Toward a Science of Learning Games

Paul Howard-Jones[1], Skevi Demetriou[1], Rafal Bogacz[2], Jee H. Yoo[2], and Ute Leonards[3]

ABSTRACT—Reinforcement learning involves a tight coupling of reward-associated behavior and a type of learning that is very different from that promoted by education. However, the emerging understanding of its underlying processes may help derive principles for effective learning games that have, until now, been elusive. This article first reviews findings from cognitive neuroscience and psychology to provide insight into the motivating role of uncertain reward in games, including educational games. Then, a short experiment is reported to illustrate the potential of reward-based neurocomputational models of behavior in the understanding and development of effective learning games. In this study, a reward-based model of behavior is shown to predict recall of newly learned information during a simple learning game.

THE ELUSIVE THEORY OF "EDUTAINMENT"

Games (Bergin, 1999; Gee, 2003) are a potential source of inspiration for teachers wishing to engage their learners, and there have been many attempts to develop experiences that combine education with the entertainment of games to generate so-called "edutainment." Yet, attempts to find the critical ingredients for such engagement have produced a bewildering array of candidates. Malone (1981) identified components of fantasy, challenge, and curiosity. Johnson (2005) drew attention to how most computer games now require no initial knowledge or manual. Garris, Ahlers, and Driskell (2002) emphasized the importance of feedback responses, reflection, and active involvement. The reason given by most gamers for pursuing their passion is fun yet some academics dismiss this as a "red herring" for educators because, they claim, children find learning fun enough already

[1]Graduate School of Education, University of Bristol
[2]Department of Computer Science, University of Bristol
[3]Department of Experimental Psychology, University of Bristol

Address correspondence to Paul Howard-Jones, 35 Berkeley Square, Bristol BS8 IJA, United Kingdom; e-mail: paul.howard-jones@bris.ac.uk.

Note: Correction added on 14 March 2011 after first publication online on 21 February 2011. The co-authors, Skevi Demetriou, Rafal Bogacz, Jee H. Yoo and Ute Leonards, and their affiliations, were erroneously missed out. The error has been corrected in this version of the article.

(Kirriemuir & McFarlane, 2004). The same authors suggest simplicity and repetition are the reasons why educational games often fail. And yet, it can be observed that many children find simple and repetitive games such as Tetris more absorbing than a well-planned lesson. A lack of established understanding may help explain the difficulties encountered by those attempting to combine learning and gaming, which have led some commentators to write the "only consensus in this whirlwind of activity seems to be that educational games are something of a failure" (Zimmerman & Fortugno, 2005). If we wish to imbue learning with the excitement experienced by computer game players, then we need to understand more about the processes linking cognition, emotion, and motivation. As pointed out elsewhere in this journal, neuroscience encourages us to consider these concepts as closely intertwined (Fisher, Marshall, & Nanayakkara, 2009; Immordino-Yang & Damasio, 2007) and potentially offers new ways of theorizing the motivated learner.

THE COGNITIVE NEUROSCIENCE OF MOTIVATION

When reviewing the cognitive neuroscience of reward and motivation, note that there are differences in how terms are applied in neuroscience compared with common usage. For example, in cognitive neuroscience, reward usually refers to short-term incentives that reinforce behavior. The relationship between motivation and learning in neuroscience has been studied chiefly in the context of reinforcement learning (Wise, 2004), a type of learning thought to support foraging among natural food sources (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006) and an ability we share with many other animals. Here, approach motivation is considered as the incentive to approach or the extent to which we want something. This appears closely related to the uptake of dopamine in a midbrain region called the ventral striatum and, in particular, a small nucleus of densely populated neurons within this region called the nucleus accumbens. Midbrain dopaminergic activity has been shown to increase when we are exposed to a variety of pleasures including food (Farooqi et al., 2007), money (Knutson, Adams, Fong, & Hommer, 2001), and computer games (Koepp et al., 1988). This visceral type of motivation may have less to do with the higher-order thinking processes that appear to motivate us toward other activities that are less gratifying in the short term, such as pursuing a difficult

course of professional development. Nevertheless, as has been pointed out by Van Geert and Steenbeek (2008), short-term motivational processes may have a powerful influence on long-term outcomes.

The value, or size, of an anticipated reward influences the motivational signal it produces in the striatum. Interestingly, this is very much influenced by context. In a study comparing the size of anticipated monetary reward with the response of the reward system, the maximum signal corresponded to the maximum available reward in that context (Nieuwenhuis et al., 2005). For example, one would expect the dopamine released by anticipating a top prize of $10 in one game is the same as anticipating the top prize of $100 in another.

In reinforcement learning, learning may be defined as the change in expected value of different actions. That is a far cry from the complex types of knowledge and understanding that educators seek to provide. There are, indeed, fundamental differences in the conceptualization of learning within neuroscience and within education (Howard-Jones, 2008). However, because reinforcement learning is intimately entangled with the visceral motivation system just described, understanding this type of learning may provide insights into how we engage with certain types of task. When we forage, we are most likely to choose the source with the highest expected reward and, when we experience the outcome of our choice, we adjust our information about the source in relation to the prediction error (PE)—or how much the actual outcome exceeds the expected outcome. In a crude sense, PE is the extent to which an outcome is deserving of "happy surprise." This error is coded by midbrain dopaminergic activity (Schultz, Dayan, & Montague, 1997). As an expected outcome is determined by recent historical contexts, this relationship between the outcome of an event and the signal it generates in the brain (correlated with PE) is another demonstration of the effect of context on our response to reward. (For example, the PE signal generated by the outcome of an action is higher if that action has generally produced poor results in the recent past. So, a positive outcome of an action will generate less PE if a similar result has already been recently experienced.) Reinforcement learning is tightly bound up with midbrain dopaminergic activity and, therefore, with very basic motivation processes.

Perhaps of more interest to educators, the uncertainty of an outcome also influences the brain's response to reward. In a study of the primate brain (Fiorillo, Tobler, & Schultz, 2003), researchers created a certain reward by frequently presenting a stimulus with a subsequent reward on 100% of occasions. Other stimuli were associated with rewards that arrived with less certainty. They showed that a stimulus associated with the imminent arrival of a 100% certain reward generated a similar spike of dopamine activity as the reward itself arriving entirely unexpectedly (i.e., 0% certainty). The actual arrival of the certain reward produced little effect at all, because

the preceding stimulus had made this a wholly predictable event. However, when a stimulus was presented that had been associated with reward on only 50% of past occasions, the stimulus generated a similarly sized spike of dopamine but then the dopamine began to ramp up again, reaching another maximum at the moment when the reward might (or might not) appear. This anticipatory ramp of dopamine, together with the previous spike, resulted in uncertain reward generating more dopamine overall than either 100% certain or wholly unexpected reward. In other words, uncertain reward appears to increase the type of dopaminergic response that has been linked to motivation (Berridge & Robinson, 1998). This effect of reward uncertainty has been suggested as an explanation of why humans are so attracted to gambling and games involving chance (Shizgal & Arvanitogiannis, 2003).

UNCERTAINTY, HUMAN BEHAVIOR, AND CONTEXT

Uncertain reward is a defining characteristic of games (Caillois, 1961; Hong et al., 2009), including computer games (Juul, 2003), and has been identified by some as an important and pleasurable aspect of their challenge (e.g., Loftus & Loftus, 1983). The scientific studies of uncertainty reviewed above may add to our understanding of how uncertainty contributes to the attraction of games, but its influence may vary with the many forms that uncertainty takes. All games include elements of uncertainty that derive from some chance factor, and experiments have shown moderate risk taking in such contexts, at around 50% success probability, heightens motivation (e.g., Atkinson, 1957). Chance-based uncertainty, therefore, appears to enhance motivation in the way that the neuroscience might predict and its attractiveness is borne out by its inclusion in the simplest traditional games. Some games (e.g., snakes and ladders) appear to rely almost entirely on this feature to engage their players, because each player's progress is completely determined by the roll of a die. However, partial knowledge about one's own abilities and those of one's competitors also contribute to uncertainty in a game, and this type of certainty can become associated with a more complex set of meanings than pure chance. A clearer example of this can be found in educational institutions, where students may be uncertain about whether they can achieve good results but can expect rewards if they do so, whether it is a gold star, a high mark, or a word of praise. In terms of their immediacy and potential to reinforce behavior, these types of reward share some similarity with the rewards often studied in cognitive neuroscience. Here, however, educators make strenuous and explicit efforts to provide reward solely on the basis of effort and ability and the consistency of this reward–virtue relationship is considered crucial to maintaining motivation and a sense of fairness (Office for Standards in Education, 2001). Uncertainty in most tasks encountered in school,

therefore, may feel less comfortable than in contexts where the role of chance is acknowledged as a strong determinant, because failure in school contexts usually has implications for social- and self-esteem. This is illustrated by students seeking greater challenge when the same task is presented as a game rather than a learning experience (Clifford & Chou, 1991). For tasks seen as educational, these students appeared most comfortable on tasks they felt about 88% confident with. Of course, such tasks can still help students perfect their skills, but this diminishment of the comfort zone for uncertainty also reduces the likelihood of surprise and may diminish the type of dopaminergic response reviewed above.

A possible solution to this problem is to introduce chance-based gaming elements into educational tasks. This can increase uncertainty without threatening esteem and should, therefore, increase their attractiveness to learners. To test this, a class of 11–12-year-olds was asked to practice their mental mathematics by playing a purpose-built computer game (Howard-Jones & Demetriou, 2009). The students answered true or false to 30 mathematical statements with the aim of maximizing their score. However, before seeing each question, they had to decide whether to ask it from Mr. Certain or Mr. Uncertain. If a student answered correctly, he/she would receive one point from Mr. Certain and either zero or two points from Mr. Uncertain, depending on the toss of an animated coin. As predicted, there was a preference for Mr. Uncertain, and this preference increased over the game. Our educational common sense is that, if a student has the correct answer, they should receive a mark. Yet here, students showed a marked preference for earning the opportunity for more uncertain reward.

The addition of chance-based uncertainty to an educational task should not be dismissed as a "sugar coating on the bitter pill of learning," because it can heighten the emotional response to the learning itself. In a study of adults playing a purpose-built computer, two dice were thrown and the resulting points could be kept if a subsequent question was answered correctly (Howard-Jones & Demetriou, 2009). Participants tried the game in two conditions, with and without the gaming element (arranged by keeping the dice fixed). Physiological measurement showed the emotional response to answering questions was greater when the gaming element was enabled and rewards were uncertain (Howard-Jones & Demetriou, 2009).

REWARD UNCERTAINTY AND THE SUBVERSION OF LEARNING DISCOURSE

It is certainly not suggested here that chance-based reward uncertainty is the only factor that serves to make computer games enjoyable. As with educational tasks, popular computer games involve many sources of uncertainty deriving from chance but they also include other, more socially sensitive factors such as the ability of the participants. Shooting for a target in a game usually requires good visual attention and motor abilities as well as a little good luck. However, within the social contexts of education, the acknowledged role of chance-based uncertainty in a task may be a critical factor influencing how students interact with it. And, because it is a factor that educators usually strive to eliminate, learning and gaming tasks can usually be distinguished according to whether it is present. This may explain why a modicum of explicit chance-based uncertainty added to a learning task can transform the discourse around the task into a type more commonly found in a football game than in a classroom.

The discourse generated when chance-based uncertain reward and learning is combined became the specific focus of another classroom study when students (13–14 years of age) collaborated in pairs in a science learning game (Howard-Jones & Demetriou, 2009). Gaming uncertainty appeared to subvert the conventional learning discourse, culturing the types of constructions and exchanges that are more often observed in sport. Failure was attributed to bad luck, while success was celebrated vigorously (often with singing and dancing) as a triumph of ability. Big losses as a result of gaming made a significant emotional impact but did not appear to deter students or generate a sense of unfairness. Furthermore, while supporting motivation, the elicitation of "sport talk" suggests gaming uncertainty may encourage additional resilience to failure. Indeed, the opponent's advantage was always vulnerable to a bad throw of the dice and students referred to this as a source of hope.

THE REWARD SYSTEM AND EDUCATIONAL LEARNING

Most formal learning benefits from declarative memory, that is, the ability to encode, store, and explicitly recall information. A link between declarative memory and reward makes evolutionary sense because it is helpful to remember, for example, the location where rich sources of food have previously been found. However, behavioral studies of rewarded encoding and recall have produced mixed results. Some early research has reported motivational effects on memory (Eysenck & Eysenck, 1982; Heinrich, 1968; Wiener, 1966), whereas other studies have been inconclusive (Nilsson, 1987) with even the effect of rewards on general performance motivation being called into question (Deci, Koestner, & Ryan, 1999). A study by Loftus (1972) showed effects of reward on encoding and suggested these arise from enhanced attention, rather from the reward itself. In other words, rewards can focus the attention of individuals more on some stimuli than others, thus making them more salient and easier to remember. In this study, Loftus showed that reward-associated items were not only better recognized in subsequent tests but also

fixated on more frequently while participants encoded them. Once longer fixation was taken into account, no other effect of reward on memory performance could be found. Cellular mechanisms for the effect of attention on memory recall have been put forward (Muzzio, Kentros, & Kandel, 2009). However, in a recent imaging study, researchers also suggested a more direct effect of reward on memory encoding (Adcock, 2006). Declarative memory formation is strongly linked to another brain system in the medial temporal lobe and, in particular, the hippocampus within it. Reward may promote memory formation via dopamine release in the hippocampus just before learning (Lisman & Grace, 2005). Evidence for this direct link comes from the observation that, in a study of adults incentivized by money to remember contextual information, higher activity in the nucleus accumbens prior to receiving the information predicted improved memory performance (Adcock, 2006). A complementary study by Callan and Schweighofer (2008) showed a correlation between midbrain dopaminergic activity during rewarded recall and performance. 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.

The concepts reviewed above suggest a basis for understanding how, in a learning game, the experience of the gaming may influence declarative memory of educational information. More specifically, improved memory performance during a game should be preceded by higher levels of estimated dopaminergic response (i.e., higher levels of PE). The demonstration of such an effect may suggest ways in which periods during computer-based learning games when learning is most likely can be identified and exploited via neurocomputational algorithms in the software. However, the chief educational significance of such a relationship would be to highlight the relevance of considering the brain response to reward, rather than just the absolute value of a reward itself, when incentivizing learning and/or developing and implementing learning games. We now report results of a study aimed at testing this hypothesis with respect to recall and which demonstrate the potential educational value of the concepts reviewed above.

METHOD

Participants were 16 postgraduate student volunteers (mean age 30 years 6 months; seven females and nine males). A computer-based quiz game was designed in which participants could win the points they found in one of four boxes, if
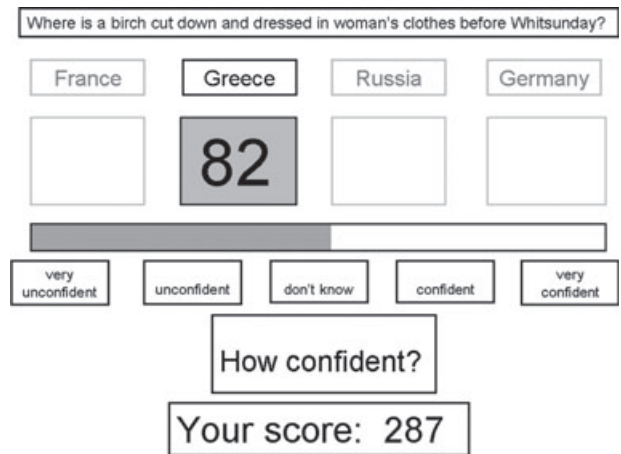

Fig. 1. Screen shot from the game used in the experiment.

they answered a subsequent multiple-choice quiz question correctly (see screen shot in Figure 1). This part of the learning game was based on the four-armed bandit task used by Daw et al. (2006) in which payouts for each arm of the bandit were generated from a Gaussian distribution whose mean diffused in a decaying random walk (so resembling the ebb and flow of natural food sources). Each participant experienced 60 trials of 30 seconds duration each. Each trial began with a 6-second window allowing participants to select one of the boxes, after which the points available from that box were immediately revealed (Figure 2). The participant was allowed to keep the points if, during the next 12 seconds, they could correctly answer a multiple-choice question. Questions concerned the content of a text which was unfamiliar to all participants ("The Golden Bough"; Frazer, 1993), to minimize the likelihood of prior knowledge. Having selected their answer, participants were given 6 seconds to indicate their confidence in their response. Feedback was then provided on the screen for 6 seconds. If their answer had been correct, the feedback would be congratulatory. If not, the correct answer would be highlighted, allowing them to study this in readiness for the question being repeated later in the quiz. Questions were randomly selected from a pool of 15 (with replacement), and the position of the four answer options on the screen would
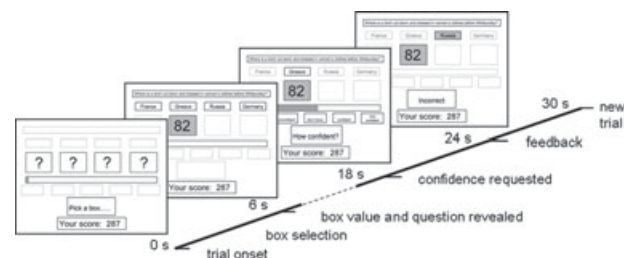

Fig. 2. Illustration of the timeline within a trial (see text for detail of trial content).

also be randomly ordered on each presentation. If a question was answered correctly twice, it would be removed from the pool. There were a total of 60 trials and participants were made aware, before taking part, that they could win a cash prize of £50 if they achieved the highest score.

MODELING OF BEHAVIOR AND ESTIMATION OF A CORRELATE OF DOPAMINERGIC ACTIVITY

In the task of deciding which of the four sources of points to select, the participant needs to learn the magnitude of reward associated with each source. The process of learning the reward magnitudes can be described by reinforcement learning models. Such models provide the estimates of participants' PE on each trial and various parameters characterizing the learning process. In our study, we have used the Kalman filter model, which was used by Daw et al. (2006) to describe learning in their experiment. The formal details of the model can be found in Daw et al. (see Supplementary Materials) and so here we provide only an intuitive description of the model (and its parameters).

The model assumes that the participant learns the expected rewards associated with each box (their values at the start of the experiment are described by parameter $\mu_{i,0}$). After choosing a box and observing the reward it provides, the expected reward associated with the box is updated proportionally to PE which, as above, is defined as the difference between the obtained reward and the expected reward. Thus, if the reward was higher than expected, PE is positive and the expected reward is increased, whereas if the reward is lower than expected, PE is negative and the expected reward is decreased.

The model also assumes that the participant has certain levels of uncertainty about the expected rewards associated with each box (their values at the start of the experiment are described by parameter $\sigma_{i,0}^2$). Each time a box is chosen, its reward uncertainty reduces, whereas for other boxes, the uncertainty increases (by amount described by parameter σ_d^2). Thus, in general, the uncertainty is lower for boxes chosen recently and higher for boxes not chosen for a long time. This uncertainty influences the degree to which the expected reward is modified after observing the reward: If the participant has great uncertainty about the expected reward associated with the chosen box, the expected reward is modified substantially (proportionally to PE—see above), so the expected reward closely approaches the value of the observed reward. In contrast, if the uncertainty is small, the expected reward is only slightly modified. The amount of expected reward modification also depends on another parameter (σ_o^2) describing how noisy the rewards seem to the participant. If they seem very noisy, they should not be greatly trusted and the expected reward is only slightly modified. However, if the rewards do not appear very noisy,

expectations are modified more strongly when a reward is observed.

Finally, the model includes forgetting of expected rewards: If boxes are not chosen, their expected rewards decay (with a rate described by parameter λ) toward a certain value (described by parameter θ).

It is also necessary to model how participants choose the boxes on the basis of their estimated rewards. When choosing between the boxes, participants faced the classic explore/exploit dilemma: They had to balance the desire to exploit (and select what seemed, according to accumulated experience, the best option) against the need to explore (i.e., select another option that may also have provided valuable information for improving future decisions). In general, there is no known optimal policy for trading off exploration and exploitation (Cohen, McClure, & Yu, 2007). It is considered that human decisions to exploit sources with highest reward involves a network that includes dopaminergic striatal processes, whereas frontal cortical regions are involved in occasionally switching to exploratory behavior that allows collection of information about other sources (Daw et al., 2006; Howard-Jones, Bogacz, Yoo, Leonards, & Demetriou, 2011). Daw et al. compared four different models of action selection that could be used to predict selections in this task. They found that the actions selected by their participants were best described by the "softmax" model, in which the choice of action, including the choice of suboptimal actions, was determined probabilistically on the basis of the actions' relative expected values. This meant, for example, that the bandit arm with the highest expected value was chosen most of the time and the arm with the lowest expected value was chosen least (but still occasionally chosen).

We fitted an identical Kalman filter model of learning and softmax model of selection as those used by Daw et al. (2006) to our data describing box selection. The values of the free parameters that gave the best fit of the softmax model for the data derived from our group of participants are provided in Table 1. Although we sought a single value of each of the above parameters for our group of participants, the parameter characterizing the extent of probabilistic exploratory behavior (β) was fitted for each individual participant, and so the mean and *SD* for this parameter across participants are provided.

We then used an internal signal of this model, the PE, as an estimated correlate of the dopaminergic activity likely to be generated in the nucleus accumbens as a result of making a box selection (Daw et al., 2006; McClure, Berns, & Montague, 2003; O'Doherty et al., 2004).

ANALYSIS AND RESULTS

Overall, there were 960 trials (16 × 60) in which 469 questions were answered correctly, 469 were answered incorrectly, and

Table 1
The Free Parameters Providing the Best Fit of the Kalman Filter and Softmax Model to the Decisions Made by Participants

Initial expected reward, $\mu_{i,0}$	Initial uncertainty in expected reward, $\sigma_{i,0}^2$	Increase in uncertainty on each trial, σ_d^2	Noise in reward observation, σ_o^2	Decay parameter, λ	Decay center, θ	Exploration parameter, β
52.656	12.991	0.20985	0.28494	0.91892	52.451	0.38835 ± 0.23663

Note: A single value of each parameters for our group of participants was determined except for exploration parameter (β), which was fitted for each individual participant, and so a mean and *SD* for this parameter across participants are provided.
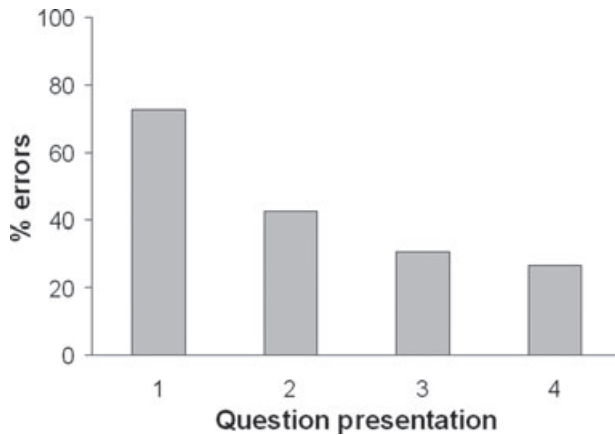

Fig. 3. Percentage error rates for questions after one presentation ($N = 362$), two presentations ($N = 280$), three presentations ($N = 194$), and four presentations ($N = 108$).

22 instances of participants failed to answer the question within the prescribed time limit. Error rates improved with repeated presentation of the same question, as shown in Figure 3.

We were interested in whether PE in the gaming component, as a correlate of reward system activity, was associated with memory performance for new information. Therefore, we restricted our analysis to those instances when participants were presented with questions they had previously answered incorrectly, on the assumption that these were opportunities to demonstrate learning achieved as a result of the game. The hypothesis under test was that PE would be greater prior to instances when participants successfully recalled the correct answer. We first applied an analysis of variance with dependent variable PE and two independent variables. The first independent (fixed) variable was recall, with two levels: successful and unsuccessful recall. Instances demonstrating successful recall were defined as occurring when correct responses were given to a question that had, on its previous presentation, been answered incorrectly. However, instances of this type in which participants rated their correct response as "very unconfident" were excluded as probable guesses. Instances demonstrating unsuccessful recall were defined as occurring when another incorrect response was given to a question that had, on its previous presentation, been answered incorrectly. The other independent (random) variable was participant (16 levels). There was a significant main effect of recall, $F(1, 17.62) = 11.65, p = .003$, and participant, $F(15, 15) = 9.38, p < .001$, with no significant interaction between these variables. When a similar analysis was carried out with the absolute value of points available in the box as the dependent variable instead of PE, the main effect for points available barely reached significance, $F(1, 16.05) = 4.77, p = .044$, with an effect for participant, $F(15, 15) = 5.34, p = .001$, and no significant interaction between participant and points. Mean values (with *SD* in parentheses) of PE for instances of successful and unsuccessful recall were 27.54 (20.35) and 23.90 (20.84), respectively. Mean values (with *SD* in parentheses) of the points available for instances of successful and unsuccessful recall were 94.98 (28.09) and 89.69 (29.29), respectively.

Figure 4 shows a simple graphical illustration of the distribution of PE and points available prior to instances of successful and unsuccessful recall, adjusted for individual differences in the game trajectories of participants. To create such an adjustment, we generated a baseline value of PE for each participant's game by calculating their mean PE over all their trials. We then calculated the mean percentage deviation of the participant's PE from this baseline for each type of trial (successful and unsuccessful recall). Across participants, these mean percentage deviations from baseline in PE prior to successful and unsuccessful recall (with *SD* in parentheses) were 5.51 (6.11) and −8.56 (9.29). In similar fashion, for each participant, we calculated the mean percentage deviation from baseline of the value of points found in the box for trials leading to successful and unsuccessful learning. Across participants, mean percentage deviations in points available prior to successful and unsuccessful recall (with *SD* in parentheses) were 1.99 (3.51) and −3.15 (5.68). A logistic regression analysis was carried out with mean percentage deviation of PE from baseline as a predictor of successful recall, which was confirmed as a predictor of recall (odds ratio = 1.54; 95% confidence intervals = 1.06–2.23, $p = .022$). In this context, the odds ratio provides the factor by which the odds of recall increases, with each percentage increase in PE. The contribution of mean percentage deviation of points from baseline as a predictor variable in the regression model did not approach statistical significance ($p = .283$).
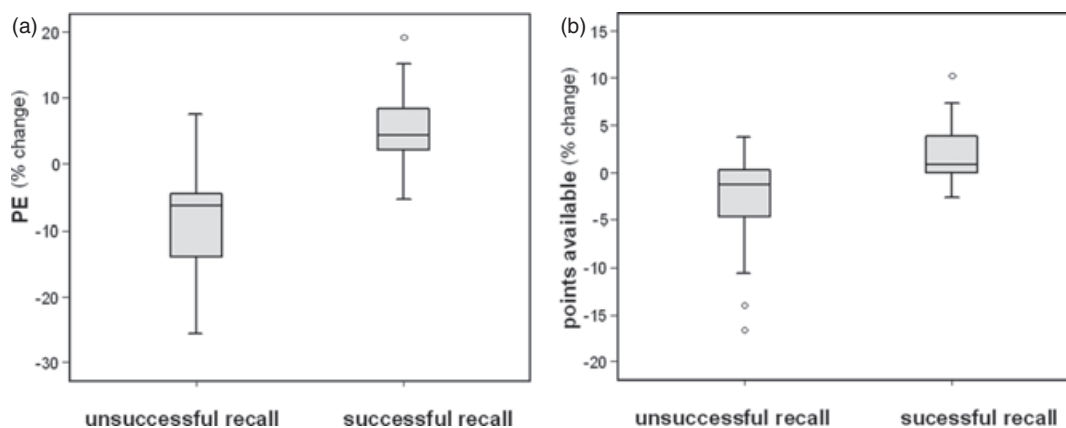
Fig. 4. Box plots of (a) mean percentage increase in PE prior to successful and unsuccessful recall and (b) mean percentage increase in points available prior to successful and unsuccessful recall.

DISCUSSION

As theorized on the basis of its correlation with dopaminergic reward activity, PE was a significant predictor of recall in our learning game. These results complement the observed relationship between midbrain dopaminergic activity and recall performance (Callan & Schweighofer, 2008), suggesting the value of neural concepts in understanding and developing learning games. However, several caveats should be noted. First, our experimental design restricted our observations to the impact of reward on recall, although the imaging literature suggests that estimating the brain's response to reward may also help predict encoding performance (Adcock, 2006). The opportunity for encoding the correct answer in our experiment occurred 18 seconds after choosing a box. The dopaminergic response to box selection would have subsided after such a delay, which was beyond the type of time period within which a dopaminergic response and an action should co-occur for learning to be enhanced (Bogacz, McClure, Li, Cohen, & Montague, 2007). Additionally, encoding would have occurred in the negative context of having just lost points. Second, the design of this study included a task involving meaningful information which helped demonstrate educational relevance. However, this learning was factual recall and did not test understanding, which is of increasingly greater educational interest. However, if improved recall occurs via enhanced attention, then deeper types of learning may also benefit and, as discussed below, uncertain reward is now being applied in the development of pedagogy that attends to these types of learning. There is considerable scope for further research on the underlying cognitive and neural processes by which reward influences learning and for evaluating applications of this knowledge in a range of educational contexts. Finally, when considering the limitations of our results, it should be borne in mind that participants in our learning game played individually. Workers reviewing the potential of neuroscience in education (Meltzoff, Kuhl, Movellan, & Sejnowski, 2009) echo the many voices in education that emphasize the importance of social interaction for learning and, indeed, classroom games are frequently collaborative and usually competitive. Little is known about the neuroscience of competitive gaming, although recent research reveals we code our competitor's losses egocentrically as our rewards (Howard-Jones et al., 2011), suggesting it may be our competitor's unexpected losses that support our declarative memory performance.

The emerging relationship between reward and learning may contribute to the next generation of commercial educational games. These may include systems that maximize dopaminergic activity for periods during a game that are critical for learning and perhaps even adapt to the behavior of their players to achieve this. However, application is not restricted to computer games and neural concepts about reward have underpinned recent efforts to develop pedagogy for whole-class teaching with immersive gaming (or "twigging") (Howard-Jones, 2010; Howard-Jones et al., under revision). This pedagogy aims to support all levels of understanding (as defined by Bloom) using learning games to generate brief periods when students may be especially engaged and receptive as a result of anticipating uncertain reward and exploiting these as special opportunities for the teacher to scaffold their learning.

In summary, our study adds to existing evidence suggesting education may benefit from revising the constructions around reward and learning that presently characterize its discourse. In particular, the potential of uncertain reward to increase motivation provides insight into an important aspect of how games, including learning games, engage their players. Moreover, computational modeling of reward system activity during gaming can help educators and developers understand how gaming events influence educational learning. Much, however, remains to be investigated, particularly in regard to the neural

and cognitive processes by which reward supports the encoding and recall of declarative memory. Although, as demonstrated here, current findings from neuroscience can already provide insight into the reward–memory relationship observed in behavioral terms, further research is needed to determine how this occurs (e.g., directly and/or via attention), and how other factors, such as emotional response, interact with these processes. More broadly, the integration of such understanding into everyday classroom practice may require the careful engineering of pedagogical approaches that attend to the many important aspects of classroom learning we already appreciate as important. This must include consideration of, for example, how techniques that exploit the engaging properties of uncertain reward conjoin with the teacher's scaffolding of learning, feedback to students, and the general dynamic properties of teacher–learner interaction (Perkins, 2009). Collaborative enterprise between education and the sciences of mind and brain may help address these questions and contribute to develop new ways to engage the learners of the future.

REFERENCES

Adcock, R. A. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron, 50*, 507–517.

Atkinson, J. W. (1957). Motivational determinants of risk taking behaviour. *Psychological Review, 64*, 359–372.

Bergin, D. A. (1999). Influences on classroom interest. *Educational Psychologist, 34*(2), 87–98.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews, 28*(3), 309–369.

Bogacz, R., McClure, S. M., Li, J., Cohen, J. D., & Montague, P. R. (2007). Short-term memory traces for action bias in human reinforcement learning. *Brain Research, 1153*, 111–121.

Caillois, R. (1961). *Man, play, and games.* New York: Free Press.

Callan, D. E., & Schweighofer, N. (2008). Positive and negative modulation of word learning by reward anticipation [Article]. *Human Brain Mapping, 29*, 237–249.

Clifford, M. M., & Chou, F. C. (1991). Effects of payoff and task context on academic risk taking. *Journal of Educational Psychology, 83*, 499–507.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration [Proceedings paper]. *Philosophical Transactions of the Royal Society B-Biological Sciences, 362*(1481), 933–942.

Daw, N., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441*, 876–879.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation [Review]. *Psychological Bulletin, 125*, 627–668.

Eysenck, M. W., & Eysenck, M. C. (1982). Effects of incentive on cued-recall [Article]. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 34*(November), 489–498.

Farooqi, I. S., Bullmore, E., Keogh, J., Gillard, J., O'Rahilly, S., & Fletcher, P. C. (2007). Leptin regulates striatal regions and human eating behavior [Article]. *Science, 317*(5843), 1355.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*, 1898–1902.

Fisher, K. R., Marshall, P. J., & Nanayakkara, A. R. (2009). Motivational orientation, error monitoring, and academic performance in middle childhood: A behavioral and electrophysiological investigation. *Mind, Brain, and Education, 3*, 56–63.

Frazer, J. G. (1993). *The golden bough: A study in magic and religion.* Ware, Hertfordshire, UK: Wordsworth Editions.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation Gaming, 33*, 441–467.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* London: Palgrave Macmillan.

Heinrich, B. A. (1968). Motivation and long-term memory [Article]. *Psychonomic Science, 12*(4), 149.

Hong, J. C., Hwang, M. Y., Lu, C. H., Cheng, C. L., Lee, Y. C., & Lin, C. L. (2009). Playfulness-based design in educational games: A perspective on an evolutionary contest game. *Interactive Learning Environments, 17*(1), 15–35.

Howard-Jones, P. A. (2008). Philosophical challenges for researchers at the interface between neuroscience and education. *Journal of Philosophy of Education, 42*, 361–380.

Howard-Jones, P. A. (2010). *The teacher's handbook of twig: Minds, brains and teaching with immersive gaming.* Retrieved from www.lulu.com.

Howard-Jones, P. A., Bogacz, R., Yoo, J. H., Leonards, U., & Demetriou, S. (2010). The neural mechanisms of learning from competitors. *Neuroimage, 53*, 790–799.

Howard-Jones, P. A., & Demetriou, S. (2009). Uncertainty and engagement with learning games. *Instructional Science, 37*, 519–536.

Howard-Jones, P. A., Tanimoto, E., Jones, C., Demetriou, S., Morgan, O., Perkins, P., et al. (under revision). Learning to "twig"—Co-constructing pedagogical concepts across neuroscience and education for teaching with immersive gaming. *Pedagogies.*

Immordino-Yang, M. H., & Damasio, A. R. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education, 1*, 3–10.

Johnson, S. (2005). *Everything bad is good for you.* London: Penguin/Allen Lane.

Juul, J. (2003). *The game, the player, the world: looking for the heart of gameness.* Paper presented at the Level Up: Digital Games Research Conference 2003, Utrecht.

Kirriemuir, J., & McFarlane, A. (2004). *Literature review in games and learning.* Bristol: Futurelab.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience, 21*(RC159), 1–5.

Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., et al. (1988). Evidence for striatal dopamine release during a video game. *Nature, 393*, 266–268.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory [Review]. *Neuron, 46*, 703–713.

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures [Article]. *Cognitive Psychology, 3*, 525–551.

Loftus, G. R., & Loftus, E. F. (1983). *Mind at play: The psychology of video games.* New York: Basic Books.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4,* 333–339.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum [Article]. *Neuron, 38,* 339–346.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning [Review]. *Science, 325*(5938), 284–288.

Muzzio, I. A., Kentros, C., & Kandel, E. (2009). What is remembered? Role of attention on the encoding and retrieval of hippocampal representations [Review]. *Journal of Physiology-London, 587,* 2837–2854.

Nieuwenhuis, S., Heslenfeld, D. J., von Geusau, N. J. A., Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage, 25,* 1302–1309.

Nilsson, L. G. (1987). Motivated memory—Dissociation between performance data and subjective reports [Article]. *Psychological Research-Psychologische Forschung, 49,* 183–188.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning [Article]. *Science, 304*(5669), 452–454.

Office for Standards in Education. (2001). *Improving attendance and behaviour in secondary schools.* London: Author.

Perkins, D. (2009). On grandmother neurons and grandfather clocks. *Mind, Brain, and Education, 3,* 170–175.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward [Article]. *Science, 275*(5306), 1593–1599.

Shizgal, P., & Arvanitogiannis, A. (2003). Gambling on dopamine. *Science, 299,* 1856–1858.

Van Geert, P., & Steenbeek, H. (2008). Brains and the dynamics of "wants" and "cans" in learning. *Mind, Brain, and Education, 2,* 62–66.

Wiener, B. (1966). Motivation and memory. *Psychological Monographs: General and Applied, 80,* 1–22.

Wise, R. A. (2004). Dopamine, learning and motivation [Review]. *Nature Reviews Neuroscience, 5,* 483–494.

Zimmerman, E., & Fortugno, N. (2005, April 5). Soapbox: Learning to play to learn—Lessons in educational game design. *Gamasutra.* Retrieved January 21, 2011, from http://www.gamasutra.com/view/feature/2273/soapbox_learning_to_play_to_learn_php