
Gamification Ethics: Exploitation and Manipulation

Author

Tae Wan Kim
Assistant Professor of Business Ethics
Tepper School of Business
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA
twkim@andrew.cmu.edu

Copyright is held by the author/owner(s). CHI 2015
April 18-23, 2015, Seoul, Korea.
Gamifying Research Workshop Papers

Abstract

Two main accusations made against gamification are exploitation and manipulation [3, 4, 5]. I explain under what circumstances gamification can be exploitative or manipulative. I maintain that gamification is characteristically vulnerable to an expressive form of exploitation and a non-reason-tracking form of manipulation. I suggest solutions.¹

Author Keywords

Gamification ethics, exploitation, manipulation.

Author Keywords

H.5.m [Information Interfaces and Presentation (e.g., HCI)]: Miscellaneous; K.4.1 [Computers and Society]: Ethics; K.8.0 [Personal Computing]: Games.

¹ This article is an abridged version of two unpublished working papers [13, 14]. I would like to thank the workshop organizers and two anonymous referees for their comments.

Problem: Exploitation

Notably, Bogost [4] dubbed gamification “exploitaionware.” Bogost claims that players gain only a small fraction of the entirely created value, while the company gains relatively too large a share of it, and the imbalance is exploitative. Yet, just because an outcome is asymmetrically distributed between employees and the employer, does not itself constitute a wrong. Suppose that a transplant surgeon gains \$35,000 by saving a patient. The patient benefits significantly more than the surgeon, but it would be odd to say that the patient exploits the surgeon. This does not mean, however, that gamification can never be exploitative. Below, I test gamification with two influential accounts of exploitation: the fairness account and the mere-means account.

According to the fairness account [21], an exploitative transaction is one in which *A* takes *unfair* advantage of *B* and a fair price is determined in a hypothetical environment in which some number of well-informed and unpressured buyers and sellers transact. Consider a famous exploitive case, *The Port of Caledonia and the Anna*, in which the master of a vessel in danger asked for assistance from a nearby tug and the master of the tug offered £1,000 or no help. The master of the vessel voluntarily agreed to pay £1,000. If there were at least one competitor, the offering price would be significantly lower than £1,000. This means that the tug master took an unfair advantage, so that the agreement was, although voluntary, exploitative.

Imagine a labor market in which labor for gamification is clearly specified within employment advertisements. In this market, some number of competing companies advertises, for instance, cashier jobs. The job descriptions of some companies include reference to the gamified working condition; other companies' descriptions do not. Would the addition of the gamified work change workers' salaries in the hypothetical market? Probably not, or, at least, it is unclear. In theory, offering a gamified working environment could allow a company to even slightly decrease wages for those who preferred gamified to non-gamified employment. From the fairness account, hence, it's difficult to say that gamification is exploitative.

The other influential view is the Kantian account of exploitation as the use of others as a mere means [1]. For Kant, what makes a person an end is the rational nature of humanity that enables her to deliberately set moral and practical ends. First of all, respecting the rational nature demands that we refrain from coercing or deceiving people. As a conceptual matter, however, gamification is voluntary [19, 20], so players are not coerced.² And it makes sense that the intent of using gamification (or any other incentives, especially at the workplace) is straightforward: gamification providers want to solve some problem *x* by making a working environment game-like.

The Kantian account requires more. Treating others as ends in themselves also demands us not to be "indifferent" to minimally maintaining their rational

² It is controversial whether or not gamification by definition is voluntary. I assume that it is. For a non-voluntary perspective, see [15].

nature of humanity [1]. Typically, this requires, in the context of organizational life, meeting minimum or reasonable safety standards and providing a minimum or living wage for employees [1]. Most companies that use gamification operate in the U.S. and other developed countries in which companies typically provide more than the legally defined minimum wage as well as other benefits including discounts on products, services, and merchandise.

The Kantian account can also be interpreted as an expressive theory, according to which an act that does not *substantively* (through e.g., deception, coercion, or minimum wage violation) use people as a mere means can *expressively* treat them as a mere means by showing disrespectful attitudes such as humiliation, contempt, offense, insult, etc. toward the rational nature of humanity [13]. For instance, treating adults as if they were children can often express the insulting attitude that the adults actually lack the rational capacity to act as adults [7]. Some cashiers who play Target's Checkout game, for instance, might reasonably believe that they are being treated as if they were in a preschool or kindergarten, in which once a four-year-old child behaves well and solves given problems, she earns a sticker or sometimes a candy or a marshmallow; the child is thrilled and wants to earn more stamps.

Solution: A Duty of Hermeneutics

In order not to express disrespect to others, we should consider how others would publicly *interpret* our acts. Hence, gamification designers have a duty to pay attention to the hermeneutical aspect of game elements. How can designers pay enough attention to the hermeneutical dimension? They can invite the input

of future interpreters, e.g., the Target cashiers, ask them how they interpret the game, and carefully consider their feedback. Perhaps a more structured way of fulfilling the duty of hermeneutics is possible. For instance, designers can develop an interpretive process based on the philosophy of Habermas's communicative approach [18], also suggested to courtroom judges who face interpretive issues [12]. No doubt, such a deliberative process cannot aim to lead involved parties to an ideally convergent interpretation. But it can at least promise a realistic and reasonably acceptable framework by which all parties can understand each other's perspectives and determine a realistic exit rather than stagnating in a dead-end interpretive debate. In addition, recognizing others' deliberative abilities is itself often valuable as an apt expression of respecting the rational nature of their humanity [17].

Problem: Manipulation

Relying on a philosophical analysis of bullshit [8], Bogost [3] claims that providers of gamification are bullshitters because they are indifferent to the justificatory quality—i.e., the truth value—of gamification. Here, Bogost, probably unknowingly, uses a reason-tracking account of manipulation [9], according to which gamification providers are manipulative to the extent that they do not track the justificatory quality of their means of influence. In typical cases, however, gamification providers seem to be committed to solving their problems using gamification. In other words, if it turns out that the gamified incentive does not work well—i.e., not true—, they find a new solution. That is, they track justificatory quality. So, Bogost's own charge is not typically true.

I develop a new perspective from Bogost's insight about bullshit. Providers who are not themselves bullshitters can attempt to influence players to be bullshitters—those who do not track the justificatory quality of their plays. Consider a well-known gamified activity, the ALS ice bucket challenge. Suppose that Alan genuinely cares about ALS patients and hopes more people become aware of the disease. He knows that his friends Ken and Taylor like to be recognized by others on Facebook. So Alan takes the ice bucket challenge himself on his Facebook and then nominates Ken and Taylor. During preparation for the challenge, Ken and Taylor realize that some filmed challenges are "liked" and "shared" many times—that is, earned points and badges, while others are simply ignored. So they decide to film their pouring ice bucket with a certain funny idea. They mention ALS in passing, but do not mean it. Many "like" and "share" their video. Ken and Taylor are excited about the points and badges. Many who would not otherwise know about ALS are now aware of ALS. Alan is happy about the outcome. So, everyone becomes happier.

Participating in the ice bucket challenge is itself a desirable act, and the worthiness of an action in part depends on the desirability of the act itself. Nonetheless, two actions that are equal in moral desirability may be of different moral worth, because the worthiness of an action also significantly depends upon the extent to which one is motivated to perform it with reasons that make it desirable [2, 14, 15]. Consider an often-invoked example: that a man wants to save a drowning child because he believes it is the right thing to do. The same man now wants to save the drowning child because he wants to seduce her mom. In both cases, the acts themselves are exactly the

same. But not many of us would deny that the moral worth of the act in the first case significantly differs from that of the inadequately motivated act in the second case.

I submit that Alan attempts to get Ken and Taylor's decision-making to fall short of a certain important moral ideal through a mental process I call "bullshitification,"³ which I define as a decision-making process in which, due to the influence of game design elements, a decision maker becomes detached from the reason that makes her action desirable, which can put the action at risk of significantly losing its moral worth. For this reason, Alan manipulates Ken and Taylor. To put it more generically, Person *A*'s act *x* manipulates person *B* when *A*, through *x*, attempts to get *B*'s decision-making to fall short of the moral ideal about moral worth and a relevant norm, "Do the desirable act with the motivation that makes the act desirable."

I maintain that a similar moral phenomenon can be attempted in many other cases of gamification. That is, providers of gamification often attempt to get workers' or customers' decision-making to fall short of the ideal about moral worth and to violate a relevant norm, "Do the right/good/desirable act with the motivation that makes it right/good/desirable." In other words, the aim

³ My notion of bullshitification differs from the so-called "motivation crowding out" [9], which says that extrinsic (especially monetary) incentives decrease intrinsic motivation. "One is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity itself"[6: 105]. Hence, a gamification player is intrinsically motivated when he or she receives no apparent reward except the game experience itself. Nevertheless, the intrinsically motivated player can be induced to be indifferent to the reason that makes the act morally desirable.

of gamification is to lead players to get motivated with points, badges, and leaderboards, and so to become indifferent to the reason that makes their action desirable, which can put their action at ethical risk of significantly losing the points of its moral worth.

Solution: Solemn Time

The problem with bullshitification is that players are induced to be motivated with a reason that does not make their action desirable, which in turn leads them to be indifferent to the right motivation that could make their job morally (more) worthy. Thus, any adequate solution must be something that can help players not to be indifferent to the right motivating reason. My suggestion is that players—that is, workers or customers—in gamified environments have what I want to call "Solemn Time," in which they learn about what their works or jobs are really doing, such as helping others, contributing to society, or enhancing important moral goods such as friendship or sustainability. For instance, the IT technicians who are excited about the points and badges in *OmniQuest* could have a chance during Solemn Time to get in touch with the reality of what they are doing through playing the game: helping senior patients in nursing homes.

One might worry that Solemn Time could distract workers from being immersed in the game and decrease the practical effectiveness of gamification. This idea about the incompatibility of fun with solemnity is not necessarily true, however. Endorsing Solemn Time can even boost the effectiveness of gamification as well as making it morally enlightened. In a field experiment about stressed out fundraising callers' performances [11], fundraising callers who were read stories about how their job could make a positive

difference in others' lives (e.g., a story about a person who would not have finished his college without the scholarship that fundraising callers created) showed

References

- [1] Arnold, D. G., and Bowie, N. Sweatshops and respect for persons. *Business Ethics Quarterly* 13, 2 (2003), 221-242.
- [2] Arpaly, N. Moral worth. *Journal of Philosophy* 99 (2002), 223-245.
- [3] Bogost, I. *Gamification is bullshit*. August 9, (2011b.)
<http://www.theatlantic.com/technology/archive/2011/08/gamification-is-bullshit/243338>
- [4] Bogost, I. *Persuasive games: Exploitationware*. May 3, (2011a).
http://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php.
- [5] Bréville, B., and Rimbart, P. *Losing on points: Do you play games, or are they playing you?* January (2014). <http://mondediplo.com/2014/01/09videogames>
- [6] Deci, E. L. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18, 1 (1971), 105-115.
- [7] Fahmy, M. S. Love, respect, and interfering with others. *Pacific Philosophical Quarterly* 92 (2011), 174-192.
- [8] Frankfurt, H. G. *On bullshit*. Princeton, NJ: Princeton University Press, (2005).
- [9] Frey, B., and Jegen, R. Motivation crowding theory. *Journal of Economic Surveys* 15, 5 (2001), 589-611.
- [10] Gorin, M. Towards a theory of interpersonal manipulation. In *Manipulation: Theory and Practice*, by C. Coons and M. Weber, 73-97. New York: Oxford University Press, (2014).

significantly higher job performance than those who did not have such an intervention or Solemn Time. Robust empirical research is required to support this idea.

- [11] Grant, A. M. The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions. *Journal of Applied Psychology* 93, 1 (2008), 108-124.
- [12] Hellman, D. The expressive dimension of equal protection." *Minnesota Law Review* 85, 1 (2000), 1-70.
- [13] Kim, T.W. Gamification and exploitation. Unpublished working paper.
- [14] Kim, T.W. Gamification and manipulation. Unpublished working paper.
- [15] Markovits. Acting for the right reasons. *Philosophical Review* 119 (2010), 201-242.
- [16] Mollick, E., and Rothbard, N. Mandatory fun: Consent, gamification and the impact of games at work. Unpublished working paper
<http://www.ivey.uwo.ca/cmsmedia/1332866/rothbard-03-13-2015.pdf>
- [17] Scanlon, T. M. *What we owe to each other*. Cambridge, MA: Harvard University Press, (1998).
- [18] Scherer, A. G., and Palazzo, G. Toward a political conception of corporate responsibility: Business and society seen from a Habermasian perspective. *Academy of Management Review* 32, 4 (2007), 1096-1120.
- [19] Werbach, K., and Hunter, D. *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia: Wharton Press (2012).
- [20] Werbach, K. (Re)defining gamification: A process approach. *Persuasive Technology: Lecture Notes in Computer Science* 8462 (2014), 266-272.
- [21] Wertheimer, A. *Exploitation*. Princeton, NJ: Princeton University Press, (1996).