



IBRO/IBE-UNESCO Science of Learning Briefings

Insights from neuroeconomics for the classroom

Jean-Claude Dreher



Title	Insights from Neuroeconomics for the Classroom
Series	IBRO/IBE-UNESCO Science of Learning Briefings
IBE Director	Dr. Mmantsetsa Marope
Author of brief	Dr Jean-Claude Dreher
Author affiliation	CNRS, Institute of Cognitive Science, Lyon, France
Date	December 2d 2021

Dr Jean-Claude Dreher is research director of the lab 'Neuroeconomics, Reward and Decision Making' (<u>https://dreherteam.wixsite.com/neuroeconomics</u>) at the Institut des Sciences Cognitives Marc Jeannerod (UMR 5229, Lyon, France). He investigates the neural mechanisms underlying decision making, motivation and reward processing in humans, using concepts from cognitive neuroscience, psychology and behavioral economics. He uses experimental tools such as model-based functional Magnetic Resonance Imaging to understand the computational processes involved when making a choice.

Executive Summary

- Neuroeconomics is a field of research seeking to explain how the brain makes decisions between different options.

- Decision making can be decomposed in several stages, including the valuation of potential options according to their costs and benefits and selection of the option having the highest subjective value.

- The concept of a common currency in the brain reflects the fact that our brain has to integrate and compare options of different types to choose between them.

- Brain imaging studies in humans have identified a valuation system, including the ventromedial prefrontal cortex and the ventral striatum, that implements this common currency in the brain to compare between options.

- One classical example to identify this valuation system consists of proposing to participants a choice between a small immediate reward and a delayed larger reward. This type of choice refers to Delay discounting.

- Delay discounting measures impatience and refers to the empirical finding that both humans and animals value immediate rewards more than delayed rewards. Discounting the subjective value of a delayed reward engages the brain valuation system, which consists of the ventromedial prefrontal cortex and ventral striatum.

- Cognitive control refers to the ability to override our impulses and to make decisions based on our goals. This ability engages the dorsolateral prefrontal-parietal network.

Adolescents are more likely than adults to opt for smaller rewards sooner than larger ones later.
 This stronger impatience is reflected by lower connectivity between the dorsolateral prefrontal cortex, part of the cognitive control network, and the ventral striatum, part of the valuation system.
 Higher cognitive control in children predicts future academic achievement.

- Working memory capacity can be increased through training programs and such training produces plastic changes in the cognitive control brain network.

Introduction

Neuroeconomics and value-based decision making

Neuroeconomics is an interdisciplinary field which combines research from neuroscience, behavioral economics and cognitive and social psychology. It seeks to explain how humans and animals make decisions between different options. Classical topics of research in the field include how individual preferences, value, risk, time preferences and social preferences are learned and represented in the brain. The focus of this brief is to introduce important concepts from the field, such as utility and common currency, and to review our current understanding of their neuronal implementation. We will also pinpoint potential implications of this research for education, focusing in particular on the neurodevelopment of cognitive control functions useful for delay discounting and risk decision making.

How does the brain make a choice between different options when there is no correct or incorrect answer, and that the choice depends entirely upon the subjective value assigned to the different options? For example, if presented with the choice between an apple and an orange, there is no correct or incorrect answer, the choices depend only upon our subjective preferences and our state of hunger.

A classical framework originally developed by Rangel and colleagues (**Figure 1**), proposed that when presented with several options, the brain assigns subjective values to each

of them, then compares between these values and selects the one having the highest subjective value. According to this model of decision-making, value-based decision making can be decomposed into at least 3 stages (Doya, 2008; Rangel et al., 2008; Sugrue et al., 2005) (**Figure 1**). The first stage involves making a representation of the current situation (or state), including the identification of internal state (eg. Hunger), external state (eg. Cold), and potential courses of actions (e.g. purchase food). Second, a valuation system attributes a subjective value to each option under consideration, weighting available options in terms of reward and punishment, as well as cost and benefit. Third, the agent needs to select an action on the basis of this valuation, choosing the option having the highest assigned subjective value. Finally, the chosen action may be re-evaluated based on the actual outcome, eventually leading to updating the other processes through learning to improve subsequent decisions. Although these processes may occur in parallel, this simplified framework is nonetheless useful to decompose basic computations performed by the brain. This framework shows that the brain has to perform multiple computations to make simple decisions. It also indicates that valuation is a key stage for the subsequent selection of a given option.

It is still unclear whether there are separate valuation systems in the brain, however a number of studies distinguish between at least two systems: Pavlovian and instrumental conditioning. In Pavlovian (or classical) conditioning, subjects learn to predict outcomes without having the opportunity to act. In instrumental conditioning, animals learn to choose actions to obtain rewards and avoid punishments. Various strategies are possible, such as optimizing the average rate of acquisition of rewards minus punishments, or optimizing the expected sum of future rewards, where outcomes received in the far future are discounted compared with outcomes received more immediately.



Figure 1. According to Rangel and colleagues (2008), value-based decision making can be decomposed into five basic processes: the construction of a representation of the decision problem, the valuation of the different actions under consideration, the selection of one of the actions on the basis of their valuations, the desirability of the outcomes that follow and finally the update of other processes to improve the quality of future decisions.

The concept of Common Neural Currency in the brain

Our behaviour is motivated by rewards of different nature among which we frequently need to choose. Because there is no single sense organ transducing rewards of different types, our brain must integrate and compare them to choose the options with the highest subjective value. It has been proposed that the brain may uses a 'common reward currency' that can be used as a common scale to value diverse behavioral acts and sensory stimuli (Sugrue et al., 2005). The need for this common currency arises from the variety of choice we are facing in our daily life. Should I go to a movie or to a restaurant tonight? In order to make a choice, our brain must be able to compare the values associated with each option.

Based on the 'common currency' concept, there should be a common brain network coding for different types of goods. Many fMRI studies are consistent with this idea, since common brain structures are involved in reward processing, regardless of reward nature. For example, increased midbrain, ventral striatum and orbitofrontal activities have been observed with different types of rewards, such as monetary gains (Abler et al., 2006; Dreher et al., 2006; J. P. O'Doherty, 2004), pleasant taste (McClure et al., 2003; J. O'Doherty, 2003; J. P. O'Doherty et al., 2003), beautiful faces (Bray & O'Doherty, 2007; Winston et al., 2007) as well as pain relief (Seymour et al., 2004, 2005, 2007). All these neuroimaging studies only investigated one reinforcer at a time and did not compare any two of these reinforcers directly. Subsequent studies using fMRI in humans and electrophysiology in monkeys have identified the common and distinct brain networks involved when making decisions between different options (Sescousse and Dreher, 2010; Sescousse et al., 2013). Together, the studies reviewed above indicate that the human ventromedial prefrontal cortex, together with the ventral striatum is involved in encoding subjective value signals, consistent with the common currency hypothesis.

Delay discounting: neurodevelopmental studies and implications for education

When deciding to engage in a given action, our choice is guided both by the prospect of reward and by the costs that this action entails. Psychological and economic studies have shown that outcome values are reduced when we are obliged to wait for them, an effect known as delay discounting. A classical experiment developed by Walter Mischel at Stanford University is known as the marshmallow experiment t[1]. In this experiment, a child had to choose between one small but immediate reward, or two small rewards if they wait for a longer period of time. The reward was either a marshmallow or pretzel stick, depending on the child's preference. The ability to wait a period of time to receive more rewards can be seen as being more patient, an indication of cognitive control to refrain from the impulsive act of choosing immediate rewards. Cognitive control refers to the ability to override our impulses and to make decisions based on our goals, rather than our habits. Numerous studies have pointed the crucial role of cognitive control in academic achievement (Duckworth et al., 2019). Individual differences in cognitive control reliably predict academic attainment, course grades and performance on standardized achievement tests. Of course, the predictive power of cognitive control for academic achievement is not unique. There are other important factors, including socioeconomic status, general intelligence, motivation and study skills that contribute to explain academic achievement. However, cognitive control, and in particular the ability to resist immediate rewards to wait for delayed rewards is a robust measure across a number of academic outcomes. For example, children who are able to wait longer for the preferred rewards tend to have better life outcomes, as measured by SAT scores, educational attainment, or body mass index (BMI) (Duckworth et al., 2010; Mischel et al., 1988). The same pattern holds for other tasks requiring inhibition of automatic responses, sustained attention and keeping instructions in working memory. This suggests that the same brain networks responsible for cognitive control, such as the dorsolateral prefrontal cortex may be at play across domains to inhibit impulsive behavior.

Many daily choices in students may reflect a choice/dilemma between academic goals that they value in the long run (eg. Doing one's homework to study to become a doctor), and non-academic goals that they find more gratifying in the moment (eg. Going outside with friends without doing one's homework). One key question is to know whether choosing the delayed option (staying at home to do one's homework) really reflects higher cognitive control ability or simply individual preferences for studying. Similarly, does a child who chose the delayed reward of two marshmallows have especially strong cognitive control or does he simply have stronger preference of marshmallows? Observation of behavior alone may not be sufficient to answer the question, neuroimaging studies from the field of neuroeconomics may help.

Early fMRI findings on delay-discounting supported that there may be two separate systems in the brain: a limbic system compution of the value of rewards delivered immediately or in the near future based on a small discount factor, and a cortical system computing the value of distant rewards based on a high discount factor (McClure et al., 2003; Schweighofer et al., 2007, 2008; Tanaka et al., 2004). Discounting would result from the interaction of these two systems associated with different value-signals. More recent studies indicate that there is a single valuation system discounting all future rewards (Kable & Glimcher, 2007). In this study, adult participants were asked to make choices between a small amount of money available now and a larger amount of money available in a few days or weeks. One problem with this type of task is that the choices are quite abstract and hypothetical since the rewards are neither delivered immediately inside the scanner, nor do participants need to wait really for long period of time while laying inside the scanner. To remediate to this problem, we have developed similar delayed discounting paradigms with primary rewards (erotic stimuli) and real delay to experience inside the scanner (Prévost et al., 2008). Our fMRI findings revealed that a similar ventromedial prefrontal cortexventral striatum brain network was engaged for delayed discounting of primary rewards. These results are consistent with the common currency hypothesis, suggesting that there is a single valuation system that discounts future rewards (Kable & Glimcher, 2007). These results, showing that the valuation system, (vmPFC/ventral striatum), implements delay-discounting is also consistent with the existence of a parallel brain system responsible for cognitive control (eg. Dorsolateral prefrontal cortex-intra parietal cortex) to inhibit impulsive behavior: fronto-parietal network (Figure 2). An important question that remains is to understand how these two brain systems (cognitive control vs valuation brain system) communicate and interact when there is a need to inhibit impulsive or habitual behavior ?



Figure 2. Scheme showing that the vmPFC/ventral striatum implements a valuation system under cognitive control of a brain system composed of the dorsolateral prefrontal cortex-intra parietal cortex whichinhibits impulsive behavior.

Adolescents and impulsivity: insights from the neural bases of delay discounting

Our understanding of the brain mechanisms underpinning delay discounting in adults have important practical implications to understand adolescent impulsivity. Many behavioral studies have reported higher impulsivity in adolescence. Early neuroscience studies suggested that adolescent impulsivity could be attributed to the immaturity of the prefrontal cortex, necessary to exert cognitive control over urges originating in the limbic system. There are changes over the adolescent decade in both gray and white matter volumes in prefrontal regions, suggesting that synaptic pruning and myelination were enabling more efficient and more effective self-regulation.

More recent studies suggest that adolescent impulsivity is not simply due to immaturity of the prefrontal cortex, which subserves cognitive control, but is accompanied by a temporary intensification of urges to pursue novel and rewarding experiences (REFS). This "maturational imbalance" view on adolescent impulsivity has guided the study of adolescent risk-taking behavior. According to this view, adolescents' disposition toward risk is because of a maturational imbalance between a brain network involved in cognitive control and goal-directed behavior and one involved in affective processes, including the anticipation and valuation of rewards. There would be a rapid development of the reward system shortly after puberty, that produces increased sensitivity to reward and which declines through late adolescence. In contrast, structures of the cognitive control brain network that inhibit impulses and direct motivation toward goal-relevant behaviors continue to develop until many years later in the 30s. Yet, the question of knowing whether choosing an immediate reward is due to a cognitive control difficulty in regulating one's desires in teenagers or to a higher sensitivity to rewards, that is more prominent during adolescence, remains. A recent fMRI study in adolescents shed light on this question (Bos et al., 2015). It used a delayed discounting experiments in which teenagers had to choose between a small monetary reward given sooner and a greater one given later. Adolescents were more likely than adults to opt for the smaller rewards sooner than larger ones later (Figure 3). This teenagers' tendency to discount the future may be due to greater cognitive control, higher reward sensitivity or both. Intertemporal preferences were correlated with self-reported future orientation but not hedonism. The choice of delayed reward option over the immediate reward was associated with increased engagement of the frontoparietal cognitive control network and, importantly, improvements in frontostriatal connectivity mediated the link between age and intertemporal preferences. That is, as people get older, there is increased communication between the dorsolateral prefontal cortex (dlPFC) and the striatum, and older participants, who have stronger connectivity between the dlPFC and the striatum, were less impulsive. Thus the increased impatience of adolescents' may be driven by weak cognitive control relative to adults, rather than heightened reward sensitivity. This does not mean that impatient behavior is not the byproduct of poor cognitive control and high reward sensitivity, but that there may be an imbalance between the two brain systems (cognitive control versus the valuation brain system) during adolescence.



Figure 3. When comparing adolescents to adults, teenagers were more likely to opt for smaller rewards sooner than larger ones later. As people get older, there is increased communication between the dorsolateral prefontal cortex (dlPFC) and the striatum, and older participants, who have stronger connectivity between the dlPFC and the striatum, were less impulsive. The increased impatience of adolescents' may thus be driven by weak cognitive control relative to adults (Bos et al., 2015).

Improving cognitive control functions through Working Memory (WM) training

Given the importance of self-control for academic achievement, intervention programs aiming to improve self-control in students is greatly needed. Identification of the brain changes occurring during such intervention programs is also crucial. One interesting perspective to understand the communication between the cognitive control brain system and the valuation brain system comes from recent findings concerning working memory, the capacity to maintain and manipulate information along a short period of time. When seeking to represent a delayed reward (eg. Becoming a doctor, receiving 2 marshmallows after waiting), working memory is needed to maintain this representation active. Some findings have shown that people having higher working memory capacity are also more patient and make less impulsive choices in delay discounting tasks. Perhaps more importantly, working memory performance can be trained through intervention programs. In a series of elegant studies, Klingberg, showed that working memory training during 14h per week during 5 weeks, increases activity of DLPFC, engaged in cognitive control and in reducing impulsivity. Moreover, working memory training not only increased the activity of neurons in the prefrontal cortex, it also increases the strength of connectivity within prefrontal cortex regions and between the prefrontal cortex and other areas. Neural changes after training are found in cortical areas that process spatial information in WM and attention, potentially providing a basis for transfer to other cognitive and behavioral tasks that rely on spatial WM and spatially selective attention. Therefore, there is a plasticity of the cognitive control brain network with WM training. When we are confronted with an environment that requires exerting our working memory abilities, the cognitive control brain system becomes more plastic and is able to modulate the valuation system engaged in being more patient. The brain becomes less sensitive to immediate rewards and more tolerant to frustration and waiting. In addition to this neuroplasticity at the system level, there are also changes in functional connectivity that occurs at rest between frontal and parietal regions associated to WM training (Constantinidis & Klingberg, 2016; Jolles et al., 2013; Thompson et al., 2016). These changes in connectivity can be also observed with

other technics than fMRI. For example, transcranial magnetic stimulation (TMS) applied over the parietal cortex propagates over the cortex, as determined by EEG. After WM training, TMS increased signals in the frontal and temporal lobe, demonstrating that training led to an increase in functional connectivity during task performance. The mechanisms underlying changes in functional connectivity could be a stronger synaptic connectivity between neurons or an activity-dependent increase in the myelination of the connecting axons (Gibson et al., 2014; Yeung et al., 2014). Consistent with this latter point, measures of white-matter volume and structure are associated with WM capacity, and there is evidence of increased white-matter density in the parietal lobe after WM training (Takeuchi et al., 2010).

Conversely, what happens when working memory capacity is reduced rather than trained? This can easily be manipulated by increasing WM by having to maintain a distracting series of numbers in memory. In this case, people became more inclined to choose immediate over delayed rewards. These results confirm that 'free' working memory capacity is needed for efficient cognitive control. In addition, children performing a delayed discounting experiment with progressively increasing delays, tend to choose the delayed option more frequently than the immediate reward option. This type of training has also been shown to modify the brain system engaged in cognitive control.



Figure 4. Summary of factors underlying training-induced increases in capacity. Training of working memory (WM) leads to a larger number of prefrontal cortex (PFC) neurons with delay activity and higher firing frequency during the delay period, stronger fronto-parietal functional connectivity with higher WM capacity and WM training. (Copyright Constandinis 2016).

Conclusions

Neuroeconomics had made great progress in the past few years to extend the study of value-based decision making (eg. Delay discounting, risky decision making, loss aversion, ...) to populations beyond the healthy adult population, including adolescents, older adults and sometimes even children. Neurodevelopmental studies help to understand the emergence and neural underpinnings of increased impulsivity and risk seeking in adolescents. Importantly, intervention programs such as working memory training programs provide evidence that the brain systems underlying this function are plastic. Improvements in working memory have been reported for a variety of populations including typically developing pre-school- and primary-school-

aged children, children with Attention Deficit Hyperactivity Disorder (ADHD) who display elevated levels of hyperactive and inattentive behavior, and in children with poor working memory. These findings obtained from research studies have started to be successfully extended by teachers in the school environment (Holmes & Gathercole, 2014).

References

- Abler, B., Walter, H., Erk, S., Kammerer, H., & Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage*, *31*, 790–795.
- Bos, W. van den, Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765–E3774.

https://doi.org/10.1073/pnas.1423095112

- Bray, S., & O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *J Neurophysiol*, 97, 3036–3045.
- Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, *17*(7), 438–449. https://doi.org/10.1038/nrn.2016.43
- Doya, K. (2008). Modulators of decision making. Nat Neurosci, 11, 410–416.
- Dreher, J. C., Kohn, P., & Berman, K. F. (2006). Neural coding of distinct statistical properties of reward information in humans. *Cereb Cortex*, *16*, 561–573.
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-Control and Academic Achievement. *Annual Review of Psychology*, *70*(1), 373–399. https://doi.org/10.1146/annurev-psych-010418-103230
- Duckworth, A. L., Tsukayama, E., & Geier, A. B. (2010). Self-controlled children stay leaner in the transition to adolescence. *Appetite*, *54*(2), 304–308. https://doi.org/10.1016/j.appet.2009.11.016
- Gibson, E. M., Purger, D., Mount, C. W., Goldstein, A. K., Lin, G. L., Wood, L. S., Inema, I., Miller, S. E., Bieri, G., Zuchero, J. B., Barres, B. A., Woo, P. J., Vogel, H., & Monje, M.

(2014). Neuronal Activity Promotes Oligodendrogenesis and Adaptive Myelination in the Mammalian Brain. *Science*. https://doi.org/10.1126/science.1252304

- Holmes, J., & Gathercole, S. E. (2014). Taking working memory training from the laboratory into schools. *Educational Psychology*, 34(4), 440–450. https://doi.org/10.1080/01443410.2013.797338
- Jolles, D. D., van Buchem, M. A., Crone, E. A., & Rombouts, S. A. R. B. (2013). Functional brain connectivity at rest changes after working memory training. *Human Brain Mapping*, 34(2), 396–406. https://doi.org/10.1002/hbm.21444
- Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625–1633. https://doi.org/10.1038/nn2007
- McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*, 339–346.
- Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*(4), 687–696. https://doi.org/10.1037/0022-3514.54.4.687
- O'Doherty, J. (2003). Can't learn without you: Predictive value coding in orbitofrontal cortex requires the basolateral amygdala. *Neuron*, *39*, 731–733.
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Curr Opin Neurobiol*, *14*, 769–776.
- O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*, 329–337.
- Prévost, C., Pessiglione, M., Cléry-Melin, M.-C., & Dreher, J.-C. (2008). Delay versus effortdiscounting in the human brain. In *Society for Neuroscience*.

- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, 9(7), 545–556. https://doi.org/10.1038/nrn2357
- Schweighofer, N., Bertin, M., Shishida, K., Okamoto, Y., Tanaka, S. C., Yamawaki, S., & Doya,
 K. (2008). Low-serotonin levels increase delayed reward discounting in humans. *J Neurosci*, 28, 4528–4532.
- Schweighofer, N., Tanaka, S. C., & Doya, K. (2007). Serotonin and the evaluation of future rewards: Theory, experiments, and possible neural mechanisms. *Ann N Y Acad Sci*, *1104*, 289–300.
- Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K.
 J., & Frackowiak, R. S. (2004). Temporal difference models describe higher-order
 learning in humans. *Nature*, *429*, 664–667.
- Seymour, B., O'Doherty, J. P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., & Dolan,
 R. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of
 pain relief. *Nat Neurosci*, *8*, 1234–1240.
- Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nat Rev Neurosci*, *8*, 300–311.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nat Rev Neurosci*, *6*, 363–375.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., Yamanouchi,
T., Suzuki, S., & Kawashima, R. (2010). Training of Working Memory Impacts Structural
Connectivity. *Journal of Neuroscience*, *30*(9), 3297–3303.
https://doi.org/10.1523/JNEUROSCI.4611-09.2010

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci*, 7, 887–893.

- Thompson, T. W., Waskom, M. L., & Gabrieli, J. D. E. (2016). Intensive Working Memory Training Produces Functional Changes in Large-scale Frontoparietal Networks. *Journal of Cognitive Neuroscience*, *28*(4), 575–588. https://doi.org/10.1162/jocn a 00916
- Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, *45*, 195–206.
- Yeung, M. S. Y., Zdunek, S., Bergmann, O., Bernard, S., Salehpour, M., Alkass, K., Perl, S.,
 Tisdale, J., Possnert, G., Brundin, L., Druid, H., & Frisén, J. (2014). Dynamics of
 Oligodendrocyte Generation and Myelination in the Human Brain. *Cell*, *159*(4), 766–774.
 https://doi.org/10.1016/j.cell.2014.10.011