Reward, Learning and Games

Authors: Paul Howard-Jones\textsuperscript{1} and Tim Jay\textsuperscript{2}

\textsuperscript{1}Graduate School of Education, University of Bristol, United Kingdom
\textsuperscript{2}Department of Teacher Education, Sheffield Hallam University, United Kingdom

Correspondence:
Paul Howard-Jones,
Graduate School of Education,
University of Bristol,
BS8 1JA,
UK
Email: paul.howard-jones@bris.ac.uk
Tel. + 44 (0) 117 331 4496
Fax. +44 (0) 117 925 1537

Word Count: 3415 (not including references)
Abstract
The link between reward and learning has chiefly been studied scientifically in the context of reinforcement learning. This type of learning, which relies upon midbrain dopaminergic response, differs greatly from the learning valued by educators, which typically involves declarative memory formation. However, with recent insights regarding the modulation of hippocampal function by midbrain dopamine, scientific understanding of the midbrain response to reward may be becoming more relevant to education. Here, we consider the potential for our current understanding of reward to inform educational learning, and consider its implications for game-like interventions in the classroom.

Introduction
Reward and education – the search for a relation between reward and educational achievement

Teachers regularly use incentives to engage their pupils, but researchers have had difficulty in developing evidence-based insight to support this practice. Partly, the difficulties derive from identifying clear educational benefits of offering rewards. Some effects of reward on memory were reported in early studies [1-3], whereas other investigations have been inconclusive [4]. Indeed, even the effect of rewards on general performance motivation have been called into question [5]. Loftus [6] reported effects of reward on encoding and suggested these arise from enhanced attention, rather than the reward itself. By showing that reward-associated items were both more remembered and also fixated on more frequently during encoding, Loftus [6] showed that rewards may focus the attention of individuals more on some stimuli than others, which may make them more salient and so memorable. Rewards over longer time scales have also shown unpromising results, with no positive effects arising when 15-16 year olds were offered financial incentives and “tickets to events” in return for raising their national examination results [7] and negative effects reported for self-regulated learning [8]. The mixed nature of these findings highlights the need for a more sophisticated understanding of reward and learning, to generate more secure principles and hypotheses to test.

In this paper, we focus upon the potential implications for game-based learning of the known effects of reward on attention and declarative memory formation. Declarative memory formation has a special significance in education, possibly because knowledge that can be made explicit is most conveniently assessable [9]. We begin by considering reward explanations of reinforcement learning behaviour in dynamic environments that require actions to optimise reward and so have a modest resemblance to popular gaming environments. We consider links between reward and attention, and how reward learning processes may explain the putative benefits of gaming environments beyond declarative memory. Finally, we review current efforts to implement this understanding in the classroom.

We should emphasise from the outset that space constraints do not afford a full review of current concepts and understanding of the relationship between reward and memory, but instead we focus on the potential relation between emerging understanding in this area and education. We hope our article outlines the current uncertainties in developing a “bridge” between neuroscience and education in this area, and may provide a useful prompt for future investigations. (For an excellent review of how reward motivation influences memory, with an emphasis on declarative memory, the reader is directed to Miendlarzewska, Bavelier and Schwarz [10].)
Motivation, reinforcement learning and midbrain dopamine

Discussions aimed at improving dialogue between neuroscience and education have identified reward as an area where new scientific insights might inform educational understanding and improve classroom practice. However, it is important to note that the meaning of terms such as ‘reward’ differ greatly between its usage in education and its meaning in cognitive neuroscience. In an educational context, rewards are usually material offerings or social symbols of recognition intended to influence behaviour, and motivation can include the desire to reach long-term goals. In cognitive neuroscience, as in the present article, we may consider reward to include both material and social reinforcers, and motivation as being associated with positive and negative affective states or stimuli, and more often with short-term behaviours that may include approach or withdrawal from stimuli [11]. Approach motivation associated with positive stimuli is the phenomenon closest to the educational use of the term ‘motivation’ (and it is in this sense that the term motivation will be used below). These differences in the use of language are augmented by those characterising different sub-fields within the scientific cognition-motivation literature [12].

Approach motivation to a positive stimulus is coded by uptake of dopamine from the midbrain to a region called the ventral striatum and, in particular, a small nucleus of densely populated neurons within this region called the nucleus accumbens. This midbrain dopaminergic activity has been shown to increase when humans are exposed to a variety of pleasures including food[13], money[14], and computer games [15]. This short-term and visceral type of motivation may have much to do with our day-to-day desire to solve problems that reap immediate benefit, but probably less to do with less immediately gratifying prospects, such as the goal of pursuing a difficult programme of study in order to further our professional or academic profile. Nevertheless, it appears a reasonable hypothesis that moment-to-moment visceral motivations do have influence on children’s learning in the classroom.

There is much we do not understand about the mechanisms by which ‘off the shelf’ games influence the reward system. Studies in the context of putative associations between computer games, addiction and the reward system have compared action-based games involving rewards [e.g. 15] to studies of DA in reinforcement learning (RL), since these games involve learning how to take actions that optimize reward in a dynamic environment. Rewarded action has been proposed as a potentially important factor in the potential of video games to influence cognitive function. Studies of reinforcement learning may, therefore, provide insight into DA function in games, although an important caveat here is that, although accepted as central, the exact role of DA within RL (and reward-related processes more broadly) remains controversial. RL is a type of learning shared by many animals and considered to support, for example, foraging among natural food sources [16]. Neural processes thought to underlie RL implicate ventral tegmental area (VTA) efferent projections that release DA to a broad range of structures such as prefrontal cortex (PFC), nucleus accumbens (NAc), amygdala, and hippocampus [17]. This dopaminergic pathway (the meso-cortico-limbic circuit) is thought to play a key role in reinforcing rewarding behaviour. When a ‘better-than-expected’ reward (positive prediction error) is signalled by activation of DA neurons, the resulting cue-reward learned associations produce a change in reward-seeking behaviour [18] helping to optimise our behaviour in a changing environment. In reinforcement learning, it is phasic DA release (i.e. a short term pulse) that is considered to code prediction error and provide this important learning signal. However, the human data for this model is somewhat circumstantial due to ethical difficulties in directly measuring DA.
transmission and the reliance, instead, on a BOLD neuroimaging signal as a proxy [19]. A first attempt to directly measure DA release in relation to prediction error suggests this may be mediated in a more complex manner by context than originally assumed [20]. Also, both phasic and tonic DA activity appear to be involved in motivational state [21], and both contribute to the extra-cellular DA levels that regulate conditioned responding.

The association between prediction error and reinforcement learning emphasises the role of recent prior experience on phasic dopamine response, in terms of the expected value of previous rewards. This expected value takes account of both the possible reward magnitude and its probability. Primate studies suggest the variance (or uncertainty) in this probability may influence tonic levels of dopamine, producing a sustained ramping between a cue that a reward may be arriving and delivery of the reward [22]. This effect appears maximal at a reward uncertainty of 50%. Evidence of a similar relationship between dopamine and reward uncertainty has also been reported in two human studies using fMRI [23,24]. This response to uncertainty has been used to explain our attraction to games of chance [25], although many other factors pertinent to playing video games, such as novelty [26] and social interaction [27], are also likely to play a role in determining midbrain DA release.

**Reward and attention**

Although prediction error forms an important part of associative learning theories involving reward, the neurobiological mechanisms by which the DA coding of prediction error contributes to this learning are not well understood [28], but are thought to involve enhanced attention to poorly predicted (or ‘surprising’) outcomes. The role of midbrain DA release in orienting attention has some support from animal studies [e.g. 29], while the role of midbrain DA in attention has generated most interest in dopamine-deficit theories of ADHD, where the failure to develop anticipatory dopamine release is thought to result in a lack of dopamine cell activity in response to attending [30]. In active paradigms, such as naturalistic scenarios involving action selection, saccades may have a bidirectional relationship with the task. They can be influenced by the nature and values of the ongoing actions and may influence the task by selecting sensory information that most strongly impacts on the observer’s actions [see 31 for review].

More straightforwardly, through learning of stimulus–reward (Pavlovian) associations (or reward learning), stimuli that are otherwise neutral to the task in hand can become imbued with value and capture attention powerfully and persistently. This provides a basis for considering how experience with a gaming environment can, irrespective of the current state of play, continue to capture attention. In other words, it can provide a scientific rationale for setting a training exercise with a gaming environment to support engagement with the training (e.g. [32,33]), irrespective of the moment-by-moment changes in the availability of rewards within the game. The possibility of increased attention broadens the potential benefits of manipulating reward to include the many types of educational learning that rely less on declarative memory formation (e.g. reasoning skills, creativity etc). Such attentional effects may also help explain why video games are reported to benefit their players in many domains typically considered as distinct, including vision, cognitive function, decision making, reaction time and speed-accuracy trade-off, attention and causality [see 34 for review]. These highly engaging games offer schedules of reward for performing many correct responses per unit of time. Increased activities in regions
targeted by dopaminergic neurons, including the ventral striatum, have been reported in fMRI studies of videogame play [35], [15] (but see [36], [37] regarding methodological concerns). This involvement of the striatum in these studies, and particularly its ventral regions, appears to suggest implication of reward processes and the neuromodulator dopamine in the reported benefits action video gaming. However, other neuromodulators, notably acetylcholine, have featured more strongly in some explanations of video game benefits [34] and the benefits themselves have not always been consistently demonstrated [e.g. 38], with some reporting of negative effects [39].

**Reward and Declarative Memory**

Midbrain dopaminergic activity also appears to influence declarative memory formation [40], which is an ability of great interest to educators [9]. In a study of adults incentivised by money to remember visual scenes, Adcock et al. [41] reported that anticipatory activation in the ventral tegmental area, nucleus accumbens, and hippocampus predicted remembering and was greater for higher rewards, and activity in the hippocampus and ventral tegmental area correlated with participants’ enhanced long-term memory for the subsequent scene. The hippocampus is part of a medial temporal lobe system necessary for the formation and consolidation of declarative memory in tasks such as the recall of facts [42-44], but also for the transfer of learned rules to novel situations [45] (however, it is less necessary for many non-declarative types of long-term memory such as skill learning and habit formation which are also of educational significance [46]). These findings support the hypothesis that reward motivation promotes declarative memory formation via dopamine release in the hippocampus just before learning [47]. The mesolimbic dopamine system is also strongly interconnected with serotonergic neurons, but serotonin does not appear involved in hippocampal memory formation [48].

Long-term potentiation (LTP), which remains the most widely accepted model for learning and memory, is usually separated into an early and late phase. The early phase is considered to comprise changes in synaptic strength, and associated retention, over a scale of minutes and perhaps hours. The later phase is considered responsible for making memories more permanent, through processes of synaptic plasticity involving protein synthesis [49]. Animal studies of long-term potentiation (LTP) in the hippocampus (HC) show that, in addition to well-known Hebbian conditions (presynaptic input and strong postsynaptic depolarization), late LTP requires action of the neurotransmitter dopamine for successful encoding [50]. D1/D5 receptors appear to gate hippocampal long-term plasticity in the mammalian brain. They play a critical role in the encoding and storage of information in the HC, and their activation in response to reward leads to increased HC processing and minimization of mismatch detection, so favouring storage [51]. Such studies reveal dopamine is less involved in the processes of early LTP. In rodents, for example, early memory is unaffected by antagonism of D1/D5 receptors [52]. In a human population suffering an age-related loss of dopamine neurons, treatment with the dopamine precursor levodopa led to a similar pattern [53]. That is, improvements in episodic memory were more robust at delayed, rather than early, testing suggesting a retroactive effect of midbrain DA on human memory supported by other human behavioural [54] and imaging studies [55-57]. In terms of the classroom, this suggests effects arising from attempts to stimulate midbrain DA release may not be observed immediately, but may be more evident a few days following the learning session.
An inverted U curve for dose has also been observed in a study of rewarded recognition involving younger participants [58]. In essence, such a curve predicts that a small DA increase may improve performance while a large one may reduce it. The curve may also explain why only a modest improvement in memory has been reported when a monetary incentive is doubled [59]. Since the baseline of individuals may be positioned at different points along this curve, individual differences may exist in terms of potential benefits of reward for memory enhancement. Genetic sources of individual variability include genes affecting dopamine transmission. Studies of correlation between genetically-determined dopamine availability and memory processes confirm the role of dopaminergic hippocampal processes in encoding motivational events [60,61]. These genes explain too little variance to provide a feasible basis for tailoring individual approaches to learners’ genetic profiles. They do, however, suggest potential value in including genetic information in educational interventions involving, for example, novel reward schedules, in order to improve detection of main effects.

This connection between dopaminergic activity and memory suggests estimates of the brain’s response to reward may provide a more accurate predictor of memory performance than the rewards themselves and help explain why behavioral studies that focus on the absolute value of the reward have produced inconsistent results. Whether the reward-memory effect requires attention as a mediator or involves a more direct process, the link between them, of course, remains of strong educational interest.

**Reward and educational learning games**

Of interest for educators is whether learners would benefit if design principles can be identified for ‘gamifying’ learning with reward schedules leading to improvements in declarative memory formation. Above, there have been several references to video games which, although known to increase mid-brain dopamine, do not provide a suitable vehicle for identifying and testing such potential design principles. Addressing this gap, Howard-Jones at al. [62] combined educational learning with a four-armed bandit task for which a neurocomputational model of changes in phasic dopamine had been validated [e.g. 16]. Adults were asked to play a quiz game in which they could win the points revealed when selecting one of four boxes, if they answered a subsequent multiple-choice quiz question correctly. Changes in phasic dopamine could be estimated on the basis of prediction error associated with each selection, and this measure predicted correct retrieval of information more effectively than the stakes themselves. This effect was observed for recall and may, therefore, be explained by attentional enhancement.

Anatomical analysis of the hippocampus suggests tonic dopamine levels may be a stronger predictor of hippocampal encoding function than phasic levels [40]. One potential way to manipulate tonic dopaminergic response is by introducing uncertainty. In essence, and in contrast to being offered a reward, the student is offered the chance to win a reward (e.g. points) according to some arbitrary mechanism (e.g. a wheel of fortune, or toss of a coin) in return for successfully completing an academic challenge (see Fig 1). For many educators, this is a counter-intuitive approach to increasing engagement. Teachers are often advised to provide reward consistency as means to ensure motivation, in the belief that disruption of the relationship between achievement and reward will be deemed unfair by students. However, an analysis of children’s dialogue when competing for uncertain rewards did not identify any such concerns. Rather motivational ‘sport-talk’ was reported, in which losses were attributed to bad luck and gains celebrated as affirmations of skill [62]. The same study reported additional emotional
response to a similar game when adults were responding to questions for uncertain, as opposed to certain, rewards, which may suggest additional emotional engagement with the task. A recent fMRI study compared the effects of interleaving short 28 second periods of adult study with exemplar questions and answers (study only), with answering questions for points (self-quizzing) and answering questions for escalating uncertain rewards determined by a wheel of fortune (game-based)[63]. As the tasks became more game-like (study-only->self-quizzing->game-based), so greater learning and self-reported engagement were achieved, with individual differences in learning gain predicted by the extent of deactivation of the default mode network, associated with mind-wandering[64]. However, the inclusion of competition prevents such gains from being wholly attributed to the manipulation of reward schedule.

The arrival of a sport-like environment when rewards are uncertain may also help combat the potential negative effects of reward motivation. The anticipation of reward can become overly stressful, neutralising its benefits and reducing striatal activity [65], possibly through mediation by serotonin [66]. Rewards that are closely linked to achievement can be assumed to reflect more strongly on self- and social esteem. This may explain why children, when given the choice, choose academic problems they are considerably more than 50% likely to solve successfully. Clifford and Chou, in a study of 4th graders, found these students appeared most comfortable on tasks they felt 79% to 96%, confident with [67]. This draws attention to how students may experience different types of uncertainty within a school environment that can impact in different ways upon their behavior and, thereby, their achievement.

Ozcelik et al. [68] tested more directly the hypothesis that uncertain reward might increase learning in contexts relevant to education. In their study, higher education students on a software engineering course were learning about databases using a computer-based game that awarded points for correct answers. The researchers randomly allocated 140 higher education students to two groups, both of whom experienced learning about database concepts in a virtual environment, which including responding to questions in return for points. One group gained points for correct answers, while the other group gained a number points determined by chance. Students in the uncertain condition achieved greater improvements in their performance. Further, researchers used statistical path analysis to demonstrate that improved motivation was a causal agent in this effect. It should be noted this study tested students on their ability to apply their knowledge to novel problems, demonstrating the effects of uncertain reward on a level of learning that was deeper than mere factual recall. In another classroom intervention (N=449), one group of 9-10 year old classes received periodic multiple choice questions during 90 minute workshops which required small teams of pupils to assign tokens to the answer(s) they believed to be correct [69]. In this “risk” group, twice the number of tokens assigned to the correct answer were returned to the group and those assigned to incorrect answers were lost. This group of classes achieved higher assessment scores at the end of their workshops than the group who were provided with a fixed number of tokens for a correct answer. Interestingly, in line with current understanding of dopamine’s retroactive influence, learning enhancement was evident only after a delay of one week, and was not immediately detectable. However, such quasi-experimental classroom interventions are notoriously difficult to control, and the authors urge caution when interpreting their results, noting that discussion of content during breaks could not be monitored and that this may have contributed to learning outcomes. They also suggest that excited discussion about the gaming context might explain differences in the two learning conditions only becoming noticeable after a week. Participants in the Ozcelik et al study were only tested
immediately following their learning experience, so this data cannot inform on the likelihood of retroactive classroom effects. Large-scale trials (N=10,000) exploring the effects of offering uncertain rewards to 12-13 year olds in science classes are now beginning in the UK [70].

Although the potential educational utility of uncertain rewards is promising, there are many scientific questions that need addressing before its theoretical basis is secure. Most work on dopamine neuron activity has been conducted on anesthetized rats, with some awake non-human primate studies and a rare number of awake human studies involving Parkinson’s patients [e.g. 20]. Differences in the approaches used make it difficult to compare the amount of bursting and characteristics of each burst across non-human primates and rodents, with much less known about potential differences between human and non-human processes underlying reward-cognition effects.

Since there is evidence for human activity associated with midbrain dopamine release varying with age [71], gender [35], genetic background [72], stress [73] and traits such as optimism [74], extroversion [75], risk aversion [24] and impulsivity [76,77], what sort of individual differences might exist in the response of individual students to educational interventions involving uncertain reward? Human processes are also likely to be strongly influenced by context. For example, in Fig. 1, how might the uncertainty associated with completing the educational task influence dopamine neuron activity? Additionally, and perhaps most daunting in scientific terms, practical interventions may involve classroom environments that are highly social. It is known, for example, that we can experience an egocentric prediction error that is coded by a phasic DA response when observe our competitor unexpectedly failing [78]. What implications should we expect for our educational learning when observing our competitor, and how are the processes influenced by the social discourse and cultural contexts of the classroom?

Finally, it is worth noting that some other features that can be, and often are, present in educational games also have the potential to enhance memory through the action of midbrain dopamine. Novelty does not often appear alongside reward in the educational discourse, but similarities in the relation of novelty and reward to memory often cause them to be considered together in the modern scientific literature. Novelty and novel contexts, like reward, engage midbrain modulation of the hippocampus and triggering of VTA activity, and studies again implicate D1/D5 receptors in gating hippocampal long-term plasticity, so enhancing long-term memory for novel events [47]. As with reward uncertainty, novelty may influence tonic dopamine response [59] and may have also implications for theorising learning games, since the effects of novelty on human memory include the types background contextual novelty that are typical in games [79]. Exploring novel environments for words improves memory for them [80] or unrelated novel but educational event [81]. Although less studied, the simple act of choosing/agency improves memory [82], also through striatal-hippocampal interaction.

Conclusions and future challenges

The emerging cognitive neuroscience of reward, memory and their interrelation promises a new perspective on the potential role of reward in education, and particularly in the development of educational games. Dopamine release from the midbrain is thought to play an important role in learning to associate rewards and actions in reinforcement learning and reward learning, and such release can also enhance declarative memory formation. Candidate processes for this enhancement include attentional orienting (which may also help explain the putative cognitive benefits of video games) and/or, more directly, through gating hippocampal function.
Currently, the science required to inform the manipulation of reward schedules for educational benefit is very incomplete. There are some important gaps in our knowledge, not least regarding the processes by which production of midbrain dopamine influences memory function. However, the existing evidence can already provide stimulus for discussing, formulating and testing new interventions. The offering of uncertain rewards, for example, in order to raise tonic dopamine response and so modulate hippocampal function, presents an easily realizable application of current understanding.

We believe the types of insight reviewed here can contribute to an understanding of how games can support learning, beyond the popular notion of “making learning fun”. While educational practice emphasizes notions such as “reward consistency” [83], this brief review has highlighted issues of reward scheduling and timescale of effects that challenge current educational perspectives. Such insights can inform on how reward might be scheduled to stimulate midbrain DA, how presentation of learning content might be sequenced to exploit such processes, and when associated effects on learning might be observable.

We have reviewed some preliminary, though limited evidence of the educational effectiveness of new approaches involving uncertain rewards. Such interventions, if carefully planned, may also contribute to extending both the educational and scientific knowledge base. Attempts at classroom implementation are likely to highlight many issues for successful transfer of knowledge between neuroscience and education that remain mostly unexplored. To address these issues, studies of reward system response will be needed involving tasks that are more educationally aligned, together with investigation of individual differences in reward-learning relationships within educational contexts. Research that seeks to further scientific understanding while providing insight into educational practice will require a transdisciplinary approach [84] involving collaboration and dialogue across these two diverse fields. The popularity of press articles about dopamine has caused it to be dubbed the “the media’s neurotransmitter of choice” [85], and it is already becoming involved with the types of neuromyth that detract from educational practice and attitudes [86]. A transdisciplinary approach will not just aid generation of educationally-relevant scientific insight, but also help construct and transmit messages to educators that ensure its appropriate application.

**Figure Legend**

Fig. 1 Current understanding suggests that, when there is uncertainty about an upcoming reward, there is a slow and sustained ramping of dopamine neuron activity between the cue predicting that a reward may (or may not) arrive and the revealing of outcome. Since this putatively tonic dopamine response is linked to greater motivation and also to modulation of hippocampal encoding function, it might potentially be harnessed for educational purposes. The diagram suggests one way in this might be achieved. In (a) an educational task (e.g. a question) cues the possibility of an uncertain reward (e.g. a number of points determined by the spin of a wheel of fortune). Ramping of dopamine neuron activity in response to uncertain reward during responding (which may be “scaffolded” by the teacher or by access to learning resources) should support learning. This should be the case if the student’s answer is correct and they are lucky with respect to receiving a
reward (a), or whether they are correct and unlucky or (b) simply offer an incorrect answer and thus receive no reward (c). However, see main text for several assumptions that underlie this suggested type of intervention.

References


51. Hansen N, Manahan-Vaughan D: Dopamine D1/D5 Receptors Mediate Informational Saliency that Promotes Persistent Hippocampal Long-Term Plasticity. Cerebral Cortex 2014, 24:845-858.


58. Apitz T, Bunzeck N: Dopamine Controls the Neural Dynamics of Memory Signals and Retrieval Accuracy. Neuropsychopharmacology 2013, 38:2409-2417.


83. Steer A: Learning Behaviour Principles and Practice – What Works in Schools, Section 2 of the report of the Practitioners’ on School Behaviour and Discipline chaired by Alan Steer

Edited by. London, UK: Department for Education and Skills (DfES); 2006. [DfES (Series Editor):


Notes on References:


This multiply-authored paper provides something approaching a consensus on the state of efforts to identify reward-cognition processes and the continuing scientific challenges in this area.


This article reports on an educational intervention in which compares the effects of uncertain and certain rewards on educational learning by a sample of undergraduate students. Analysis of the data links beneficial effects of uncertain reward to increased motivation.


This article explores the practical realities of trying to introduce neuroscientific concepts about reward into educational games in the classroom. It reports on a a design-based research approach to the development of learning game technology informed by neuroscience to judiciously interrelate insights associated with diverse theoretical perspectives (e.g., neuroscientific and pedagogical).


This is a helpful paper for understanding sources of potential variation in dopamine neuron activity. Reflecting current literature, it draws chiefly on rodent and non-human primate studies but links to human data in a critical manner that helps inform about the current gaps in understanding.


This is a comprehensive review of how dopamine may boost the formation of declarative memory for rewarded information but also control the generalization of reward values to related representations.


Provides arguments for a transdisciplinary approach to research and to constructing messages for communicating across neuroscience and education, including consideration of how dopamine is already beginning to feature in unhelpful educational neuromyth.

This article reviews potential mechanisms by which DA transmission in frontostriatal systems modulates associative learning, cognitive flexibility, and motivation.