Is “Value Capture” a Convincing Objection to Gamification? A Response to Nguyen

1. Introduction

The implications of ‘gamification’ have come under considerable philosophical scrutiny – especially of what ways, if any, gamification might be morally problematic. For our purposes, we will follow the definition used in the writings of C. Thi Nguyen, namely that gamification is the “introduction of game-like elements into practical life” (Nguyen, 2020, 189). This definition has a few notable characteristics, which differ from other meanings that have been used in the literature. First, it is possible, but not necessary, that if an activity is gamified, it thereby becomes a game. Second, there is a possibility of unintentional gamification.[1]

In a book and unpublished paper, Nguyen puts forward an intriguing moral objection to gamification, namely the notion of ‘value capture’. By value capture, he refers to a circumstance in which a person has subtle values, is put in a social environment which presents simplified versions of those values, and where those simplified values come to dominate her practical reasoning (Nguyen, Forthcoming, 1). That gamification will lead to value capture is presented as a serious objection against its use in many contexts.

Value capture may sound familiar from the adage of Goodhart’s law that “when a measure becomes a target, it ceases to be a good measure.” But, as Nguyen points out, value capture is more insidious than Goodhart’s law (Nguyen, 2020, 215). Goodhart’s law occurs when a simplified objective incentivises the wrong behaviour. Value capture occurs when we *internalise* those oversimplifications.

This paper responds to several reasons why value capture may be bad for us, and considers their strength as an objection to gamification. Notable in its omission from this paper is a discussion of the extent to which gamification promotes or inhibits autonomy.[2] This is an important question that has attracted much attention. However, I do not interpret autonomy concerns as being among the unique aspects of the value capture objection, and so they are omitted for brevity. §2 of this paper discusses whether value capture deprives us of a period of reflection about our values. §3 explains how hyper-explicit values don’t capture the full richness of human values. §4 offers an economic reframing, considering ways in which institutional values are systematically misaligned with our best interests. §5 suggests Nguyen ought to be more concerned about the *maximisation* element of value capture. §6 offers concluding remarks.

---

[1] Two influential definitions in the literature to contrast this with are “the use of game design elements in non-game contexts” (Deterding, 2011, 1), which would seem to rule out the possibility of a gamified activity itself being a game, and (Sicart, 2014, 225)’s “the design of services and products with the methods of game design, with the intention of engaging users in way similar to those of games”, which seems to make unintentional gamification impossible.
[2] Autonomy concerns are discussed in (Kim, 2018), (Rey, 2014), and (Sicart, 2014).

2. Value capture limits reflection

The goals that we internalise because of value capture were formulated by someone else. When we are value captured, we do not engage in reflection and formation about our goals and motivations. Nguyen repeatedly uses the metaphor that value capture is like purchasing your values 'off the rack', as opposed to creating them yourself (Nguyen, Forthcoming, 5).

A response to this concern is to say that Nguyen is simply confused about *when* this reflection takes place. Consider the Fitbit's step counter, an archetypal example of a gamified metric. My Fitbit motivates me to walk and run more, but only to increase a number on a screen, and not because (in the moment) I am inherently motivated by regard for my physical fitness. This is a case of value capture that most people would consider to be benign. In this case, the reflection that Nguyen seeks *did* occur – it occurred when I decided to buy and set up the Fitbit. The use of a Fitbit might be seen as analogous to pre-committing to eating the same foods for breakfast every day, or wearing the same suit. It is not that I thoughtlessly adopted the values of another person or institution, but that I decided to only engage in such reflection on my values *once*.

We might, for the time being, loosely sketch the following theory: The purportedly reflection-depriving elements of gamification are not morally problematic, under the following conditions. First, that you periodically engage in reflection to check whether the reasoning at the time you adopted the metrics is still sound. Second, that the gamified metric was adopted willingly.

It's partly an empirical question to what extent people deliberate before setting up gamified systems. But I take it that Nguyen's implicit contention is that there is something valuable about *certain kinds* of deliberation, which is lost in circumstances of value capture. Deeply reflecting on one's values is effortful, and perhaps unpleasant. The trouble with various forms of tracking technology is that they tempt us to adopt pre-fabricated values; doing so would be a great deal less work (Nguyen, Forthcoming, 3). The case for being concerned is strongest about whether people will be – to put it crudely – too lazy to engage in the most valuable forms of reflection.

Why is the adoption of simplified metrics so tempting? This is an interesting question for cognitive psychology, but it is worth briefly noting that one plausible cause is a cognitive bias for excessive trust in numbers; "Many people seem to trust quantified data simply because it is quantified" (Nguyen, Forthcoming, 34). This would appear to stem from the heuristic that ease of comprehension is *somewhat* correlated with correctness. Insufficient reflection, then, would be downstream of a misfiring heuristic.

3. Legibility and hyper-specific values

One of the most promising elements of Nguyen's research programme is its relevance to James Scott's influential and heterodox writings on anthropology and early state formation (Scott, 2017) (Scott, 2020). Scott emphasises what he calls 'legibility' – the property of being usable and understandable by a large centralised bureaucracy. States, by necessity, push significantly in the direction of legibility. It is for this reason that Scott views historic trends like the move toward the use of surnames and standardised units with some suspicion. The use of legible measures (which may be less useful than other measures which rely on context-sensitive knowledge) is what Scott calls the state's view of the world. By 'state', he refers to any large institution. Legibility is the antithesis of what Scott calls 'metis', or local knowledge.

Nguyen discusses Scott in the context of the selective pressures that gamified metrics are under. Gamified metrics are often created by private companies with a profit motive – but governments and universities engage in gamification too. Regardless of how benign the motives of an institution might be, gamified metrics are subject to demands for *usability* (Nguyen, Forthcoming, 4). Any hyper-specified metric will necessarily fail to capture the subtlety of human values.

Nguyen is considering in unpublished work whether hyper-specific values also discourage value *exploration* (Nguyen, Forthcoming, 17). A hyper-specific value makes it seem as though there is one exact set of values which is known to be correct.

Nguyen's emphasis on legibility strengthens his overall value capture thesis. It clarifies that the problems of value capture arise *irrespective* of what we assume about the motives of the persons or institutions producing the metrics; "It is possible to be value captured by a fully socialist bureaucracy" (Nguyen, Forthcoming, 33). This makes it a much more general worry. It also suggests that government regulation could not save us from the value capture dilemma.

4. A microeconomic re-framing

A third reason why Nguyen considers value capture to be ethically problematic is his expectation that the institutions which present the simplified metrics will have values systemically different from our own. In some form, this is undoubtedly true. Not even the most committed libertarian is likely to argue that private companies will create gamified metrics that are maximally aligned with your long-term eudaimonia (nor, I suspect, would the companies themselves). Cases in which institutions have goals which systemically differ from persons have been extensively studied, in the familiar guise of principal-agent problems from microeconomics. In this case, the agent is you, and the principal is the person or organisation creating metrics to get you to engage in a certain behaviour. More exotically, the principal and agent might *both* be you, at different stages in time, as you use tools like Fitbit in an attempt to engage in more future-oriented behaviour. Forms of gamification in employment like Target's 'Checkout Game' are an attempt to more closely align the incentives of the employer (the principal) and the cashiers (the agent) (Kim, 2018). In some ways, this aspect of Nguyen's contribution can be seen as a more philosophically rich reformulation of the principal-agent problem.

None of this is to allege ignorance on the part of Nguyen. He frames parts of his argument in economic terms, and even notes the clear line of inspiration on the perils of 'legible' metrics coming from the classic Hayekian critique of central planning (Hayek, 1945). The price system aggregates dispersed illegible knowledge into legible form; market prices are gamification's more successful cousin.

Nguyen's economic reframing of value capture comes toward the end of his paper. He divides value capture into the problem of *externalities* and the problem of *scale*.

First, value capture can produce significant negative externalities. Nguyen opens his paper with the example of his relative, who went on a holiday with friends who neglected her for the sake of increasing their Fitbit step count. Following his personal anecdote, I was in secondary school when the social media Snapchat introduced 'streaks'. Streaks increase for each consecutive day that users exchange photographs with one another – an approximation, perhaps, of Nguyen's fictitious 'FriendBit' (Nguyen, Forthcoming, 2). To find something interesting to photograph and reliably be distracted by your phone each day was a form of costly signalling that one cared about the maintenance of one's friendships. This imposed a large negative externality on anyone who did not use the streaks feature.

Second, Nguyen views value capture as synonymous with applying the wrong *scale* of values to a certain problem. On a large scale, the costs of non-tailored values are often outweighed by their benefits – for example, the use of something like tradeable pollution credits to abate carbon emissions (Nguyen, Forthcoming, 44). On a smaller scale, like in our aesthetic lives, the costs generally outweigh the benefits. Value capture is an *excessive* preference for the *largest-scale* values. Value capture is not always wrong; it is a tradeoff between efficiency and fine-tuning (Nguyen, Forthcoming, 16). Nguyen uses this to argue for a view he calls 'value federalism': that we should want different values suited to different scales. This name is not just a metaphor: the arguments for value federalism are structurally similar to arguments for federalism and devolution in political systems (Nguyen, Forthcoming, 41). James Scott, unsurprisingly, is an advocate for ceding greater powers to local government.

The externality framing of Nguyen is more convincing than the scale framing. In the case of an individual adopting values, it is in principle straightforward to test why we find certain (usually quantitative) presentations of information tempting to adopt as values. In the case of mismatched scale, the causal pathway is less clear: *why* precisely should we have such a preference for the largest-scale values? In my view, Nguyen fails to demonstrate that scale and the net benefits of simplified metrics are anything more than loosely correlated.

5. The perils of maximisation

The recent work on gamification parallels an earlier literature, namely on the alleged perils of ethical maximisation. A perfect utilitarian (or, follower of any other maximising theory with a well-specified objective) would be a kind of slave to their ideals. Arguably, no room would be left for such a person to engage in the idiosyncratic reflection, hobbies and interests necessary for the good life. Hence, Susan Wolf's famous contention that, if a moral saint did exist, you would certainly not want to be his friend (Wolf, 1982). She writes that "if the moral saint is devoting all his time to feeding the hungry or healing the sick or raising money for Oxfam, then necessarily he is not reading Victorian novels, playing the oboe, or improving his backhand" (Wolf, 1982, 421). Moral saints are value captured, with the gamified metric here being utility, or whatever other moral attribute is being maximised.

Wolf points out the ways in which we usually enjoy spending time with people because of their unique combination of interests and values and preferred conversation topics; the cultivation of these things is antithetical to maximisation. For the same reasons Wolf mentions, a person who has been value captured is more likely to be boring and unpleasant.

Moral saints strengthen value capture as an objection to gamification, insofar as they provide Nguyen with more ammunition. Even if you somehow *did* have a gamified metric which fully captured the richness and subtlety of your values, truly maximising that metric would make you a kind of slave. Some would argue that you would even lose your autonomy, your authenticity, or your *self*, in the process.

Moral saints underscore an important point, namely that it's difficult to say whether gamification is positive or negative on net. As I understand the term, moral saints are close to being good by definition. And yet, Wolf pointed out serious downsides of such people, which might temper our enthusiasm. Scholars in the Wolf tradition have been interested not so much in whether a scenario would be an improvement as in whether something of value and significance was lost along the way. Similarly, Nguyen is not against gamification in all cases, but urges us to be more cautious about its consequences.

6. Conclusion

This paper considered Nguyen's notion of value capture as an objection to gamification. It was shown that, while it is not necessarily problematic that we engage in less reflection because of value capture, there are significant grounds to be worried that people will avoid engaging in the most valuable forms of reflection. Simplified metrics are tempting to adopt as values – quite possibly because of demonstrable cognitive biases. With his understanding of legibility, Nguyen demonstrates that we should expect to see value capture in a wide range of circumstances, not just from private companies. Nguyen is correct to point out the ways in which you should be suspicious of the motives of the institutions producing gamified metrics. It is less clear that this element of value capture is anything other than the well-trodden ground of the misaligned incentives studied in political economy. Nguyen in certain other respects *understates* the strength of his objection, by failing to tie it in with the worries not only about flawed or over-simplified goals, but with *maximisation*. Value capture amounts to an intriguing but incomplete objection, with multiple promising threads to further explore.

Word count: 2483

References

Deterding, Sebastian. 2011. *Situated motivational affordances of game elements: A conceptual model*, Gamification: Using Game Design Elements in Non-Gaming Contexts, a Workshop at CHI. https://www.semanticscholar.org/paper/Situated-motivational-affordances-of-game-elements%3A-Deterding-Bredow/39e1dbd0db14a0f3a4667d0badcd9623331fadd2.

Hayek, F. A. 1945. "The Use of Knowledge in Society." *The American Economic Review* XXXV, no. 4 (September).

Kim, Tae Wan. 2018. "Gamification of Labor and the Charge of Exploitation." *Journal of Business Ethics* 152 (1): 27-39. https://www.proquest.com/docview/2103463746?accountid=10673&parentSessionId=rFKD0RA9XYz4Hg1D9TkO%2F%2BVu1ndogUSX4B656k0jZNM%3D&pq-origsite=primo&sourcetype=Scholarly%20Journals.

Nguyen, C. T. 2020. *Games: Agency as Art*. N.p.: Oxford University Press.

Nguyen, C. Thi. Forthcoming. "Value Capture." *Journal of Ethics and Social Philosophy*. https://philarchive.org/archive/NGUVCH.

Rey, P.J. 2014. *The Gameful World: Approaches, Issues, Applications*. Edited by Steffen P. Walz and Sebastian Deterding. N.p.: MIT Press.

Scott, James C. 2017. *Against the Grain: A Deep History of the Earliest States*. N.p.: Yale University Press.

Scott, James C. 2020. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. N.p.: Yale University Press.

Sicart, Miguel. 2014. *The Gameful World: Approaches, Issues, Applications*. Edited by Steffen P. Walz and Sebastian Deterding. N.p.: MIT Press.

Wolf, Susan. 1982. "Susan Wolf." *The Journal of Philosophy* 79 (8): 419-439. https://doi.org/10.2307/2026228.