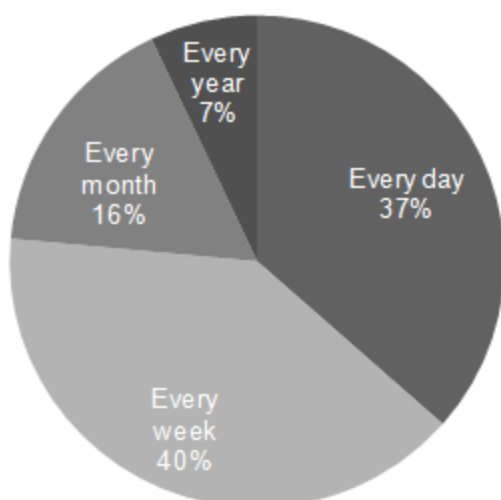


Figure 3 : Game-playing frequency.

## Game rating vs. performance

Based on automated registration of player performances, player responses were placed into three categories, from a bad performance to a good performance:

- Players that did not complete the game.
- Players that completed the game, losing some lives.

- Players that completed the game without losing any lives.

By comparing the average game ratings with the performance of the players (Figure 4), we can see an indication that *winning isn't everything*: the most positive players were the ones that *failed* some, and then completed the game. Completing the game without failing was followed by a *lower* rating of the game (the statistical significance was the slightly weak  $p < 0.06$  for all three categories of player performance combined).

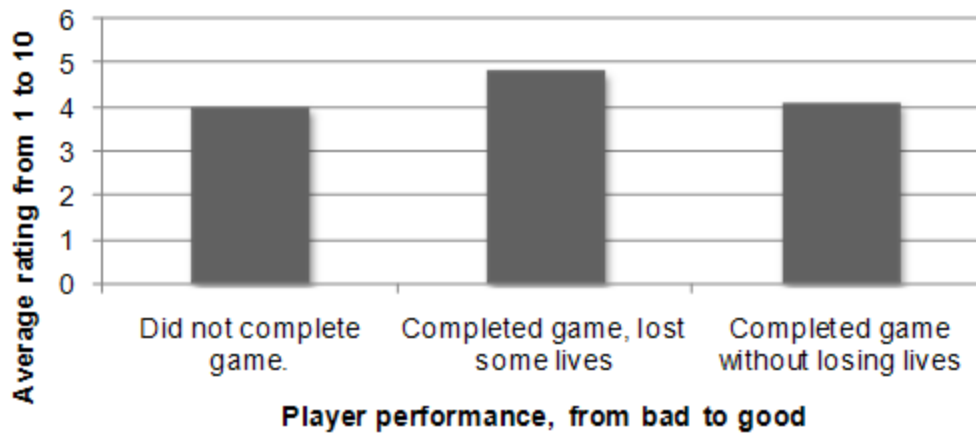


Figure 4 : Player rating of game as function of performance.

This runs counter to the simple idea that players enjoy a game more the better they do, but it vindicates the game design imperative that a game must be neither too hard nor too easy as argued by, for example, Fullerton, Swain, and Hoffman (2004, 249). This returns us to the second question, of whether feeling responsible for failing in a game will make players like the game less. In the test, players were asked why they failed or succeeded. Categories were based on attribution theory, but expanded into smaller subcategories:

- **Person** was split into “I am bad at this kind of game” and “I am bad at games in general” to capture difference between general player skills and player knowledge of specific genres.
- **Entity** was asked as “The game was too hard”.
- **Circumstance** was split into “I was unlucky” and “I made a mistake” in order to distinguish between the experience of losing due to chance and losing due to a strategic mistake.

As can be seen in figure 5 and figure 6, players were slightly more likely to report being responsible for success (“figured out how to play right”) than being responsible for failure (“made a mistake”). This is well-known phenomenon called *attribution asymmetry*, whereby individuals are more likely to attribute success to personal factors, and failure to external factors (Försterling, 87-91).

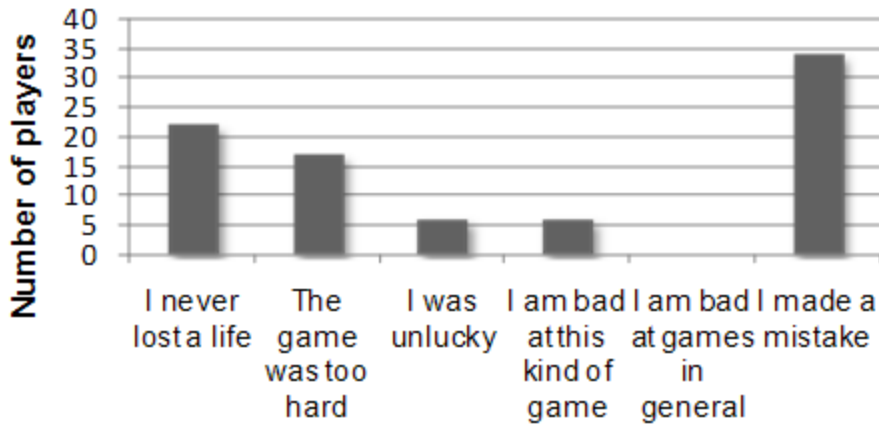


Figure 5 : Player attribution of failure

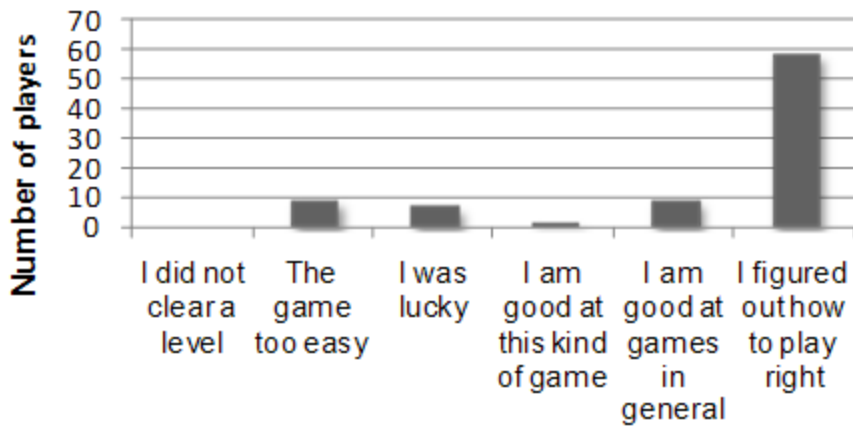


Figure 6 : Player attribution of success.

Do players prefer games where they do not feel responsible for failing? This seems not to be the case. On the contrary, even though players presumably on some level dislike being personally responsible for failing, the feeling of being responsible for failing was nevertheless tied to a *positive* rating of the game (see Figure 7).

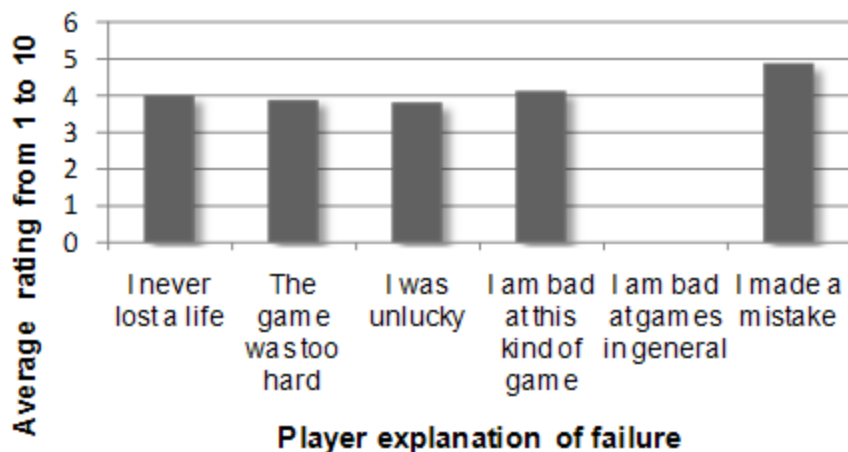


Figure 7: Rating as function of failure attribution.

Since players who never lost a life are not relevant, and too few players answered “I was unlucky” or “I am bad at this kind of game” for the results to be meaningful, we can see how players who answered “The game was too hard” rated the game compared to those who answered “I made a mistake”. In this case, there is a clearer significance of  $p < 0.016$ . In effect,

this answers the second question of this paper – players *prefer feeling responsible* for their own failure. Or at least the negative emotions from failing are more than cancelled out by other factors. This result is parallel to a study of players playing the bowling mini-game in *Super Monkey Ball 2* (Amusement Vision, Ltd. 2002), in which players exhibited positive reactions when falling off the edge of the playing field, but negative reactions of watching the replay of the same event (Ravaja et al. 2005). Although players do not want to fail, they may nevertheless enjoy it when feeling responsible for it. [5](#)

## Players Reactions When Not Failing

Do players have theories of the function of failure, and in that case, how do they frame them? To find out, players were asked if they had ever experienced a game that was too easy, and “*How do you know if a game is too easy?*” Answers were sorted into four categories based on their primary content, listed here with example answers and percentages:

Answer type	Examples
Too easy as lack of a challenge. (36%)	<p>“not challenging enough”</p> <p>“boring... doesn't provide further challenges”</p> <p>“If I don't feel challenged. Of course that's a pretty predictable answer, but it's hard to put it any other way.”</p> <p>“I get bored”</p>
Too easy as not failing. (6%)	<p>“When you never die. And beat it in a day”.</p> <p>“It doesn't seem to challenge me – I never lose”</p>
Too easy as not being measured on performance. (5%).	<p>“I can do things I know are “wrong” and don't get punished.”</p> <p>“A game is too easy when you are progressing through the game automatically no matter how good you are playing.”</p>
Too easy as not having to rethink strategy. (27%)	<p>“When I know exactly what to do and I can do it optimizing the result without (big) effort”</p> <p>“No challenge, go through motions to complete it without any thought”</p> <p>“If the challenge and thought require to complete it objectives become second nature quickly or there is no need for such contemplation”</p> <p>“If the method for solving it is obvious and never fails.”</p>

The first response type, lack of challenge, is somewhat tautological. Response 4 gives room for more interpretation: if a game being too easy is experienced as the game being shallow and uninteresting, it means that the role of failure is much more than a contrast to winning – failure pushes the player into reconsidering strategy, and failure thereby subjectively *adds content* to the game. The game appears deeper when the player fails; failure makes the game more strategic.

It would be nice to know if the results from this experiment map to players of published commercial games. In a discussion of the initial disappointing reception of the game *Shopmania* (Gamelab 2006), Catherine Herdlick and Eric Zimmerman discuss how much of the criticism of the game came from the fact that it was perceived as too easy:

In the original version of *Shopmania*, we approached the first several levels of the game as a gradual tutorial that introduced the player to the basic game elements and the core gameplay. This approach was based on the generally held casual game wisdom that downloadable games should be very easy to play, and that the frustration of losing a level should be minimized. However, the problem with going too far in this direction is that the game ends up feeling like interactive muzak: you can play forever and not really lose, and the essential tension and challenge of a good game are lost. From our analysis, players were telling us that the first seven

or eight levels felt like a tutorial. By the third or fourth level, we had playtesters exclaiming out loud, “I get this game. Can I skip the tutorial” Catherine Herdlick and Eric Zimmerman, “Redesigning Shopmania: A Design Process Case Study,” *IGDA Casual Games Quarterly* 2, no. 1 (2006), [http://www.igda.org/casual/quarterly/2\\_1/index.php?id=6](http://www.igda.org/casual/quarterly/2_1/index.php?id=6)

One of the negative comments on *Shopmania* was about having seen the whole game too early:

“After 20 minutes, felt like I saw the whole game...” (*Redesigning*)

The “see” here probably does not refer directly to concrete graphics or level layouts, as much as it ties into some of the player comments in my experiment: The players complain about the game not pressuring them, not threatening with failure. Again, while players may dislike failure, not failing can be as bad as never succeeding.

## Flow: The Standard Theory of Failure and Challenge

The standard psychological explanation for game failure and challenge is Mihaly Csikszentmihalyi’s theory of *Flow* (see Figure 8), according to which the challenge of a given activity forms a narrow channel in which the player is in the attractive *flow* state (Csikszentmihalyi 1990).

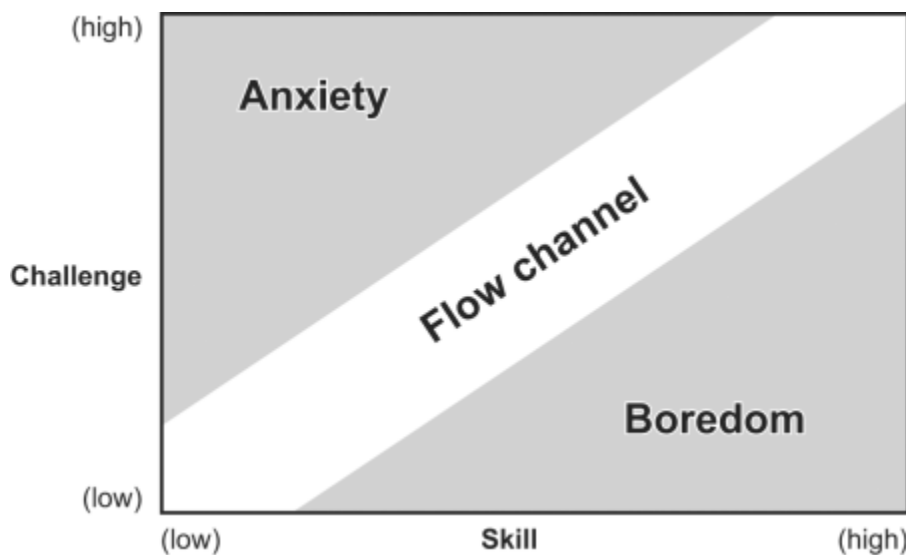


Figure 8: The flow channel. (Csikszentmihalyi 1990, 74)

While flow theory does suggest that the player may oscillate between anxiety and boredom, it poses the banal problem that the standard *illustration* suggests a smooth increase in difficulty over time. Noah Falstein (Falstein 2005) has refined this to say that game difficulty should vary in waves – sometimes the game should be a little easy, sometimes a little hard, and that irregularity leads to enjoyment, as illustrated in Figure 9. An irregular increase in difficulty makes the player more likely to experience both failure and successes.



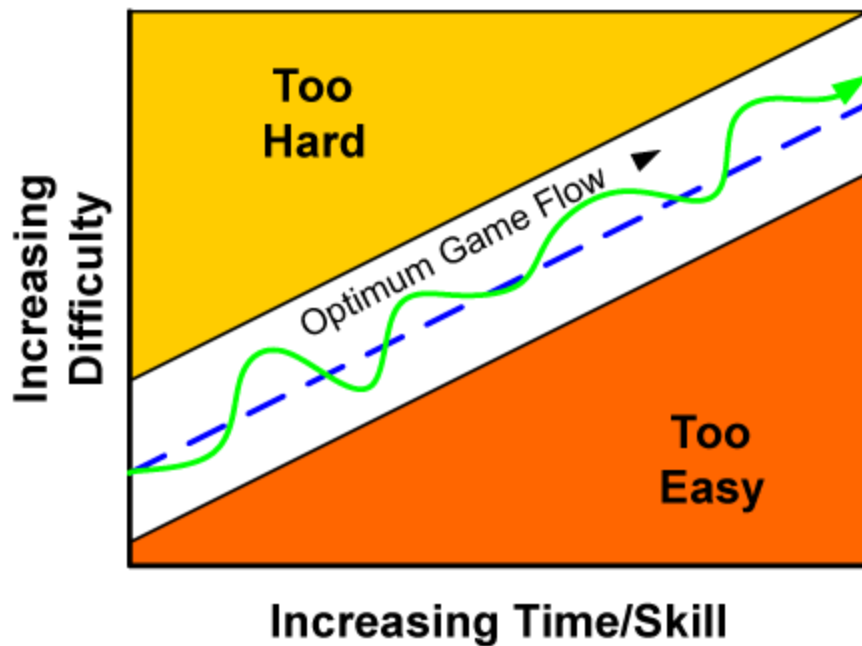


Figure 9: A better flow. (© 2004 Noah Falstein)

## Conclusions: The Contradictory Desires of Players

I initially discussed a contradiction between the observation that players want to win and the observation that players prefer games where they lose some, then win some. This leaves us with several opposing considerations indicating that games should be both easier and harder than they are:

1. The player does not want to fail (makes player sad, feels inadequate).
2. Failing makes the player reconsider his/her strategy (which makes the game more interesting).
3. Winning provides gratification.
4. Winning without failing leads to dissatisfaction.

Points 1 and 3 suggest that games should be very easy, whereas points 2 and 4 suggest that games should not be *too* easy. The actual relationship of game design and game playing is probably not as antagonistic as this seems. A more productive view is that games derive their interest from the interaction between these different considerations, and that the apparent contradiction comes from the fact that games can be viewed from two distinct frames of reference (see Figure 10). Playing a game entails (1) a goal-orientation as part of the activity, but a player also has (2) an outside view of the game that entails an aesthetic evaluation of game challenge. This is the source of the contradiction discussed in the introduction, between players wanting to win, and players wanting not just to win.

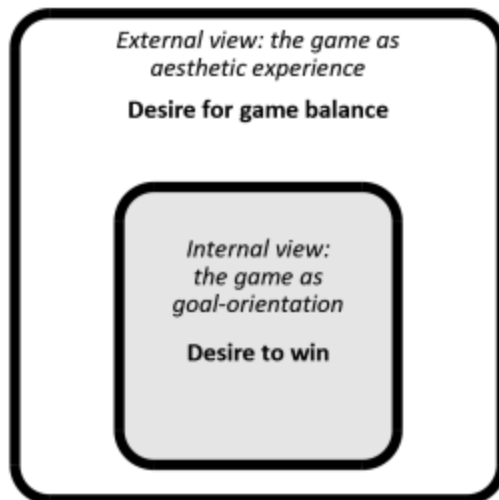


Figure 10: Internal and external views of a game.

The second question was whether players would prefer *not* feeling responsible for failing, and whether the success of casual games consequently could be attributed to the fact that they tend to have energy punishment rather than life punishment, making failure seem less of a direct consequence of player actions. This idea seems to be largely disproved – player appreciation of the game was tied positively to feeling responsible for failure. This suggests that I had been focusing on the wrong part of the punishment system, and that the attraction of casual games is better explained as sparing use of *setback punishment*: failing in casual games is rarely tied to any substantial setback, and never to having to mechanically replay a game sequence.<sup>6</sup> Players still feel responsible for failing, but they are less likely to feel stuck in the game, being forced to replay a part of the game.

Finally, this research points to another layer of complexity in player psychology. That failure and difficulty is important to the enjoyment of games correlates well with Michael J. Apter's *reversal theory*, according to which people seek low arousal in normal goal-directed activities such as work, but high arousal, and hence challenge and danger, in activities performed for their intrinsic enjoyment, such as games (Kerr and Apter 1991, 17). This yields an extra complication in relation to the game *Shopmania* discussed previously: if the role of failure is to force players to discover new strategies in a game, why is this even necessary? Given that players enjoy a challenge, why do players not simply challenge themselves by finding new ways to play the game? Game designer David Jaffe goes as far as asserting that players are basically lazy and "WILL NOT use ANY mechanic they do not need to use. They will take the path of least resistance to get from A TO B" (Jaffe 2007)..

The conclusion must still be that players want to fail as well as win, but that players of the *single-player* games discussed here do *not* seek out additional challenge or depth if they do not have to. Perhaps single player games are perceived as designed experiences that players expect to be correctly balanced without having to seek additional challenges themselves?

By contrast, although the focus here has been on single player games, Jonas Heide Smith has documented how players of multiplayer games frequently handicap themselves to create an even playing field, effectively opening themselves to failure (Smith 2006, 217-227). Multiplayer games and more open sandbox games seem to encourage players to undertake more challenge-seeking behavior.

The study raises a number of additional questions, but I believe the following are the most obvious ones to explore further:

- Is the relation between game rating and performance also consistent if the game is made easier or harder?
- How do players perceive difficulty in games without time pressure or failure states, such as "endless" mode in *Bejeweled 2* (PopCap Games 2004) or *Sudoku*?
- In game development experience, it is certain that small changes to game designs *do*

matter to players. To what extent can individual elements of a game design be isolated?

- To what extent can we extrapolate from one game to all games?
- Will the results of the test be different with a more “casual” audience?

I have argued that failure is central to player enjoyment of games. This is not that surprising, given conventional wisdom that a game should be balanced to match the skills of players. However, it is notable that failure is more than a contrast to winning – rather failure is central to the experience of *depth* in a game, to the experience of improving skills. The study supports the idea that *growth*, the experience of learning, of adjusting strategies, of trying something new, is a core attraction of video games.<sup>7</sup> Hence the desire for game balance, losing some, winning some, is also a desire for variation in the challenge and difficulty of the game. *Failure adds content*.

If the classic tenet of storytelling is Aristotle’s that a story should have a beginning, a middle, and an end, the core tenet of games must be this: *a game should be neither too easy nor too hard*. This is more than the simple truism it sounds like. It reveals much deeper and more complicated facts about games, and players.

## Acknowledgements

This research was done in collaboration with Gamelab in New York City, who provided facilities, discussion, feedback, and playtesting. Thanks to T.L. Taylor, Jonas Heide Smith, Eric Zimmerman, Nick Fortugno, Chris Bateman, and Matthew Weise for comments. Thanks to Svend Juul for statistical expertise.

## Appendix 1: Offline Test Procedure

- Participants were tested one at a time, and did not see or talk to other participants. Participants were informed that “We are working on a game, and we would like to hear your input. This is not a test of your skill; we would simply like to know what you think about the game.”
- Each player was asked to play the game until the game was over. It was noted on what levels players lost lives.
- Each player was asked “Why did you fail?” and “Why did you complete the level?” The explanations were coded as being either due to *ability* (personal factor), *performance* (circumstance), or the *game* (entity).
- After one game had been played, the player was interviewed.
- Each player was asked to rate the game as follows: “If this was a game you found on the web, how would you rate it on a scale from 1 to 10, with 1 being the worst and 10 being the best?”<sup>8</sup>
- Each player was asked to explain if he or she had ever played a game that was too easy.
- Each player was asked how he or she could tell if a game is too easy.
- Participants were not paid, but as game testing is often described as a way of entering the game industry, testers may have strong motivation for pleasing the company. This affects the confidence in the absolute judgments of the players, but since the testers’ interest in pleasing the company will be statistically uniform, the data can be used relatively in correlation with other data from the test.

## Appendix 2: Online Test Procedure

- Players were recruited via the author’s blog.
- Players were told that “This is *not* a test of your skills, but a test of how you feel about playing a little game experiment”; players were not aware that the test concerned failure.
- Players were directed to a page with instructions, as can be seen at <http://www.jesperjuul.net/test/rpt2/>.
- Players were directed to the game. The game consisted of four levels. The player had

three lives.

- When a player reached game over, either by completing all four levels or by losing all three lives, the player was directed to an online questionnaire. In the questionnaire, the player was asked to rate the game as follows: "Say you found this game on the Internet. On a scale from 1 to 10, with 1 being the worst game ever, and 10 being the best game ever, how would you rate this game?"
- Only players who completed the entire questionnaire were included.

## References

- Amusement Vision, Ltd. 2002. *Super Monkey Ball 2*. SEGA Corporation (GameCube).
- Csikszentmihalyi, Mihaly. 1990. *Flow: The Psychology of Optimal Experience*. New York: Harper & Row.
- Falstein, Noah. 2005. "Understanding Fun—The Theory of Natural Funativity". In *Introduction to Game Development*, ed. Steve Rabin, 71-98. Boston: Charles River Media.
- Försterling, F. 2001. *Attribution: An Introduction to Theories, Research and Applications*. London: Psychology Press.
- Fullerton, Tracy, Chris Swain, and Steven Hoffman. 2004. *Game Design Workshop: Designing, Prototyping, and Playtesting Games*. San Francisco: CMP Books.
- Gamelab. 2006. *Shopmania*. (Windows).
- Gremlin. 1977. *Hustle*. (Arcade).
- Herdlick, Catherine, and Eric Zimmerman. 2006. "Redesigning Shopmania: A Design Process Case Study". *IGDA Casual Games Quarterly* 2, no. 1. [http://www.igda.org/casual/quarterly/2\\_1/index.php?id=6](http://www.igda.org/casual/quarterly/2_1/index.php?id=6).
- Jaffe, David. 2007. Aaaaaaaaaannnnnnndddddd Scene! *Jaffe's Game Design*. November 25. [http://criminalcrackdown.blogspot.com/2007\\_11\\_25\\_archive.html](http://criminalcrackdown.blogspot.com/2007_11_25_archive.html).
- Jolly Bear Games. 2007. *Big City Adventure: San Francisco*. (Windows).
- Kerr, J. H, and Michael J Apter. 1991. *Adult Play: A Reversal Theory Approach*. Amsterdam: Swets & Zeitlinger.
- Lazzaro, Nicole. 2004. "Why We Play Games: Four Keys to More Emotion in Player" Experiences. Paper presented at the *Game Developers Conference*, San José, 2004.
- Malone, Thomas W. 1982. "Heuristics for designing enjoyable user interfaces: Lessons from computer games". In *Proceedings of the 1982 conference on Human factors in computing systems*, 63-68. Gaithersburg, Maryland, United States: ACM.
- Namco. 1980. *Pac-Man*. Namco (Arcade).
- PopCap Games. 2004. *Bejeweled 2 Deluxe*. (Windows).
- Ravaja, Niklas, Timo Saari, Jari Laami, et al. 2005. "The Psychophysiology of Video Gaming: Phasic Emotional Responses to Game Events". In *DiGRA Conference Proceedings*. <http://www.digra.org/dl/db/06278.36196.pdf>.
- Smith, Jonas Heide. 2006. *Plans and Purposes: How Video Games Shape Player Behavior*. PhD dissertation, IT University of Copenhagen.

## Notes

1. For studies of players in multiplayer settings, see (Smith 2006) and Lazzaro (2004).
2. Not all failure is punished in games - many smaller types of failure go unpunished, such as bumping into a wall.
3. Casual games are understood here as downloadable games that the player can play freely for typically 60 minutes, after which the game must be purchased to continue playing.
4. Via the *Ludologist* blog, [www.jesperjuul.net/ludologist](http://www.jesperjuul.net/ludologist).

5. The conclusions from the *Super Monkey Ball 2* study may not map to questions discussed in this essay, as *Super Monkey Ball 2* has a rewarding audiovisual feedback when the player fails compared to the more basic representation in the game prototype used here.
6. This is also due to the fact that casual games tend to contain much randomness, making every replay of a single level is a bit different from the previous.
7. This is close to what Nicole Lazzaro calls “hard fun” (2004).
8. Since there is no universal scale for rating games, little can be deduced from the individual rating, but ratings can be used comparatively to examine player perceptions of game quality.