Boredom, Information-Seeking and Exploration

Andra Geana (ageana@princeton.edu)[*+] Robert C. Wilson (bob@email.arizona.edu)°
Nathaniel Daw (ndaw@princeton.edu)[*+] Jonathan D. Cohen (jdc@princeton.edu)[*+]

[*]Department of Psychology/[+]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540 USA
°Department of Psychology, University of Arizona, Tucson, Arizona 85721 USA

Abstract

Any adaptive organism faces the choice between taking actions with known benefits (exploitation), and sampling new actions to check for other, more valuable opportunities available (exploration). The latter involves information-seeking, a drive so fundamental to learning and long-term reward that it can reasonably be considered, through evolution or development, to have acquired its own value, independent of immediate reward. Similarly, behaviors that fail to yield information may have come to be associated with aversive experiences such as boredom, demotivation, and task disengagement. In accord with these suppositions, we propose that boredom reflects an adaptive signal for managing the exploration-exploitation tradeoff, in the service of optimizing information acquisition and long-term reward. We tested participants in three experiments, manipulating the information content in their immediate task environment, and showed that increased perceptions of boredom arise in environments in which there is little useful information, and that higher boredom correlates with higher exploration. These findings are the first step toward a model formalizing the relationship between exploration, exploitation and boredom.

Keywords: boredom, exploration, information-seeking

Introduction

The complexity and uncertainty of the real world makes choosing between behavioral options challenging. We often have numerous alternatives from which to choose, and we often have incomplete information about many of these. When making a choice, therefore, we must often consider two competing goals: one is earning as much reward as possible, and the other is to gain information about the alternatives, that may improve our choices in the future. This dilemma is known as the exploration-exploitation tradeoff (Cohen, McClure & Yu, 2007) that has been the subject of growing investigation.

Ample evidence has shown that humans and animals engage in information-seeking, even at the cost of current reward (Behrens et al. 2007; Bromberg-Martin & Hikosaka, 2009). Theoretical models show that the information acquired under these scenarios can improve the computation of value estimates, leading to better choices and greater reward over the longer term (Wilson et al. 2014). In this context, it is worth noting that evidence from the literatures on curiosity and creativity suggests that humans and other animals find new information valuable even when it is not possible to immediately use it to acquire better reward (Gottlieb, 2012; Kidd & Hayden 2015), and that this desire for information can lead to the disengagement from activities that are otherwise rewarding. This type of disengagement is a widespread and long-documented phenomenon, and has been considered to be an important factor in boredom (Csikszentmihalyi, 2000; Eastwood et al., 2012). That is, boredom may reflect a bias toward the pursuit of behavior that is not immediately rewarding, but that may provide information useful for increasing long-term reward — in other words, exploration.

Consistent with this proposition, monotonous, repetitive, or insufficiently informative or stimulating tasks increase the perception of boredom (Hill & Perkins 1985; Pattyn et al. 2008), and they are valued and attended to less (Schmidhuber, 1997; Eastwood et al., 2012). Furthermore, observed behavioral correlates of boredom seem to suggest a link to exploration: boredom has been found to prompt the search for new stimulation (Fowler 1967; Meagher & Mason 2012), the increased drive to discover new goals and resources (Eastwood et al. 2012), and the tendency toward innovation and creativity (Bench & Lench 2013).

Complementing these empirical findings, studies in the reinforcement learning and machine learning literatures have suggested that boredom might serve an adaptive function – for instance by signaling an increased opportunity cost of choosing the current option compared to other available options (Kurzban et al. 2013), or by ensuring that too-well-known (i.e. insufficiently informative) options are penalized in value, which can lead to better learning and a higher long-term reward rate (Schmidhuber, 1991; Simsek & Barto, 2006). While these empirical and theoretical lines of work are consistent with the assertion that boredom reflects a signal biasing behavior toward exploration, to date there has been no direct test of this hypothesis. Empirical work has largely been qualitative and observational (Hill & Perkins 1985), while theoretical predictions have not been tested in human participants.

To address this gap in the literature, we conducted three experiments that parametrically manipulated the information content of participants' task environment, examined perceptions of boredom, choices to engage with the task or abandon it in favor of an alternative, and the link between self-reported boredom and overall exploration behavior. Results showed that varying information content elicited correlated changes in boredom, and that people showed increased exploration in response to boring (i.e., less informative) contexts. We conclude by discussing the relationship of our findings to a theoretical model that parallels optimal foraging theory (which focuses on immediate reward) to formalize the value of information (as a proxy for future reward) in explore-exploit decisions.

Experiment 1

This experiment tested for correlations between self-reports of boredom and the amount of useful information that can be gained by continued engagement in the current task. Previous theoretical work has suggested that prediction error (PE) can be used to measure the amount of useful information left to learn in a given task environment (Schmidhuber, 1997). Specifically, it was proposed that asymptotically low prediction errors signal that a good representation of the task has already been learned, and persistently high PEs signal that the task environment is too random and errors cannot be reduced. Both cases were suggested to cause boredom, and the increased drive to disengage and search for more informative alternatives. (Schmidhuber, 1990; Luciw et al., 2013). Accordingly, we used change in prediction error as an index of learning and, correspondingly, participants' estimates of information available in the task environment,

We found that, in line with theoretical work, people's self-reported boredom ratings correlated negatively with change in prediction errors, consistent with the hypothesis that boredom increases as the amount of information that can be gained from the task decreases. When participants played a computer game for which they already had all the information necessary to perform perfectly (so there was nothing left for them to learn), or the outcomes were completely random (so there was no task structure to learn), they reported being more bored than in a task in which they could acquire useful information as they played.

Methods

Participants Twenty-five Princeton University undergraduates (ages 18 to 22) performed the experiment in exchange for course credit.

Task Participants were asked to predict numbers generated by a virtual machine (for a similar design, see Nassar et al.2010). On each trial, they made their predictions by adjusting a vertical slider (the "prediction slider", see fig 1A) between 0 and 100 to indicate the next value that the virtual machine would generate. After they adjusted the slider, they pressed the space key to confirm their prediction, and the machine generated the number for that trial. Games in the task consisted of thirty trials, and changes between games were signaled to the participants; there were twenty-four games in total, with the session lasting approximately one hour.

The critical experimental variable was the difference between the participants' prediction and the actual generated number, that we refer to as the Prediction Error (PE). Participants were rewarded based on these prediction errors: the smaller the error (i.e., the closer their prediction was to the actual number), the more points they received.

The machine generated numbers according to an underlying distribution, which differed between conditions. In the "Gaussian" condition, numbers were generated from a Gaussian distribution with a fixed mean and standard deviation; however, each number was not displayed until after the participant recorded their guess. In the "Certain" condition, numbers were generated from a Gaussian, and displayed on the screen before the participant made their response. In the "Random" condition, numbers were generated uniformly between 0 and 100, but again not displayed until after the participant recorded their guess. Therefore, the underlying distribution of the number-generating machine was such that participants had to either learn the generative process to gradually reduce their prediction error (in the Gaussian condition), they were already told the next number and did not need to learn anything to make perfect predictions (the Certain condition), or the numbers were randomly generated and participants could not reduce their prediction error (in the Random condition). Participants were also asked to self-report their level of boredom on every fourth trial (for a total of ten times throughout each game). They did so by adjusting another slider (the "boredom slider") at the bottom of the screen (figure 1A).

Results

The "Certain" condition elicited the highest boredom ratings in all participants (repeated measures ANOVA, $F(2, 60) = 5.03$, $p = 0.01$). The ratings for the Certain condition were consistently higher than for the Gaussian and Random conditions in both early trials (first six games) and late trials (last six games), as shown in figure 1B, and for the average ratings within a game (1C). The Gaussian condition, by contrast, was consistently rated as the least boring. The Random condition was rated in-between the other two.

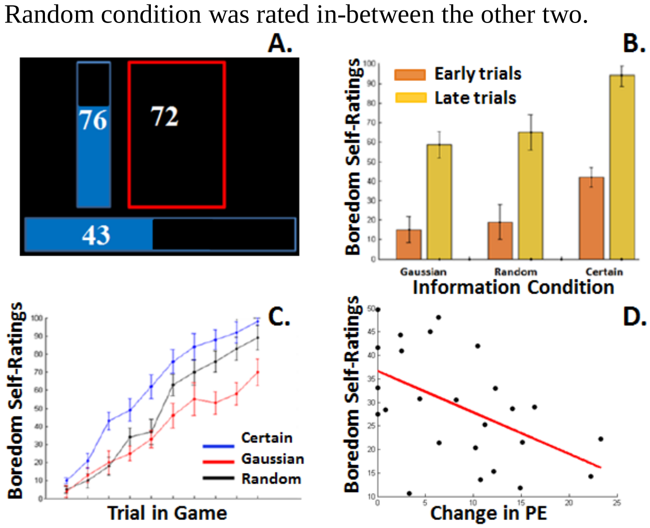

Figure 1: A. **Task design**: participants had to predict the next number generated by the virtual machine (the red rectangle). They predicted by adjusting the vertical slider to the left. They rated their boredom using the horizontal slider at the bottom. B, C. Ratings for the Certain condition (no useful information content) were significantly higher both within a game, and for early and late games. D. Average change in prediction error within a game correlated with the average boredom rating for that game.

There was also a significant observed main effect of time on boredom ratings: for all three conditions, later ratings were significantly higher than earlier ratings, both within a game , and across the entire session in early versus late trials (figures 2A and 2B, two-way repeated measures ANOVA, $F(2,18) = 13.39$, $p < 0.01$).

Absolute prediction errors were computed for each game (as the absolute value of the difference in participants' prediction from the number generated on each trial), and the average change in prediction error for each game was computed as the average difference between PEs on consecutive trials. These values were then binned for changes in PE, and the average boredom ratings corresponding to those games were calculated (regardless of which condition those games were in – although, as explained in the methods, the participants were only able to significantly reduce their PE in the Gaussian condition). As shown in figure 1D, there was a significant negative correlation between the change in prediction error and the boredom ratings (R^2=0.2613, $p < 0.01$).

Discussion

This first experiment revealed a correlation between the information content available in a task (operationalized as change in prediction error), and the subjective perception of boredom. The results suggest that the amount of information that can be learned from a task is linked to how boring the task is perceived: the Certain condition – i.e., the one in which there was no useful information to be learned by performing the task, because all the information was already given to participants – elicited the highest boredom ratings (fig. 1B). Conversely, the Gaussian condition, in which it was possible to improve predictions by learning the underlying number-generating distribution, was rated as the least boring. This is consistent with previous theories on 'too much or too little information' causing suboptimal levels of arousal (Schmidhuber, 1997; de Rijk, Schreurs & Bensing 1999), as well as with the notion of "desirable difficulty" – i.e., the notion that there is a certain amount of effortful information-processing that helps learning and is perceived as desirable (Bjork & Bjork, 2011). To our knowledge, this is the first direct empirical demonstration of a correlation between state boredom and a quantitative manipulation of information-content.

Experiment 2

This experiment examined the extent of task disengagement and exploratory behavior in response to information-content. Using the same three conditions that elicited differential boredom levels in participants in Experiment 1, we modified the number-prediction task to allow participants to decide, on their own, whether they wished to persist in the current game or quit and move on to another game. This afforded a more direct examination of the relationship between information content and exploration, and how this traded off against present reward.

Methods

Participants Twenty Princeton University undergraduates (ages 18 to 22) performed the experiment. They were compensated with $12 for their time, plus a performance-dependent bonus of up to $7.

Task Participants played a variant of the number-prediction task used in Experiment 1. However, in this version, games did not have a fixed length. Rather, participants were told that a game could go on for up to one hundred trials, but they could choose to end it earlier and move on to a new game at any time by pressing the "reset" button on the screen (figure 2A). If they pressed the "reset" button, they would see a brief inter-game screen, and then start a new game with a new number-generating process. Participants were told that the task would take approximately fifty minutes, regardless of how many games they went through in that time: the task finished at the end of the current game once the fifty-minute time period was up. There was no boredom slider in this design. In all other respects, the tasks and conditions (Gaussian, Certain and Random) were the same as in Experiment 1. After each game ended (either because the participant pressed the "Reset" button, or after 100 trials), the next game was drawn from one of the three conditions with equal probability.

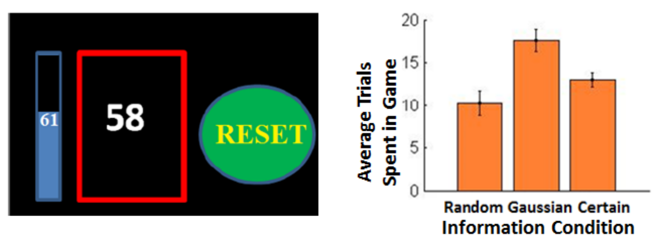

Figure 2: A. Variant of the Experiment 1 task. Participants could click **a** 'Reset' button to end current game and start a new one. B. Participants spent most time in the Gaussian games (where information content was most useful), despite the fact that the Certain games were the most rewarding.

Results

No participants chose to stay in any game for the entire duration of one hundred trials; all pressed the "reset" button to move on to a new game well before the total number of possible trials in the current game had elapsed. There was a significant difference, however, in the average number of trials spent in a game before choosing to switch (repeated measures ANOVA, $F(2,69) = 7.04$, $p < 0.01$; fig 2B), with most participants spending significantly longer on games in the Gaussian condition than in either of the other two conditions ($F(2,69)$=9.22, p<0.01). The difference between the Certain and Random conditions was not significant (paired t-test, $t(23) = -1.29$, $p = 0.206$). Furthermore, the probability of switching away after the first trial of a new game was significantly higher for the Certain games than either of the other two conditions; again, the values for the

Certain and Random and Certain conditions were statistically indistinguishable ($t(23) = 0.11$, $p = 0.91$).

Discussion

The results of this experiment suggest that the conditions with low information content, and associated with boredom in Experiment 1, carry a "penalty." The three conditions elicited a U-shaped curve for quitting times that mirrored the curve for boredom in Experiment 1 (figure 2B). This behavior is particularly striking for the Certain condition, in which the potential for reward was highest (participants had access to the correct prediction on every trial). Despite this, none of the participants stayed in a Certain game until it terminated, choosing instead to switch away from these games *more* frequently than from the other two conditions. Thus, participants were willing to take a point loss in order to quit the Certain game early.

This switching behavior resembles exploration – foregoing current reward-maximizing behavior in favor of options associated with a greater likelihood to gain information. There are at least two ways in which such behavior could be viewed as adaptive. First, it might help improve participants' representation of the task environment, and thus make better decisions about which tasks to perform: Experiencing more games could lead to faster discrimination between the Gaussian and Random games, by reducing estimation uncertainty (Payzan-LeNestour & Bossaerts, 2011). This could help participants determine the condition they were in earlier in the game which, in turn, would allow them to quit Random games earlier — a reward-maximizing strategy. Second, even if participants were not consciously trying to improve their representation of the task environment, but rather just aimed to terminate boring games earlier, it is possible that this drive to escape boring situations reflects an endogenous bias toward exploration, acquired over the course of evolution and/or development. Such a bias may reflect the prevailing value of exploration in the real world which is complex and rich in opportunities to gain information. This is consistent with previous work suggesting that humans and animals show an inherent aversion to low-stimulation, information-poor tasks (Fowler, 1965). Experiment 3 was conducted to further explore this possibility and, in particular, the idea that boredom, exploration and the value of information are sensitive to the alternatives available in the environment.

Experiment 3

This experiment examined whether boredom and exploration are dependent not only on the information content of the current task, but also on the (perceived) alternatives in the task environment. This builds on prior suggestions that engagement in a current task engagement is sensitive to global properties of the environment (Fowler 1967; Csikzentymihalyi 1997), and that motivation and boredom are associated with the opportunity cost of current behavior compared to alternative possible behaviors (Eastwood et al. 2012; Kurzban et al. 2013). However, to our knowledge, these ideas have not yet been tested in controlled laboratory experiments..

Here, we tested whether it was possible to change people's perceptions of boredom by manipulating the availability of other more or less attractive options in the task environment, and the extent to which this impacted exploratory behavior.

Methods

Participants Forty participants recruited from the Princeton University undergraduate community received course credit for participating in this experiment.

Task Participants played a task that consisted of two parts (referred to as part A and part B, figure 3A), played in order. They were told at the beginning of the experiment about both parts, and what each part would entail. They were also regularly reminded about part B while playing part A (they received three reminders, every five minutes, for the twenty-minute duration of part A).

Part A involved a two-armed bandit horizon task used in previous work to quantify exploratory behavior (Wilson et al., 2014). Participants had to choose between an ambiguous bandit (from which they had seen only one reward sample), and an unambiguous bandit (from which they had seen three reward samples). The decision horizon of the task was manipulated to be either short (a total of five trials: four forced-choice trials during which participants received three samples from one bandit, and one sample from the other, and one free-choice trial in which they could choose whichever bandit they wished), or long (a total of ten trials: four force-choice, six free-choice trials). This task can be used to quantify exploratory behavior (defined as the frequency with which the ambiguous option is chosen when its estimated value falls below that of the unambiguous option), and has shown that such behavior increases with task horizon. Every seven games, participants received a query screen that asked them to assess task-related factors such as difficulty, the average number of points they earned, or how many games they have played so far. Among these questions, there were regular queries about their level of interest in the game, which they were asked to rate a total of six times over the course of the task. Part A was the same for all participants, and lasted for a total of seventy-two games and approximately twenty minutes.

Part B differed between participants. There were four conditions, each involving one of a set of tasks that had been previously rated by a different sample of participants in a brief interest-rating study. As noted above, participants were instructed from the start that they would be performing Part B after Part A was finished, and they were reminded about it three times during part A. Each participant was randomly assigned to one of the four conditions, as follows: 10 participants watched a "CrashCourse" YouTube video (previously rated as a highly interesting task); 10 participants counted the number of words in a two-page mathematical typography article (a task previously rated as highly boring); 11 participants played a simple colormatching game (previously rated at medium levels of interestingness), and 9 participants played another round of a bandit task very similar to the one played in part A.

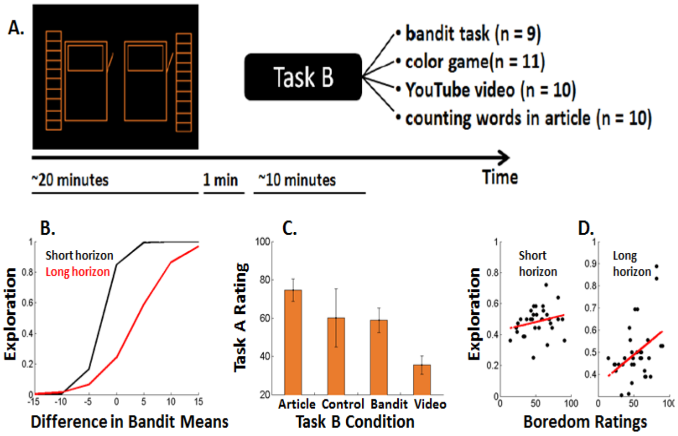

Figure 3. A. Participants played 20 minutes of the two-armed bandit horizon task (Wilson et al., 2014), followed by a break, and 20 minutes of Task B. B. Choice curves for the short horizon (black) and long horizon games (red)show more exploration in long horizon. C. People's ratings of how boring Task A was differed depending on which Task B they had to perform. D. Higher boredom ratings led to more exploratory behavior in the long horizon.

Results

Participants rated the interestingness of the bandit task (played by all participants during part A) differently, depending on the task they were told they would play during part B: the YouTube video (rated as highly interesting), the word-counting task (highly boring), or one of the control tasks (intermediate ratings). A one-way ANOVA showed a significant main effect of condition ($F(2,26) = 9.07$, $p < 0.01$), with participants who expected to watch the video (in part B) rating the bandit task (while performing it during part A) as more boring than those participants who expected to perform the word-counting task (paired t-test, $t(1,15) = 5.21$, $p < 0.01$). The ratings for the two controls fell in an intermediate range (fig 3C).

Exploration in the bandit task was defined as choosing the ambiguous bandit. Participants replicated the results from previous studies (Wilson et al., 2014) regarding the impact of decision horizon on exploration: the decision curve for the long horizon (fig. 3B, red) is flatter and shifted to the right compared to the curve for the short horizon (black), suggesting higher exploration for long horizons, both in the form of decision noise and of information bonus.

Average exploration within each correlated significantly with participants' ratings of the task: the higher the boredom rating, the more likely participants were to explore (figure 3D). This pattern was only observed in the long decision horizon, and not in the short horizon ($F(1,33) = 6.01$, $p = 0.02$ for Horizon 6, $F(1,33) = 1.88$, $p = 0.17$ for Horizon 1), and it was observed for estimates of both exploration-related parameters: decision noise and information bonus (which were both significantly higher in the long horizon with high boredom ratings).

Discussion

Experiment 3 showed that it was possible to manipulate participants' perceptions of boredom and bias toward exploration within a given task, based on a manipulation of the task environment (in this case, the next task to be performed). All participants played the same bandit task for the same period of time, but those participants who had been told they would perform an interesting task (watch a YouTube video) after finishing the bandit task, rated the bandit task as significantly more boring than those who had been told that they would have to perform a boring task (word-counting; figure 3C). This is consistent with previous findings regarding the effect of increased available stimulation on relative motivation (Fowler, 1967), as well as the theoretical framework proposed recently by Kurzban et al (2013) that relates boredom to perceived opportunity cost.

A second result was the strong correlation between participants' boredom and their exploratory behavior in the two-armed bandit horizon task: participants who rated the bandit task as more boring also showed significantly higher exploration (fig 3E). Increased exploration in response to a boring situation has previously been suggested in the literature (Cohen, McClure & Yu 2007), but this is the first report of an empiricaclly measured correlation between self-reported boredom and a quantitative measure of exploration.

Interestingly, Task B's identity is, by design, irrelevant to the value of exploration *within* the bandit task (since Task B only occurs after it is over), but nevertheless affected it. This suggests the perceived value of alternatives is a globally estimated quantity that generalizes freely and potentially inappropriately to particular, more constrained choices, much as has been suggested for opportunity costs in other domains (Niv et al., 2007).

General Discussion: An Information-Sampling Account of Boredom and Exploration

The results of the three experiments described above provide convergent support for a relationship between information context, boredom, and exploration. We showed that levels of reported boredom correlate with the amount of useful information that participants can extract from the environment, such that when there is too much or too little available information, they become more bored (Experiment 1). We also showed that perceptions of boredom can be modulated based on the environment in which the task is performed (Experiment 3). This strongly suggests that when people compute the value of staying with the current behavior, they take into account some measure of relative value between the local and global environments. Lastly, we showed that when boredom levels are high (due to low information content), people show a greater propensity to switch away from the current task – both in situations when that may be useful (Experiment 3), and even when it would appear to be suboptimal (Experiment 2).

These findings suggest two factors contribute to boredom and the tendency to switch behavior: a lack of useful information in the current task, and the availability of more (valuable) information outside the current task. This is consistent with both previous theoretical work linking boredom to too much or too little available information (Schmidhuber, 1997; Csikszentmihalyi, 2000), and with more recent accounts that suggest that boredom might signal an increased opportunity cost of performing the current task (Kurzban et al, 2013). Our results are also consistent with the idea that humans need constant access to a certain amount of information in order to maintain a satisfactory level of adaptive behavior (Zakay 2014), and that that information comes in the form of optimal levels of variability in the environment (Kidd et al., 2012).

The work presented here provides an important first step toward formalizing the link between information, boredom and exploration. Our treatment of exploration — as the decision to switch away from a current task with the specific goal of acquiring more information — parallels the formal theoretical treatment of foraging behavior, in the marginal value theorem (MVT; Charnov, 1976), as the decision to switch away from a current task with the specific goal of acquiring greater reward. In related work, we have sought to formalize this definition of exploration in a model that relates it to information acquisition in a way that parallels the definition of foraging for reward in MVT, and to test this model in an empirical study. In aggregate, this line of work promises to offer a normative understanding of human exploratory behavior and, within this framework, to cast boredom as an adaptive response to situations that should favor exploration in the service of learning, and the maximization of long-term reward.

References

Behrens. T.E.J., Woolrich, M.W., Walton, M.E. Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience* 10: 1214 – 1221

Bench, S. W., & Lench, H. C. (2013). On the function of boredom.*Behavioral Sciences*, *3*(3), 459-472.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*

Bromberg-Martin, E. S., & Hikosaka, O.(2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, 119-126.

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. *Philosophical Transactions of the Royal Society: Biological Sciences. 362*, 933–942

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety*.

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

de Rijk, A. E., Schreurs, K. M., & Bensing, J. M. (1999). Complaints of fatigue: related to too much as well as too little external stimulation?. *Journal of Behavioral Medicine*, *22*(6), 549-573.

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The Unengaged mind defining boredom in terms of attention. *Perspectives on Psychological Science*, *7*(5), 482-495.

Fowler, H. (1967). Satiation and Curiosity: Constructs for a Drive and Incentive-Motivational Theory of Exploration. *Psychology of learning and motivation*,

Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*,*76*(2), 281-295.

Hill, A. B., & Perkins, R. E. (1985). Towards a model of boredom. *British Journal of Psychology*, *76*(2), 235-240.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449-460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(06), 661-679.

Luciw, M., Kompella, V., Kazerounian, S., & Schmidhuber, J. (2013). An intrinsic value system for developing multiple invariant representations with incremental slowness learning.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience*, *30*(37), 12366-12378.

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue?. *Physiology & Behavior*, *93*(1), 369-378.

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. PLoS Computational Biology, 7, e1001048.

Schmidhuber, J. (1991, November). Curious model-building control systems. In *Neural Networks*, *1991. 1991 IEEE International Joint Conference on* (pp. 1458-1463). IEEE.

Şimşek, O., & Barto, A. (2006). An intrinsic reward mechanism for efficient exploration. *Proceedings of the 23rd international conference*, 833-840.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General, 143*(6), 2074.

Zakay, D. (2014). Psychological time as information: The case of boredom.*Frontiers in psychology*, *5*.