Using Dark Gameplay Design Patterns for Game Ethics Education

Anonymous First Author
Institutional Affiliation
Address line 1
Address line 2
Telephone
firstauthor@institution.com

ABSTRACT
It is important for students in game programs to understand the ethical issues and responsibilities they will assume as they become professional game creators. What concepts might game educators use to support their students in learning game ethics? We argue that dark game design patterns might be one such concept. This is because they can scaffold ethical speculation around existing game features and technologies, assist in the exploration of potential ethical issues and concerns in near-now game features and technologies, and support reflection regarding existing game features perceived as “troublesome” in order to articulate whether or not they should be considered as dark. We provide a examples of how this could be implemented in the classroom.

Keywords
Dark patterns, dark gameplay design patterns, game ethics education, pedagogy

INTRODUCTION
There is a responsibility faced by educators in production-oriented game programs: by the time students graduate they should (hopefully) be able to identify and reason about the ethical issues and responsibilities they assume as game creators. These responsibilities are related to their games’ content and gameplay and the uses (and misuses) of the technologies developed for and during game development. For these reasons, ethics-related learning objectives have long been seen as important in post-secondary games education (IGDA, 2008).

However, there is not much work on how to best teach game ethics. For example, which concepts might be productive? We know that gameplay design patterns have been used productively to teach game design (Holopainen et al., 2007), therefore dark game design patterns might similarly be useful for teaching game ethics. Dark game design patterns are commonly reoccurring game design elements that are “used intentionally by a game creator to cause negative experiences for players which are against their best interests and are likely to happen without their consent” (Zagal et al., 2013). As a concept they have been used in games industry primarily as a way to identify problematic (i.e. unethical) practices that should be avoided (e.g. Sanusi, 2017). Much effort has looked into identifying patterns (e.g. Dahlan & Susanty, 2020), better understanding how/when they operate (e.g. Lewis, 2014), how their use should be regulated (or not) (e.g. Goodstein, 2021), and how to help players develop “resistance” to them (e.g. Bongard-Blanchy et al., 2021). New dark patterns have even been proposed to critique existing practices in game/software development (Linehan et al., 2015).
Dark patterns in game design are widely used in a variety of ways as an object of study across a variety of university coursework and workshops (e.g. Dunlap, 2019; Sparrow, 2020; Hassan, 2021; Levy, 2021; Zagal, 2021; Prax, 2022; Hammer et al., 2022; Chen, 2023). In this article we propose a few ways they can be used for supporting student’s development as future ethical professionals in the game industry. Specifically, we believe the dark patterns can be used to help engage in forward thinking ethical reflection (e.g. consider the potential ways that their games (and associated technologies) could be misused or result in unintended consequences) and a scaffolding for articulating the issues that certain features or design elements may have – including those that already exist in publicly available games. In a sense, what we are arguing is that the final determination of whether a pattern is dark or not (i.e. telling students that X is a dark pattern) is likely less useful pedagogically than having students engage in the reflective process of identifying and/or producing their own patterns (existing, plausible, and speculative) and arguing how and why they are dark (or not).

<table>
<thead>
<tr>
<th>Considerations</th>
<th>Guiding Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer Intent</td>
<td>• How likely is it that the game’s creators intended for the use of the pattern to result in a negative experience for the player?</td>
</tr>
</tbody>
</table>
| Negative Experience | • What kind of negative experience might use of the pattern result in? (e.g. pain, frustration, anger, fear, embarrassment, grief, sense of loss, etc.)
  • How likely is it that the negative experience will in fact occur? |
| Player’s Best Interests | • In what way(s) is the use of the pattern in the best interest of the player? (e.g. greater likelihood of game experience desired by the player?)
  • If the player was aware of the reasons for implementing the pattern how likely is it that they would they agree to its use? |
| Player’s Consent    | • How reasonable is it to assume that they player fully understands what kind of play experience use of this pattern will likely lead to? (i.e. Is the pattern well known by players? Is its use discussed in player communities? |

Table 1: Components from definition of dark game design patterns and guiding questions for students

The definition of dark game design pattern provided by Zagal et al. (2013) provides a framework for student’s reflective processes. The definition consists of four considerations: developer intent, negative experience, player’s best interests, and player’s consent. From these considerations we can derive some guiding questions (see Table 1 for some examples). The answers to these questions (given a particular pattern) provide the framework of an argument for (or against) the “darkness” of a particular pattern. By answering these kinds of questions students can also begin to tease out subtleties and nuances that might be relevant for a particular pattern. In the following sections we illustrate a few techniques for using dark design patterns in the classroom.

**SPECULATIVE DARK(?) PATTERN DEVELOPMENT**

Design patterns are not generally developed, rather they are discovered from existing structures in games (Björk & Holopainen, 2005; Thomas et al., 2020). We can turn this around, however, and ask students to elaborate their own patterns from their experiences playing games, news sources, and more. In this way students can reflect on design elements before they become patterns in the “commonly reoccurring” sense.

As an assignment, students could be asked to investigate videogame-related news articles, press releases, and so on for announced features or developments that might be problematic. Next, students would articulate design patterns of those features and

---

1 Additional questions can and should be developed of course. Furthermore Zagal et. al. (2013) also provide guiding questions for specific kinds of dark patterns in the original article that are not included here for reasons of space.
then examine their patterns using the lens of dark gameplay patterns. Their analysis would obviously be speculative – their patterns would be based on limited knowledge of any real-world implementation and they would have few (if any) actual game examples to rely on. Therefore, their examination would pose a multitude of new questions – for example, if the pattern worked in a certain way, it might be problematic, but not in others, and so on.

For example, in 2012 an article published on *Ars Technica* described how Rockstar games announced they would implement a novel way of dealing with cheaters in their game *Max Payne 3*: instead of banning cheaters, they would be “quarantined” in a cheaters-only pool, only being able to play against other people similarly found to have been cheating (Orland, 2012). From this article (and Rockstar Games’ official announcement), students could propose the following pattern:

**Cheater’s Pool** - *Players caught cheating in a PvP game are silently banished to a “cheater’s pool” where they can only play against other people in the pool. Players are not notified they are sent to the cheater’s pool (based on Rockstar Games, 2012).*

In their analysis of this pattern students would have to comment on, and connect, the developer’s intent to a negative outcome, discuss what the player’s best interests might be, and what player consent could look like. For example, the above pattern was (most likely) intentionally designed and implemented as a way to separate players caught cheating from those not. This is done in order to “protect” the quality of the play experience of non-cheaters (i.e. playing against cheaters is a bad play experience). However, what sort of experience does it create for those caught cheating? Is it a dark pattern because secretly forcing cheaters to play against each other results in a worse play experience for cheaters? When a cheater plays against non-cheaters, they do so presumably because it is a positive experience for them. Is this also the case when everyone is cheating? If the play experience is a good one (cheaters enjoy when they play against cheaters), then it is hard to argue that this pattern is dark (see commentary in Orland, 2012). However, if it is not – then this pattern purposefully leads to a negative experience for some players (the cheaters) that is against their best interests: is it dark? Well, the definition of dark pattern does not articulate how to consider what should happen when distinct groups of players are affected different. How should that be weighed? Does the “darkness” of the pattern depend on how many cheaters there are in the game because that is where the negative experience occurs? Or, should we ignore the negative experience for the cheaters by virtue of them being cheaters? If so, how can we morally justify that? Furthermore, should we account for the possibility that the cheating-detection system may have errors and falsely flag some players as cheaters? Also, this pattern operates secretly – without players’ direct knowledge or consent. Arguably it is in the best interest of non-cheaters, but not so in the case of cheaters. So, how do we handle consent in this case? And so on, and so forth.

The analysis above is not meant as an exhaustive and rigorous analysis – rather it is intended to show how (1) identifying a pattern, (2) articulating it, and then (3) discussing how it might be dark or not (on what aspects of the patterns would the darkness depend?) can scaffold a more nuanced and deliberate discussion regarding the ethics of certain design features.

In another example, could be asked to:

1. Research recent patents awarded for videogames and related technologies.
2. Develop patterns from these recent patents
3. Discuss how/why these patterns could be dark and how games that implement the patterns could avoid the moral issues (if any)

For instance, in 2021 Sony Corporation was granted a US patent titled “Spectators Vote to Bench Players in a Video Game” (Osman & Chow, 2021). As the patent’s title suggests, the idea is that people watching a livestream videogame event can vote and provide feedback to competitors in the game and the voting process can result in the removal of a player from the game. Furthermore, the patent describes an embodiment wherein a viewer can pay a certain amount using in-game or real-world currency to have a player removed directly from the game being watched. This patent was widely discussed in the gaming media (e.g. Chalk, 2021; Jiang, 2021) at the time for its potential ethical issues and ramifications. From the patent specification a student could articulate the following patterns:

**Vote-to-Boot** - A livestreamed esports game allows spectators to vote to remove a player from a game. The player cannot veto the vote although a certain 60% voting threshold needs to be met (e.g. 40%, 70%, etc.).

**Pay-to-Boot** - A livestreamed esports game allows spectators to vote to remove a player from a game. To remove a player, audience members must simply pay a pre-determined (by the game’s operator) “fixed price” that can be paid in in-game currency or real-world currency. As soon as an audience member pays the price, the player is removed from the game.

Instead of a written assignment, students could engage in a class discussion that would proceed by teasing apart the potential moral issues and considerations of implementing such patterns in a game (e.g. what negative player experiences might result from these patterns? Do these result from elements included for the player’s best interests?, etc.). As part of this discussion students could discuss specific implementations – arguing for certain ones that might be more (or less) objectionable. By envisioning themselves as game creators implementing the above patterns in their games they can also begin to consider what they might need to do when it comes to unexpected consequences. For example, should anything be done if certain players (e.g. women, under-represented minorities) are consistently “booted” more often than other players?, Might implementing these patterns go against notions of “fair-play” by possibly encouraging players to turn on each other and rile up the audience in order to get them to boot their rivals?).

The point of these exercises is to ask students to speculate on something that is under-specified (e.g. they have no details on how the *Max Payne*’s cheater’s pool actually works) or plausible (to our knowledge, the Sony patent above has not been implemented in a videogame). Therefore, at the time they are doing their analysis, no one really knows how people would react to their implementation, what the actual resulting experiences will be, and so on. This type of creative speculation has been used in other areas, e.g. computer ethics education, as a way to encourage students to consider unintended negative consequences and ramifications of present-day technologies. One such example involved having students write summaries of (hypothetical) episodes of the science fiction TV show *Black Mirror* – known for its episodes exploring the potential harms of current technologies (Klassen & Fiesler, 2022). In this sense, this speculative exercise could also be applied to existing patterns where students could be asked to:

1. Choose an existing (and generally seen as benign) game feature or technology
2. Describe a dark pattern that could be derived directly from the above
3. Write a short description of an imaginary videogame that implements that feature in a morally problematic way.
4. Explain how that implementation is problematic and how it is derived from the dark pattern.

In the case of some benign patterns, it might not even be necessary for students to derive potential morally problematic misuses. Some of these patterns might benefit from examination from an alternate perspective. For example, consider the following:

**Deliberate Misses** – In a challenging reflex-based 2D action shooting game the first few shots fired by enemy characters are manipulated so that they deliberately barely miss the player (as seen in Luftrausers).

**Secret Buffs** – In a PvP match, players secretly start the match with a few buffs (e.g. extra damage, improved auto-aim) that wear off after a player gets a few kills. The buffs are increased if it’s your first PvP match (as seen in some titles in the Gears of War series).

Elements such as the above are commonly used in games. The justification for lying/deceiving the player is often that it provides the player a feeling of success and skill and is often used early in a game experience. In other words, it is done in the player’s best interests (have a feeling of success in the game). Even if we assume that the feeling of success is real (this would be an empirical question), how long does this effect last for and what happens after that “wears off”? In other words, what happens to the player after the deliberate misses/secret buffs are removed. Does the player have a negative experience as they realized that their (actual) skill is less than they imagined? How likely are players to feel let down afterwards when their actual skills must face the challenge of non-assisted gameplay? Are these patterns examples of manipulating the player into believing they are more skilled than they really are? Are design elements like these inherently patronizing? In the context of other games well known for their challenge and difficulty (e.g. Dark Souls, Cuphead), players often comment on their appreciation for how the game’s design is “fair”, player’s know where they stand in relation to the game in terms of their skill and mastery. While there aren’t moral concerns about games that are “easy” – do undisclosed game systems that dynamically modify a player’s understanding of their skill in relation to the game pose a moral conundrum?

**ANNOYANCE-BASED PATTERN DEVELOPMENT**

Another way to use dark patterns in game ethics education is as a lens for questioning current and existing practices and design elements that people complain about but that may or may not have moral considerations to them. We’ve labelled these as “annoyances” because, as a pedagogical exercise, it can be useful in helping students disentangle their feelings, opinions, and perceptions regarding a certain design element from what moral objections they may have about it: is a design element merely an annoyance (e.g. I don’t like this), or is there some fundamental moral principle that supports students’ notions? (e.g. this is wrong). Consider for example:

**Already-on-the-Disc DLC** – When buying a physical copy of a game, bonus DLC content is already on the disc but needs to be “unlocked” by paying an additional fee. The DLC content was never advertised as being included in the game purchase (as seen in many titles).

**Uninterruptible Cut Scenes** – In-game cinematics/movies that are not interactive. Players also cannot pause them or skip past them (as seen in many titles).
The above are things that game players often complain about. Using the dark pattern definition can help students articulate their objections to such elements and frame them in a context that addresses (supposedly), the moral source of their distaste. Are uninterruptible cut scenes objectionable because they demonstrate a lack of respect for the player’s time and interest? (after all, they are “forced” to watch them) Or perhaps they are in the best interest of the player? (i.e., when the cut-scene relays important information the player will need to better understand the game and succeed) Conversely, perhaps already-on-disc DLC should be framed as allowing the player faster/easier access to content because it is merely “unlocked” rather than requiring the additional effort of downloading and installing? (i.e., it is in the player’s best interest as you do not have to spend time downloading the DLC) Also, what is the negative experience resulting from already-on-disc DLC? If this cannot be articulated, then arguing that the pattern is dark becomes much more challenging.

We do not argue that the above are dark patterns (or not) – rather that an approach that begins with students raising objections about commonly recurring design elements they do not like that then proceeds with them examining them through the lens of dark patterns can help students begin to understand the difference between a design element that is morally objectionable from one that is not preferred for reasons of taste or preference. See Appendix A for additional patterns that can be used for classroom discussion and analyses.

DISCUSSION AND CONCLUSIONS
We have described different ways that the concept of dark patterns in games could be used in the classroom for game ethics education. Specifically, we argue that dark game design patterns can serve as a framework for:

1. Scaffolding ethical speculation around existing game features and technologies.
2. Exploring potential ethical issues and concerns in near-now game features and technologies.
3. Supporting reflection regarding existing game features perceived as “troublesome” in order to articulate whether or not they should be considered as dark.

However, there is an additional consideration hinted at earlier: in what ways might the definition of dark patterns itself be problematic? As Deterding et. al. (2020) argue, the definition of dark game design patterns is not grounded in a specific ethical framework. Similarly, Dupont and Malliet (2021) argue that the definition is vague and potentially unusable due to an inconsistency between the “darkness” of a pattern being objectively determined via the definition while also being dependent on contextual details and information. In many of our earlier examples we have, in fact, argued that the speculative exercise of proposing potential dark patterns can be guided by questions that examine contextual details (and particulars of implementation) as a way to determine “darkness”. We have also shown how the definition does not provide guidance for certain kinds of situations that may arise – e.g. the Cheater’s Pool pattern that presumably affects cheaters and non-cheaters differently. In the context of this article, the merits (or lack thereof) of Zagal et. al.’s definition (2013) are a feature – especially when they are used to spur discussion and reflection on how the definition may fail to encompass design elements deemed morally problematic or perhaps include some that may be benign. Therefore, the meta-question of what is a dark design pattern? And, how should one define them can also be pedagogically productive – alerting students to identify and engage with “alternative ethical issues or vantage points on game design” (Deterding et al., 2020) that are not considered by the original definition. Encouraging students to articulate and propose their own improved definitions –
especially after engaging with exercises such as those described earlier – could promote even deeper reflection and understanding. A better and deeper understanding of the moral implications and concerns of game design elements is, of course, something to embrace and look forward to.

**BIBLIOGRAPHY**


Levy, L. (2021). *Ethics and Gaming Course—Georgia Institute of Technology, USA*. -- 7 --


Orland, K. (2012, June 13). Rockstar condemns Max Payne 3 cheaters to play only against each other. *Ars Technica*. https://arstechnica.com/gaming/2012/06/rockstar-condemns-max-payne-3-cheaters-to-play-only-against-each-other/


APPENDIX A – SAMPLE PATTERNS FOR DISCUSSION
The following are additional sample patterns that [First Author] has used in games ethics education over the years for assignments, classroom discussions and more. No claims are made regarding their “darkness” or lack thereof.

Daily Playing Time Bonuses – Player is provided rewards consisting of in-game items and game currency based on how much time they spend playing the game in a 24 period. The more time spent playing, the better the rewards (as seen in many games-as-a-service).

Forced Additional Online Account Creation - Before being able to play a game that was purchased or provided via one an online service (e.g. Microsoft GamePass, Steam) the player must create an online account with an additional (different) online service (as seen in Doom Eternal on Microsoft GamePass).

Difficulty-Setting Dependent Achievements – A game has meta-rewards (e.g. achievements/trophies as seen on Playstation, Xbox, and Steam) for reaching certain moments in a game’s campaign or story mode (e.g. finishing the first chapter, completing the game), but these are not awarded if the player plays the game on an “Easy” difficulty setting (as seen in many console games).

Multiple Playthrough Required for True Ending – A game requires the player to play through the game to the end multiple times before they can see/experience the game’s “true ending” that explains the game’s narrative in a satisfying way. Playthroughs are non-trivial in terms of playing time/duration (as seen in NieR Automata and other games).

Location Changing Special Vendor - In-game vendor who sells special in-game items appears in a different in-game location every week. Vendor is only available for a limited time and the location “rotates” randomly between a handful of places that are very far from each other (as seen in Destiny 2).

Insufficient Rewards – Game occasionally provides premium currency rewards, but these are in amounts that are not enough to purchase anything in the in-game store. It would take months of gameplay before a player could save enough premium currency to buy something (as seen in many F2P mobile games).

Unrevealed Main Character – In a sequel to a popular game the main character is different from the one used in promotional materials, free game demo available prior to launch, and trailers. After the opening section of the game the actual main character is revealed and is the sole character used for the rest of the game (as seen in Metal Gear Solid 2).

Fake Online Play - Players create a username and connect to a multiplayer lobby and play against another player.

Always Online Single-Player – A single-player game requires that the player maintain an “active” internet connection while the game is being played for security and authentication purposes. The internet connection is not used for any other reason (as seen in many games).

Cycled-out Content – A online paid game has been “live” for many years and has had multiple expansions that added a lot of new paid content to the game. The developers have decided to overhaul the game by “cycling out” certain paid content (game locations, missions, etc.) to streamline the game experience and make it more welcoming to new players who are otherwise overwhelmed (as seen in Destiny 2).

Locked Bonus Rewards – The game provides rewards labelled as “bonus” that the player cannot obtain unless they are paid for separately. The locked bonus rewards
expire or disappear after a certain amount of time has passed (e.g., Buy “keys” to unlock “bonus chests” that disappear within a week if they are not opened).

**Timed Achievement for Non-Play** – The game provides a trophy/achievement that is only unlocked or obtained when a certain amount of time has passed during which the player has not logged in or played the game (e.g., Achievement is awarded after 5 years have passed as seen in Stanley Parable).

**Incompletable Collection** – An (otherwise) single-player game encourages players to collect virtual items, but the collection is not completable without trading virtual items with other players. (As seen in many Pokémon games).

**In-Game “Rate Me” Requests** – Occasionally, a game is interrupted to ask the player to provide a rating and review in the app store. The player can dismiss the request or respond to it. If they dismiss it, it will appear again sometime in the future during another player session. If the player rates the game, the interruption never appears again. (as seen in many games on the Apple appstore).

**Epic Lootbox Unlock Notification** – Game periodically notifies players of their friends’ epic finds when they open lootboxes purchased with real-world money. However, it does not provide information for when those finds occurred – they might be recent, or from a long time ago. (as seen in some mobile games).

**Battlepass Progression** – Players pay a one-time fee to activate a Battlepass that allows them to “progress” by obtaining increasingly more desirable in-game rewards for regularly playing a game. The Battlepass is active from the moment it is purchased until its pre-set expiration date (e.g. July 1st). Battlepass rewards are useful for competitive play (as seen in many mobile games).

**Faked Die Rolls** – The game informs the players of the odds of success of their actions (e.g. 75% chance to hit an enemy). However, when the game actually “rolls the dice”, the chance to hit is higher (e.g. 90% chance to hit an enemy, instead of 75%) (as seen in many strategy games).

**Critical Path Limited Availability Collectibles** – In games in which collecting virtual items is encouraged and a part of the critical path towards advancement, certain critical collectibles may only be available via timed special events and/or through micro transaction payments.

**Gated Reviews** – The game explicitly asks players to rate it on an app store. However, if the player rates the game anything other than the maximum score, a compulsory prompt appears asking the player to instead send a message to the developers for them to improve the game. (as seen on Dungeon Keeper iOS).

**Capped Inventory** – Game provides limited inventory space in a game that relies heavily on items and item collection. The inventory size can be increased by paying extra.

**Power Creep** – The in-game items purchased via micro-transactions slowly depreciate in value over time as new items (that are more powerful/valuable) are introduced to the game.

**Pre-Order Exclusives** – When customers pre-order a videogame, they are provided “unlock codes” for in-game content that is not available (or obtainable) by customers who did not pre-order the game.

**Multiple Currencies** – Many free-to-play games use multiple currencies for in-game purchases. One currency (aka “premium currency”) is usually really difficult to obtain via ingame actions and must be purchased with real world money. The other currency is regularly obtainable through regular play. Broadly speaking, premium currency is
used to purchase highly desirable in-game items, while the regular currency is used for “normal” items.