

Do Game Bots Dream of Electric Rewards?

The universality of intrinsic motivation

PENNY SWEETSER, Australian National University, Australia

MATTHEW AITCHISON, Australian National University, Australia

The purpose of this paper is to draw together theories, ideas, and observations related to rewards, motivation, and play to develop and question our understanding and practice of designing reward-based systems and technology. Our exploration includes reinforcement, rewards, motivational theory, flow, play, games, gamification, and machine learning. We examine the design and psychology of reward-based systems in society and technology, using gamification and machine learning as case studies. We propose that the problems that exist with reward-based systems in our society are also present and pertinent when designing technology. We suggest that motivation, exploration, and play are not just fundamental to human learning and behaviour, but that they could transcend nature into machine learning. Finally, we question the value and potential harm of the reward-based systems that permeate every aspect of our lives and assert the importance of ethics in the design of all systems and technology.

CCS Concepts: • **Human-centered computing** → **HCI theory, concepts and models**; • **Computing methodologies** → **Reinforcement learning**; • **Applied computing** → **Computer games**.

Additional Key Words and Phrases: rewards, reward-based systems, gamification, games, play, intrinsic motivation, machine learning, reinforcement learning, flow, self-determination theory, learning

ACM Reference Format:

Penny Sweetser and Matthew Aitchison. 2020. Do Game Bots Dream of Electric Rewards?: The universality of intrinsic motivation. In *International Conference on the Foundations of Digital Games (FDG '20), September 15–18, 2020, Bugibba, Malta*. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3402942.3402965>

1 INTRODUCTION

Reward-based systems are used extensively throughout our societies¹, starting in early childhood, continuing through our education, and following on to our workplaces and government policies. It is an almost unquestioned fact within these systems that rewards (and punishment) are an effective way to encourage the behaviours that we want and to deter the behaviours that we do not want. However, we know from decades of psychological research into motivation and learning that extrinsically focused rewards might produce short-term behaviours that meet our desires, but that they can potentially cause long-term detrimental impacts that are contrary to our goals. This is especially true when we are dealing with behaviours that should be driven from intrinsic motivation, such as learning and positive social interactions. Interestingly, this faith in reward-based systems has flowed over into our design of technology, with examples such as gamification and machine learning.

¹Referring predominantly to the Global North.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

In this paper, we recount the history of reward-based systems and their role in our societies and systems. We discuss the relevant psychological theories that have evolved over the last century, from behaviourism, to self-determination theory, and flow. We examine how these theories have manifest in our design of software and technology, using gamification and machine learning as case studies. We compare the common threads related to curiosity, play, and intrinsic motivation that run through each of these systems. We question the ethics and effectiveness of the reward-based systems that underpin how we learn, work, play, and design technology. Finally, we propose that we have limited ourselves by thinking in terms of reward-based systems and that intrinsic motivation could hold the key to unlocking opportunities and advancements in human and machine potential.

2 REINFORCEMENT AND CONTROL

The concept of reinforcement emerged from the field of behaviourism and the work of B.F. Skinner in the 1950s, following on from Watson [46] and Thorndike [44]. Skinner's work related to operant conditioning and experimentation in the operant conditioning chamber, which became known as the Skinner box [39]. Skinner investigated how reinforcements, positive and negative, could alter and ultimately control the behaviour of what were termed organisms. Skinner mostly experimented on rats and pigeons, but humans were also organisms, no different to the subjects of his experiments. The underlying philosophy of behaviourism is that organisms are collections of behaviours, responding to their environment, and that controlling the environment will control the behaviours and, therefore, the organism.

In behaviourism, there is no concept of self or self-determinism. Rats, pigeons, and humans are collections of behaviours, with no consciousness, mind, or choice. In the Skinner box, the rat or pigeon (starved to 75% of its body weight) was placed in a sealed environment and was required to 'learn' that pressing a lever, in the right conditions, would result in the dropping of a food pellet (see Figure 1). The organism could also 'learn' to press the lever to remove negative reinforcement (electric shock). Although Skinner's experiments successfully demonstrated that behaviour could be manipulated and controlled with reinforcements, he also found that once the reinforcements stopped, the behaviours stopped. Skinner subsequently experimented with different schedules of reinforcement and found that a variable ratio reinforcement schedule (i.e., gambling) could slow the extinction of the behaviour. Although operant conditioning is referred to as a process of learning, it is more accurately described as a process of training, manipulating, or controlling an organism.

Although we might consider ourselves to differ from rats and pigeons, our societies abound with human-created Skinner boxes. Schedules of reinforcement (rewards) and punishment underpin our education systems, workplaces, and how we raise our children. Kohn [22] analyses decades of research on the use of rewards in classrooms, workplaces, and families and concludes that rewards are unethical and ineffective. They are unethical because they are designed to control an individual's behaviour to suit the person or organisation doing the controlling and rarely for the benefit of the individual. They are ineffective as they diminish interest in and devalue the behaviour itself and the individual will cease the behaviour when the rewards stop. They also have adverse effects on risk-taking, creativity, and social behaviour. The benefits are often short-term and narrow, but the detriments can be long-term with broader impact.

3 INTRINSIC MOTIVATION AND FLOW

Intrinsic motivation means doing something because it is inherently interesting or enjoyable, whereas extrinsic motivation means doing something because it leads to a separable outcome [33]. Skinner's reinforcements were extrinsic (or external) rewards (i.e., separable outcomes) that exist outside of the self. Although this was not a differentiation that Skinner made, as behaviourism does not have a concept of self. Following on from behaviourism, other theories

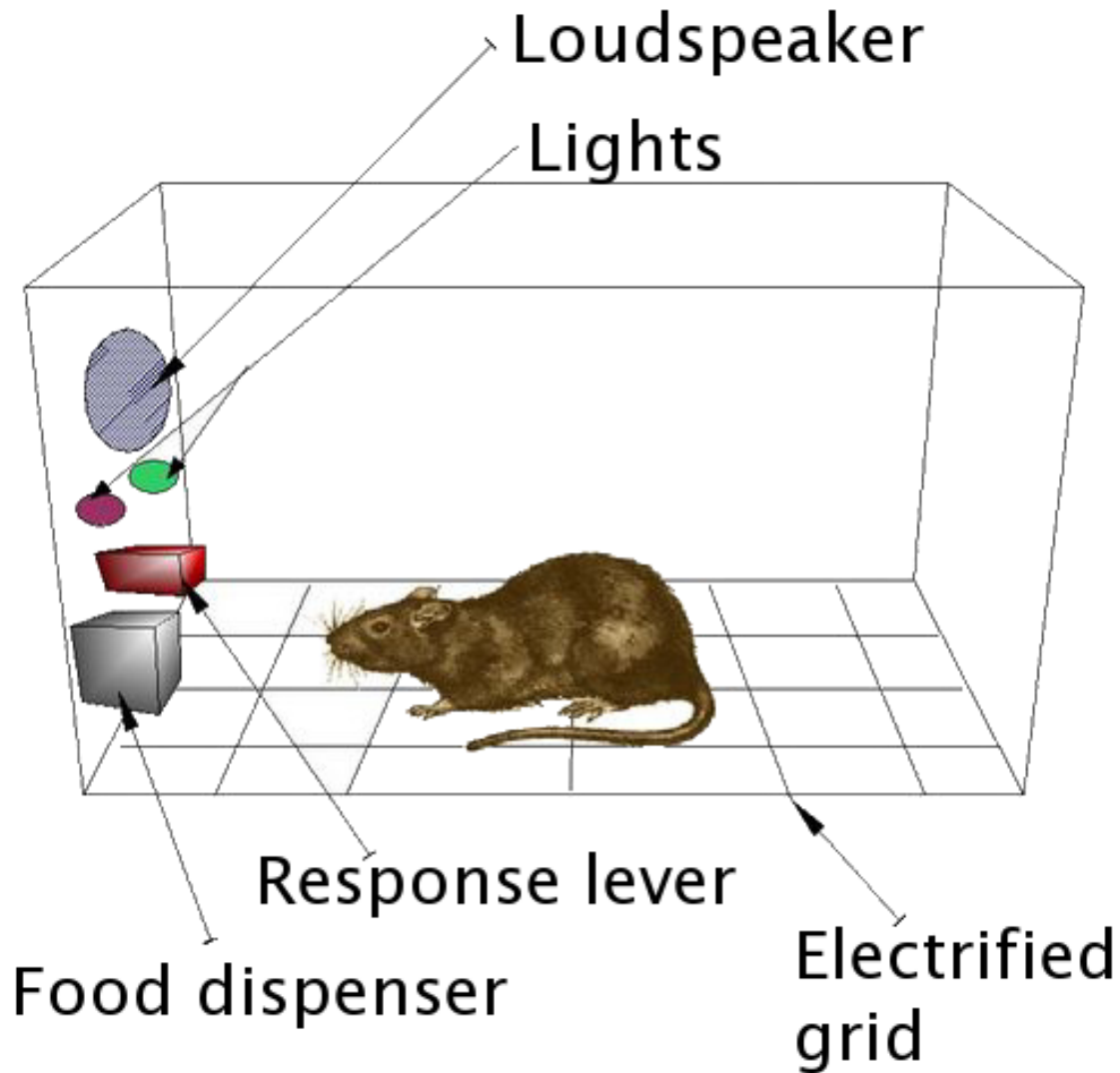


Fig. 1. An illustration of Skinner's operant conditioning chamber, also known as the Skinner box [48].

began to emerge in relation to why organisms might engage in certain behaviours. The concept of intrinsic motivation was first acknowledged in animal studies in the late 1950s, where it was observed that many organisms engage in exploratory, playful, and curiosity-driven behaviours without reinforcement [47]. Edward Deci pursued this theory in the 1970s and argued that some activities provide their own inherent reward, so that motivation is not dependent on external rewards [11]. Importantly, he raised the question of how extrinsic rewards might affect intrinsic motivation.

In the mid-1980s, Self-Determination Theory (SDT) was proposed to explain motivation, based on psychological needs, including innate needs for competence, autonomy, and relatedness [13]. Competence is a need for challenge and feelings of effectance, autonomy is a sense of volition or willingness, and relatedness is a sense of belongingness and connectedness to others. SDT distinguishes between intrinsic and extrinsic motivation and explains that satisfying the psychological needs can enhance intrinsic motivation, whereas external rewards can undermine intrinsic motivation [33]. The rewards are not the only factor, as they can be experienced to be more or less controlling depending on the contingencies and the environment [12]. Threats, deadlines, directives, evaluations, imposed goals, and competition pressure can also diminish intrinsic motivation as they are experienced as controllers of a person's behaviour [12, 33]. Rewards can also be given informationally (e.g., acknowledging good performance) without attempting to control and strengthen the behaviour. The rewards themselves are almost incidental, with the greater issue being the control versus self-determination of human behaviour in a social context [12].

Intrinsic motivation is manifest in behaviours that people often do for no external rewards, such as play, exploration, and challenge seeking [12]. This bears a close resemblance to the concept of flow [9], also originating in the 1970s, which seeks to understand what makes experiences enjoyable and why people engage in certain activities. Mihaly Csikszentmihalyi began his research with people who spend large amounts of time and effort on activities that are difficult, but provide no external rewards, such as composers, chess players, and rock climbers. His later studies were conducted with a broader set of people, asking them to describe how it felt when their lives were at their fullest and when what they did was most enjoyable. He found that optimal experience, or flow, is the same around the world and irrespective of social class, age, or gender. Very different activities are described in similar ways when they are being enjoyed. Flow is an experience “so gratifying that people are willing to do it for its own sake, with little concern for what they will get out of it, even when it is difficult or dangerous” [10]. The key element in flow is that it is an end in itself – the activity is intrinsically rewarding and autotelic. In flow, the goal is often an “excuse” for the process [27].

4 GAMES AND GAMIFICATION

One domain in which motivation, flow, rewards, and learning, along with manipulative design, all come into play is that of games and gamification. Gamification has been defined as the design strategy of using game design elements in non-game contexts [15]. However, the selection and design of these elements are key in determining whether the designed system will play to intrinsic or extrinsic motivation and the resulting impacts and implications for the end user. It would be a mistake to consider that simply taking individual game elements and applying them in a non-game context will result in increased motivation and engagement for the user, let alone considerations for their well-being and long-term interest in the tasks or system. The first thing to consider before employing “game design elements” is the nature and relationship of games and play.

There are many definitions for games, but they are commonly defined as rule-bound activities with goals and at least one player who tries to fulfil the goals [21]. Games have also been contrasted with play, as a form of play with goals and structure [24]. Caillois [7] used the terms “paidia” and “ludus” for playing and gaming, as two poles of play activities, with paidia referring to uncontrolled, free-form, and improvisational activity and ludus to restricted, structured, and goal-focused play. Games have been described in relation to play as both a formalised subset of play and as an activity that includes play [21]. However, the core learning is that play is a fundamental component of games and that games do not exist without play. Rules, structure, and goals alone do not constitute a game. It has been noted that gamified applications often focus on rule-bound, goal-oriented play, rather than open, exploratory, free-form play [15].

Games have been long acknowledged for their motivational pull and potential for use in other domains, with an early example being educational games in the 1950s [21]. Games are an example of the autotelic experience described by Csikszentmihalyi [10], as players are willing to play games for the experience itself, without need for external reward [43]. This, in part, explains the popularisation of the concept of gamification, since the first documented case in 2008 [15]. However, the evolution of gamification can also be tied to reinforcement and reward-based systems that predate modern game design. Many corporations have had loyalty programs running for decades, in which they reward customers with points, status, freebies, and perks for their continued patronage [29]. To refer to these systems as gamifications is to misunderstand the nature of games and the centrality of play to these experiences. It has been suggested that “pointsification” would be a more accurate name for such systems [32] or perhaps even “exploitationware” [3]. We suggest that “Skinnerification” could be another appropriate term. The expectation that this approach will increase motivation to engage with a system, or at least engagement that continues when the rewards stop, is also contrary to the decades of research discussed in the previous sections.

The most commonly used forms of gamification have been points, levels, leaderboards, and achievements, also referred to as reward-based gamification [29]. It is not surprising then, that these types of systems play to extrinsic motivations and can be prone to the same issues as other reward-based systems. Many gamification systems seem to draw inspiration from achievement systems on gaming platforms, such as Steam (see Figure 2), Xbox, and Playstation [17]. For example, some of the early drivers of the “new wave of gamification” [17], mobile application Foursquare and website StackOverflow (see Figure 2), have borrowed their design elements from game platform achievement systems [26]. It has been suggested that these systems themselves are gamifications of games, in the form of a shared meta-game layer across all games on the platform, and that they closely resemble corporate loyalty programs [17]. Game designers have questioned whether these achievement systems obscure the core game experience and shift playing motives from intrinsic enjoyment of the game to extrinsic hunting of achievements [17]. One possible reason for the narrow and reward-based approach of many gamification systems is that gamification has been driven by industry publications and frameworks [26], many of which have lacked validation or grounding in research [14]. It has been noted that reward-based gamification is tempting to use, as it is easy to implement and has an immediate effect (rewarding in itself), but reward-based systems have caused harm over the years [22] and reward-based gamification is no exception [29].

Nicholson [28] observes that gamification seems to assume that using the term “game” means that an activity will become a more engaging experience, despite generally using the least interesting part of a game (i.e., the points system). He notes that there are more effective ways than a scoring system to engage users. There has been substantial research in the field of games design and player experience that seeks to understand what makes a gaming experience engaging or enjoyable. For example, the PENS (Player Experience of Need Satisfaction) is a measure of need satisfaction in play, elaborated from SDT [34]. Similarly, GameFlow [42, 43] is a model of player enjoyment in games that includes eight elements (e.g., Challenge, Control, Clear Goals), which is based on Csikszentmihalyi’s flow. Zwart [53] notes that in applying GameFlow to a gamification system that a choice can be made for each element in whether to achieve the criteria via designing for intrinsic or extrinsic motivation.

It is also important to consider the extent to which the gamification is seeking to control versus support the user. Consistent with the findings of SDT, it is possible for reward-based gamification systems to be experienced to be more or less controlling. For example, leaderboards can be experienced as a form of goal setting [23] and game platform achievements can be experienced either intrinsically (as competence boosts) or extrinsically (as ego boosts), depending on the design and the player [8]. Nicholson [28] asserts the importance of taking a user-centred approach to designing

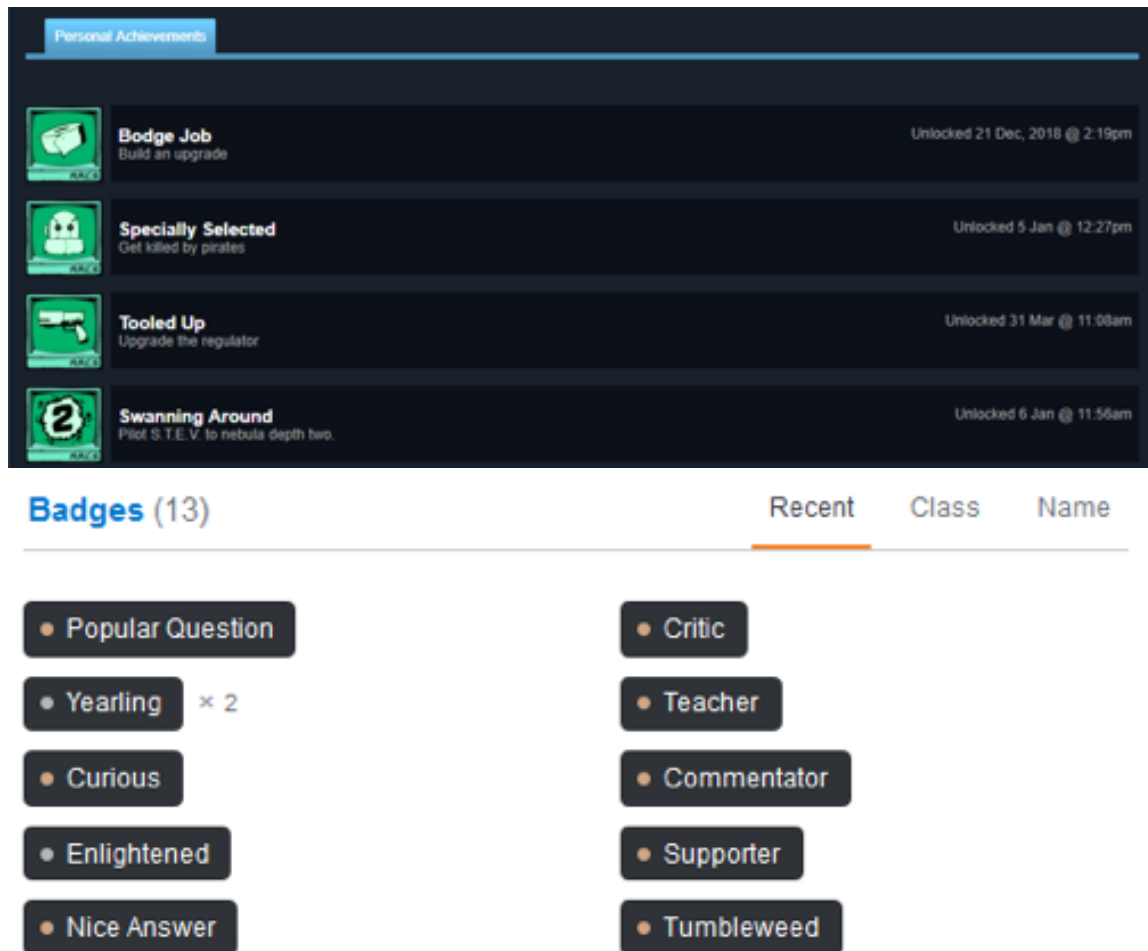


Fig. 2. Players earn achievements on game platform Steam (top), while website StackOverflow (bottom) includes badges and a reputation points system to motivate users to engage [40].

gamification systems, asking “how does this benefit the user?” for each design decision. User-centred gamification design is not about controlling the user, it is about empowering the user. The benefit to the user should be first and the organisation second [29].

Nicholson [29] contrasts reward-based gamification with meaningful gamification, which is defined as using game design elements to help build intrinsic motivation and meaning in non-game settings. Performing tasks for intrinsic reasons leads to a healthier mental state and well-being than performing tasks for extrinsic rewards [29, 33]. Meaningful gamification encourages a deeper integration of game mechanics into the non-game context, with a focus on the underlying activity and an understanding of where the integration of game elements makes sense [28]. Meaningful gamification uses gameful or playful layers to help a user find their personal connections that will motivate their engagement with an activity for long-term change [29]. What is “meaningful” depends on the individual and no one gamification fits all. This resonates with SDT, in which intrinsic motivation exists in the relation between individuals

and activities [33], and also with flow, in which the experience is a function of the individual and the task [10]. Users need to be empowered by the gamification system to create, learn, and demonstrate mastery in different ways. The system should afford meaningful engagement, but not prescribe it.

It has also been proposed that extrinsic motivation provided by gamification systems can help a user to establish their intrinsic motivation for an activity [52]. However, decades of research has shown that the impact of extrinsic rewards is usually to decrease intrinsic motivation [33]. Some recent research has suggested that extrinsic rewards in gamification systems can have a positive effect on intrinsic need satisfaction [49], but further research is needed in this area. There are also some situations in which reward-based gamification systems can be useful – when the user has no way to develop intrinsic motivation for an activity, when the goal is to teach a skill with a real-world value (which the user will recognise upon mastering the skill), or when it is necessary to create immediate or short-term change [29]. Recent research has proposed gameful design heuristics that include intrinsic motivation heuristics, extrinsic motivation heuristics, and context-dependent heuristics [45].

It is also important to consider the relationship between the gamification layer and the activities and tasks that it is aiming to support [21]. If the gamification and activities are not well integrated, then the user might end up focusing on the gamification to the exclusion of the activity. If the gamification represents an overly simplified model of the activity, then the user can develop less-effective behaviours. If the gamification has flaws, then users are likely to deliberately exploit weaknesses that they find, in effect “gaming the gamification”, which could provide a source of intrinsic motivation that the system itself lacks. Knaving and Bjork [21] recommend that the gamification model should not overshadow the main activity, that it should be opt-in or invisible, that users should not be penalised for not engaging with the gamification, and that any mandatory actions should be meaningful in regard to the main activity. Nicholson [29] also highlights the importance of choice in interacting with gamification systems, noting that play is optional and that users need to be able to choose to engage with the system on their own terms. This relates back to the SDT need for autonomy.

The considerations for gamification raised in this section resonate with the issues put forward by Kohn [22], leading to the questions of not only “is gamification effective”, but also “is gamification ethical”? Perhaps the latter is a question that can be applied not only to gamified systems, but to games more broadly. Many games employ design practices that can be considered manipulative, such as dark patterns [50], taking our understanding of psychology, HCI, and design, and applying them for subversive, rather than supportive, purposes. These questions of ethics in game design and gamification are well placed in the current broader commentary around ethics in AI and technology.

5 MACHINE LEARNING

Given the pervasiveness of reinforcement schedules in our everyday lives, it is not surprising, then, that reinforcement has spilled over into how we train machines. Reinforcement learning is an approach to machine learning that aims to learn which actions in a given situation are most likely to lead to maximum reward in the future [41]. Reinforcement learning was inspired by The Law of Effect put forward by Thorndike [44], and the concept of actions and consequences, or trial-and-error learning [41]. Unlike other forms of machine learning, it models an intelligent agent that can influence its environment. Reinforcement learning has applications in medicine [51], robotics [1], and advertising [6]. However, it has shown the most promise for tasks with clear rewards, such as the Atari computer games [25] and traditional board games (e.g., Go [36] and Chess [37]). In these domains, reinforcement learning methods greatly outperform any other algorithms and are often superhuman in ability.

Despite this success, these algorithms consistently perform poorly at tasks that have sparse or hard to find rewards [4]. Even games that are relatively simple for humans, such as Montezuma's Revenge (see Figure 3) are challenging for reinforcement learning in that the agent must learn a set of skills in order to find its first reward [2]. Hand-crafting dense rewards on a case-by-case basis has proven difficult and error-prone [4], leading to building interest in developing intelligent agents that can learn to self-reward, via intrinsic motivation. An intrinsically motivated agent engages in behaviour for its own sake, even if those actions are unlikely to have practical value in terms of the specified task [38]. However, what the agent learns from these self-motivated behaviours is often essential to developing the skills required to solve the intended task.

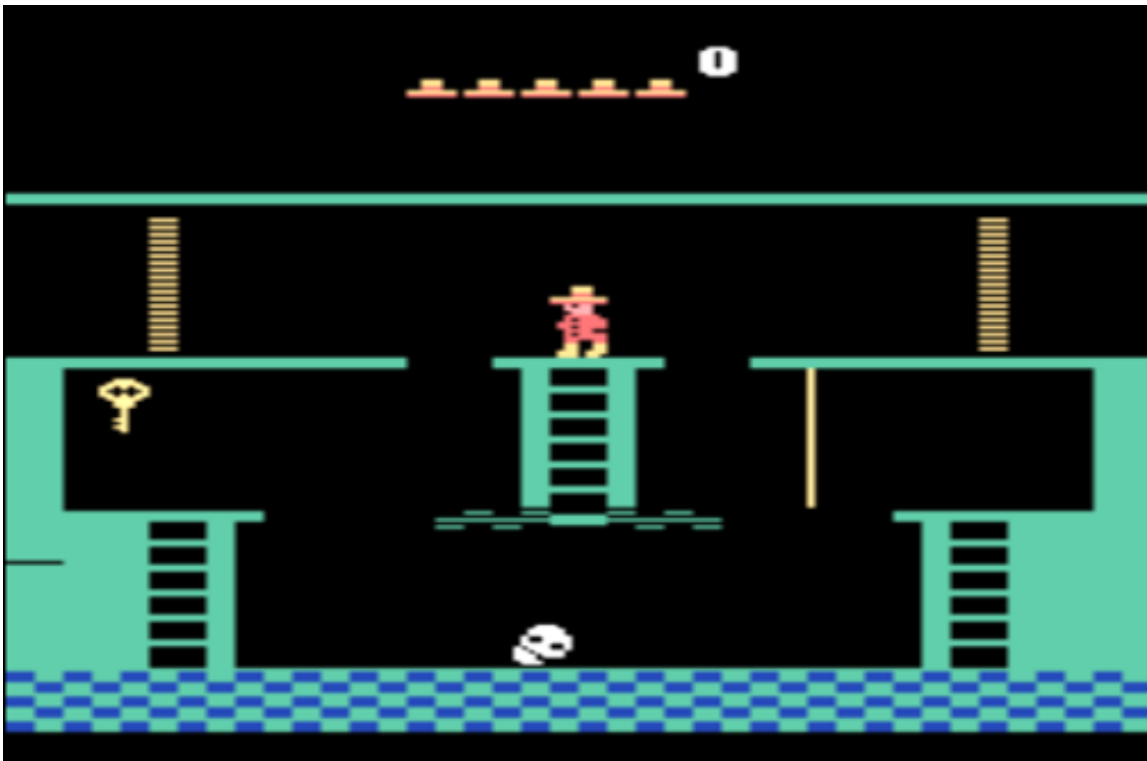


Fig. 3. First room of Montezuma's Revenge. The player must climb down the ladder, use the rope to jump to the platform on the right, avoid the skull, and collect the key in order to obtain the first reward.

Extrinsic rewards in reinforcement learning are premised on the belief that it is easier to recognise good behaviour than to describe it. For example, winning a game of Chess is much harder than identifying who won. However, in practice, structuring these extrinsic rewards has several problems. Agents might shortcut the reward mechanism, known as wireheading, allowing them to gain a reward but avoid the implied objective [16]. For example, an agent that was intended to learn to win a boat race instead learned to perform small loops, forever picking up speed boosts due to the rewards provided [30]. Extrinsic rewards can not only lead to unwanted behaviour but can potentially be so challenging to achieve that an agent might never reach a single reward. One agent trained to solve mazes almost never found the

exit, even after millions of interactions, and so failed to learn the task [35]. This gives rise to a 'chicken and the egg' problem as rewards are needed to learn appropriate skills, but those skills are needed to gain the first reward.

Both wireheading, and the difficulty of achieving the first reward can be addressed by giving the agents intrinsic motivation. Wireheading may not be alleviated completely, but as skills learned intrinsically are not dependant on the extrinsic reward, they tend to be broader and more generalisable [38]. Furthermore, intrinsic motivation can provide rewards for the agent to learn the basic skills required to achieve an initial reward, bootstrapping the learning process. Exactly what kind of intrinsic motivation is best in reinforcement learning is still not well understood, but two general ideas have emerged: learning about the environment and learning how to control it.

Agents that seek to learn about the environment, often referred to as novelty-seeking or curiosity-driven agents, reward the agent for learning about its world. Recent research has shown that providing an agent with a curiosity reward, via an Intrinsic Curiosity Module, enables an agent to explore enough of the game to reach its first rewards [31]. The curiosity bonus is awarded to the agent based on how poorly it can predict a future state given its current state and the selected action. That is, the agent is encouraged towards actions that have unexpected consequences. A surprising result was that agents trained using purely this intrinsic curiosity reward performed quite well on the extrinsically motivated task, even though they were never made aware of their real objective. One downside with the technique is a problem called the 'Static TV' problem [5], which is a thought experiment that argues that an agent with a remote control to a noisy TV would become stuck watching TV due to its unpredictable behaviour. This proves to be a serious problem, as environments often contain large sources of uncertainty that can never be learned and that contain no useful information.

A different approach to intrinsic rewards is to motivate an agent to be able to control its future. This is achieved by maximising an information-theoretic concept called empowerment. Empowerment is a measure of how much influence the agent's actions will have on its future state [20]. This causes the agent to prefer states from which it is able to keep its options open. One implementation of this is UNsupervised REinforcement and Auxiliary Learning (UNREAL), which supplements extrinsic rewards by learning how to control observed pixels while navigating a maze [19]. The combination of pixel control and extrinsic rewards allows the agent to greatly outperform previous algorithms in sparse reward problems such as labyrinth navigation.

Intrinsic motivation has shown a great deal of promise in solving sparse reward problems in reinforcement learning, but there are still many unanswered questions. For example, what kind of intrinsic motivation is most useful, and most generalisable? How should intrinsic rewards be combined with extrinsic rewards? That is, should an agent only rely on intrinsic motivation or transition somehow from intrinsic to extrinsic motivation over time? Outside of games, reinforcement learning has not yet delivered on its promise of optimal control. The reality may be that extrinsic rewards alone are simply not enough to properly train agents to robustly solve tasks with real-world complexity.

6 DISCUSSION

There has been a great deal of research and debate related to reward-based systems and their implications for humans over the last few decades. The focus of this paper is to consider how we take these insights and lessons learned and apply them to our design of technology. Debate and discussion around reward-based systems often centres on the effectiveness of these systems. In this paper, we have considered issues of effectiveness related to the impact of short-term rewards on long-term behaviour and motivation. We have also explored the situations in which reward-based systems might be more or less effective. The question that is not often explored and that is, perhaps, even more pertinent to the broader context of designing technology currently is the question of whether reward-based systems are ethical. The answer to

this question must be considered with respect to the design intent of the systems, are they to support or to control? Are they to benefit the individual, in terms of learning and development, or are they to support the organisation in having the individual comply and perform? Are we designing systems that seek to achieve short-term organisational-focused outcomes (e.g., increasing profits), at the cost of the well-being of the individual? There is no clear answer to the question of whether reward-based systems are ethical, because the answer is that it depends on the context. In the same way that machine learning and data collection are not inherently unethical, we must consider how, why, and what future impacts could occur. What we can conclude is that reward-based systems can be designed and used in an unethical way and that we must place ethics front and centre in our design process. It is unethical to leverage our understanding of psychology to manipulate people into engaging in behaviour that serves an organisation or government, but that is to their own detriment. Valuable future work lies in further defining and refining the ethical design of reward-based systems in technology, including games and gamification.

With respect to machines, reinforcement learning has demonstrated that rewards alone are insufficient for learning effective behaviours to solve tasks with real world complexity. When rewards are sparse and it takes a long time to receive the first reward, the machine needs to first develop skills, in order to be able to find the first reward. For the phenomenon of wireheading, machines are susceptible to learning the wrong thing or focusing too narrowly on a reward. This mirrors the behaviour of humans in gamified or other reward-based systems. Presenting a reward in either case is inviting reward-seeking behaviour, for which organisms and machines are incentivised to find the minimal path to the reward. Depending on the design of the system and rewards, this could involve short-cutting or missing the designed behaviour or outcome entirely. As discussed with respect to gamification, humans can take great joy in the behaviour of gaming the system in this way. It is also possible that the resulting behaviours will be counter to the purpose or intent of the system.

The shortfalls of reinforcement learning reflect the incomplete understanding inherent in using reinforcement, or Thorndike's Law of Effect [44], as the foundation for modelling learning. Humans do not learn via a process of reward-seeking, we learn through an innate drive for discovery, in which the experience is its own reward. There is an element missing in our understanding and modelling of learning behaviours in the context of reinforcement learning. Researchers have attempted to bridge this divide with the same approach that has proven to be most effective for human learning, intrinsic motivation, in the form of curiosity and empowerment. Curiosity (seeking new states) and empowerment (increasing choice of state) can be compared to SDT needs of competence and autonomy. Interestingly, when using intrinsic motivation, machines have incidentally increased their performance at the desired task, of which they were unaware. We propose that the element missing from reinforcement learning is the same element missing from behaviourism on which it is premised, that is, the concept of self. To learn and to be effective at a task, you need more than a set of behaviours and you need to be more than a set of behaviours.

We have conducted our analysis primarily from an individualistic perspective, grounded in psychological theory. We have focused on the rewards of the individual, both human and machine, to the exclusion of the collective. Although out of the scope of this paper, valuable future work lies in a sociological analysis, including both human collectives and multi-agent systems. There are many interesting examples of social reward systems within video games, such as massively multiplayer online role playing games (e.g., World of Warcraft). Likewise, there have been interesting recent advances in multi-agent reinforcement learning, including applications of intrinsic motivation [18].

7 CONCLUSION

Biologically-inspired computing, including machine learning, affords us the ability to hold a mirror up to ourselves and to test our fundamental understanding of how the processes we are modelling work in our natural world. There is value in not just improving our computing techniques, but also in questioning our assumptions about our world and ourselves when designing these systems. Machines, like humans and other organisms, struggle to learn, or to learn accurately and effectively, in an environment where external rewards are the only motivator. Like humans, machines will focus too narrowly on the extrinsic rewards and learn the wrong thing or be unable to generalise what they learn. Like us, machines must first learn how to learn, before they can learn any external task effectively. Furthermore, it seems that there is a universality to the satisfaction of needs, particularly with respect to learning and play, in that organisms and machines can be deeply motivated by curiosity, exploration, empowerment, and self-determination. In answer to the question “do game bots dream of electric rewards?” - maybe they do, but only because they have been limited by our faith in reward-based systems. Let us be more than “*a society of loyal Skinnerians, unable to think our way out of the box we have reinforced our way into*” [22].

REFERENCES

- [1] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. 2017. Hindsight experience replay. In *Advances in Neural Information Processing Systems*. 5048–5058.
- [2] Yusuf Aytar, Tobias Pfaff, David Budden, Thomas Paine, Ziyu Wang, and Nando de Freitas. 2018. Playing hard exploration games by watching youtube. In *Advances in Neural Information Processing Systems*. 2930–2941.
- [3] Ian Bogost. 2013. *Exploitationware*. Palgrave Macmillan US, New York, 139–147. https://doi.org/10.1057/9781137307675_11
- [4] Yuri Burda, Harri Edwards, Deepak Pathak, Amos Storkey, Trevor Darrell, and Alexei A Efros. 2019. Large-scale study of curiosity-driven learning. In *International Conference on Learning Representations*.
- [5] Yuri Burda, Harrison Edwards, Amos Storkey, and Oleg Klimov. 2019. Exploration by random network distillation. In *International Conference on Learning Representations*.
- [6] Han Cai, Kan Ren, Weinan Zhang, Kleanthis Malialis, Jun Wang, Yong Yu, and Defeng Guo. 2017. Real-Time Bidding by Reinforcement Learning in Display Advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (Cambridge, United Kingdom) (WSDM '17)*. ACM, New York, NY, USA, 661–670. <https://doi.org/10.1145/3018661.3018702>
- [7] Roger Cailliois. 1961. *Man, play and games (Meyer Barash, Trans.)*. Free Press of Glencoe, New York.
- [8] Carlos Cruz, Michael D. Hanus, and Jesse Fox. 2017. The need to achieve: Players’ perceptions and uses of extrinsic meta-game reward systems for video game consoles. *Computers in Human Behavior* 71 (2017), 516–524. <https://doi.org/10.1016/j.chb.2015.08.017>
- [9] Mihaly Csikszentmihalyi. 1975. *Beyond Boredom and Anxiety*. Jossey-Bass Publishers, Washington.
- [10] Mihaly Csikszentmihalyi. 1991. *Flow: The Psychology of Optimal Experience*. Harper Perennial, New York.
- [11] Edward L Deci. 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology* 18, 1 (1971), 105. <https://doi.org/10.1037/h0030644>
- [12] Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125, 6 (1999), 627.
- [13] Edward L. Deci and Richard M. Ryan. 1985. The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality* 19, 2 (1985), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)
- [14] Sebastian Deterding. 2015. The lens of intrinsic skill atoms: A method for gameful design. *Human-Computer Interaction* 30, 3-4 (2015), 294–335. <https://doi.org/10.1080/07370024.2014.993471>
- [15] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining “Gamification”. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (Tampere, Finland) (MindTrek '11)*. ACM, New York, 9–15. <https://doi.org/10.1145/2181037.2181040>
- [16] Tom Everitt and Marcus Hutter. 2016. Avoiding wireheading with value reinforcement learning. In *International Conference on Artificial General Intelligence*. Springer, 12–22. https://doi.org/10.1007/978-3-319-41649-6_2
- [17] Juho Hamari and Veikko Eranti. 2011. Framework for Designing and Evaluating Game Achievements. In *Proceedings of the 2011 DiGRA International Conference: Think Design Play (DIGRA '11)*.
- [18] Max Jaderberg, Wojciech M Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castaneda, Charles Beattie, Neil C Rabinowitz, Ari S Morcos, Avraham Ruderman, et al. 2019. Human-level performance in 3D multiplayer games with population-based reinforcement learning. *Science* 364, 6443 (2019), 859–865.

- [19] Max Jaderberg, Volodymyr Mnih, Wojciech Marian Czarnecki, Tom Schaul, Joel Z Leibo, David Silver, and Koray Kavukcuoglu. 2016. Reinforcement learning with unsupervised auxiliary tasks. In *International Conference on Learning Representations*.
- [20] Alexander S. Klyubin, Daniel Polani, and Chrystopher L. Nehaniv. 2005. All Else Being Equal Be Empowered. In *Advances in Artificial Life*, Mathieu S. Capcarrère, Alex A. Freitas, Peter J. Bentley, Colin G. Johnson, and Jon Timmis (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 744–753.
- [21] Kristina Knaving and Staffan Björk. 2013. Designing for Fun and Play: Exploring Possibilities in Design for Gamification. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (Toronto, Ontario, Canada) (*Gamification '13*). ACM, New York, NY, USA, 131–134. <https://doi.org/10.1145/2583008.2583032>
- [22] Alfie Kohn. 2018. *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes (25th anniv. ed.)*. Houghton Mifflin Harcourt, Boston.
- [23] Richard N. Landers, Kristina N. Bauer, and Rachel C. Callan. 2017. Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior* 71 (2017), 508–515. <https://doi.org/10.1016/j.chb.2015.08.008>
- [24] Kevin Maroney. 2001. *My entire waking life*. Retrieved September 18, 2019 from <http://www.thegamesjournal.com/articles/MyEntireWakingLife.shtml>
- [25] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533. <https://doi.org/10.1038/nature14236>
- [26] Lennart E. Nacke and Sebastian Deterding. 2017. The maturing of gamification research. *Computers in Human Behavior* 71 (2017), 450–454. <https://doi.org/10.1016/j.chb.2016.11.062>
- [27] Jeanne Nakamura and Mihaly Csikszentmihalyi. 2014. *The Concept of Flow*. Springer Netherlands, Dordrecht, 239–263. https://doi.org/10.1007/978-94-017-9088-8_16
- [28] Scott Nicholson. 2012. A User-Centered Theoretical Framework for Meaningful Gamification. In *Games+Learning+Society (GLS) Conference 8.0*. 223–229.
- [29] Scott Nicholson. 2015. *A RECIPE for Meaningful Gamification*. Springer International Publishing, Cham, 1–20. https://doi.org/10.1007/978-3-319-10208-5_1
- [30] OpenAI. 2016. *Faulty Reward Functions in the Wild*. Retrieved September 18, 2019 from <https://openai.com/blog/faulty-reward-functions/>
- [31] Deepak Pathak, Pulkit Agrawal, Alexei A. Efros, and Trevor Darrell. 2017. Curiosity-Driven Exploration by Self-Supervised Prediction. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 488–489. <https://doi.org/10.1109/CVPRW.2017.70>
- [32] Margaret Robertson. 2010. *Can't play, won't play*. Retrieved September 18, 2019 from <https://www.kotaku.com.au/2010/11/cant-play-wont-play/>
- [33] Richard M. Ryan and Edward L. Deci. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25, 1 (2000), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- [34] Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (2006), 344–360. <https://doi.org/10.1007/s11031-006-9051-8>
- [35] Nikolay Savinov, Anton Raichuk, Damien Vincent, Raphael Marinier, Marc Pollefeys, Timothy Lillicrap, and Sylvain Gelly. 2019. Episodic Curiosity through Reachability. In *International Conference on Learning Representations*.
- [36] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 7587 (2016), 484–489. <https://doi.org/10.1038/nature16961>
- [37] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, and Demis Hassabis. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144. <https://doi.org/10.1126/science.aar6404>
- [38] Satinder P. Singh, Andrew G. Barto, and Nuttapon Chentanez. 2005. Intrinsically Motivated Reinforcement Learning. In *Advances in Neural Information Processing Systems 17*, L. K. Saul, Y. Weiss, and L. Bottou (Eds.). MIT Press, 1281–1288.
- [39] Burrhus F. Skinner. 1953. Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist* 8, 2 (1953), 69–78. <https://doi.org/10.1037/h0054118>
- [40] StackOverflow. 2019. *Badges*. Retrieved September 18, 2019 from <https://stackoverflow.com/help/badges>
- [41] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement learning: An introduction*. MIT press, Cambridge, MA.
- [42] Penelope Sweetser, Daniel Johnson, Peta Wyeth, Aiman Anwar, Yan Meng, and Anne Ozdowska. 2017. GameFlow in Different Game Genres and Platforms. *Comput. Entertain.* 15, 3, Article 1 (April 2017), 24 pages. <https://doi.org/10.1145/3034780>
- [43] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: A Model for Evaluating Player Enjoyment in Games. *Comput. Entertain.* 3, 3 (July 2005), 3. <https://doi.org/10.1145/1077246.1077253>
- [44] Edward L. Thorndike. 1898. "Animal Intelligence". *Nature* 58, 1504 (1898), 390–390. <https://doi.org/10.1038/058390b0>
- [45] Gustavo F. Tondello, Dennis L. Kappen, Marim Ganaba, and Lennart E. Nacke. 2019. Gameful Design Heuristics: A Gamification Inspection Tool. In *Human-Computer Interaction. Perspectives on Design*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 224–244.
- [46] John B. Watson. 1913. Psychology as the behaviorist views it. *Psychological review* 20, 2 (1913), 158–177. <https://doi.org/10.1037/h0074428>

- [47] Robert W. White. 1959. Motivation reconsidered: The concept of competence. *Psychological review* 66, 5 (1959), 297–333. <https://doi.org/10.1037/h0040934>
- [48] Wikipedia. 2019. *Operant conditioning chamber*. Retrieved September 18, 2019 from https://en.wikipedia.org/wiki/Operant_conditioning_chamber
- [49] Nannan Xi and Juho Hamari. 2019. Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management* 46 (2019), 210–221. <https://doi.org/10.1016/j.ijinfomgt.2018.12.002>
- [50] José P. Zagal, Staffan Björk, and Chris Lewis. 2013. Dark patterns in the design of game. In *Foundations of Digital Games 2013 (FDG 2013)*. 39–46.
- [51] Yufan Zhao. 2009. *Reinforcement Learning Design for Cancer Clinical Trials*. Ph.D. Dissertation. University of North Carolina, Chapel Hill, NC.
- [52] Gabe Zichermann and Christopher Cunningham. 2011. *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O'Reilly, Sebastopol.
- [53] Nynke Zwart. 2018. *Assessing the Effect of Gamification on various aspects of Sport Exercises through an Augmented Reality Exertion Interface (Bachelor's Thesis)*. Technical Report. Radboud Universit, Nijmegen, Netherlands.