



Published in final edited form as:

J Neurosci Res. 2020 June ; 98(6): 998–1006. doi:10.1002/jnr.24545.

Fractionating the all or none definition of goal-directed and habitual decision-making

Drew C. Schreiner¹, Rafael Renteria¹, Christina M. Gremel^{*,1,2}

¹Department of Psychology, University of California San Diego, La Jolla California, 92093, United States

²Neurosciences Graduate Program, University of California San Diego, La Jolla California, 92093, United States

Abstract

Goal-directed and habitual decision-making are fundamental processes that support ongoing adaptive behavior. There is a growing interest in examining their disruption in psychiatric disease, often with a focus on a disease shifting control from one process to the other, usually a shift from goal-directed to habitual control. However, several different experimental procedures can be used to probe whether decision-making is under goal-directed or habitual control, including outcome devaluation and contingency degradation. These different experimental procedures may recruit diverse behavioral and neural processes. Thus, there are potentially many opportunities for these disease phenotypes to manifest as alterations to both goal-directed and habitual control. In this review, we highlight examples of behavioral and neural circuit divergence and similarity, and suggest that interpretation based on behavioral processes recruited during testing may leave more room for goal-directed and habitual decision-making to co-exist. Furthermore, this may improve our understanding of precisely *what* the involved neural mechanisms underlying aspects of goal-directed and habitual behavior are, as well as how disease affects behavior and these circuits.

Keywords

decision-making; goal-directed; habits; outcome devaluation; contingency degradation

Introduction

Within the past decade, there has been growing interest and success in examining psychiatric conditions through the lens of instrumental control processes gone awry, namely transitions

*Corresponding Author: Christina M. Gremel, University of California, San Diego, 9500 Gilman Drive #109, La Jolla, CA, 92104, Phone: 858-534-8933, cgremel@ucsd.edu.

Author Contributions

Conceptualization, D.C.S., R.R., and C.M.G.; *Writing – Original Draft*, D.C.S., R.R., and C.M.G.; *Writing – Review & Editing*, D.C.S., R.R., and C.M.G.; *Funding Acquisition*, C.M.G., and R.R.

Conflict of Interest Statement

The authors have no conflict of interest to declare.

Data Accessibility

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

between goal-directed and habitual decision-making processes. This is in part due to the elegant work delineating experimentally-defined behavioral definitions, as well as the identification of distinct and separable cortico-basal ganglia loops supporting each process. These foundations have provided avenues with which to investigate how disease may target one or the other of these decision-processes. At the same time, many have found the hypothesis that the decision-making process is under *either* goal-directed *or* habitual control unsatisfying.

For example, a growing literature suggests that drug dependence produces a bias towards reliance on habitual decision-making processes (Everitt and Robbins, 2005; 2016; Gremel and Lovinger, 2017; Hogarth et al., 2012). At the same time, other reports in the literature (e.g. Ersche et al., 2016; Hogarth et al., 2018) suggest that addicts may be goal-directed in some aspects of their drug-seeking and drug-taking behaviors, and compulsive in others. This understandable frustration has led to both a disregard for the habit hypothesis, as well as to the further development of habit hypotheses; for example, that goal-directed and habitual processes may exist in a hierarchical selection framework, i.e. one could habitually select goal-directed actions to execute (e.g. Cushman and Morris, 2015) or that there might be goal-directed selection of habitual action sequences (Dezfouli and Balleine, 2012). However, even a hierarchical framework still suggests that the measured behavior could be goal-directed or habitual in its entirety, leaving us once again in this unsatisfactory position. Here we suggest that restricting interpretations to the actual experimental manipulations performed may provide more space to identify aspects of differing decision-making processes which may coexist.

As goal-directed control is a widely-used descriptor in neuroscience research, it is important to first briefly review what the accepted instrumental definitions of goal-directed and habitual decision-making processes encompass. The initial definition of goal-directed control stipulated that an action should be sensitive to *both* outcome value and contingency (e.g. Dickinson and Balleine, 1994). When under goal-directed control, there is an explicit use of the goal (or outcome) representation, and the relationship (or contingency) between the action and its outcome. In contrast, habitual actions are made with less dependence on the value of the outcome, and are relatively insensitive to the contingency between the action and the outcome. This highlights an unsatisfying aspect of habitual control; it is commonly defined as a loss of goal-directed control. However, these definitions have been explicitly operationalized via tests which manipulate either outcome value through outcome devaluation tests (Adams and Dickinson, 1981) or the action-outcome contingency (Dickinson et al., 1998) through contingency degradation, omission, and extinction testing. This does provide the advantage of being experimentally defined behaviors one can probe.

The majority of current studies investigate *either* outcome value *or* action-outcome contingency. This is potentially problematic. While goal-directed definitions arose through experimental psychological analysis of behavior, where sensitivity to outcome devaluation and contingency degradation could both be easily observed, the responsible neural mechanisms could plausibly differ. Indeed, there is a dissociation between the neural mechanisms of these two processes described in the literature. For instance, one of the earliest studies investigating the neural substrates of goal-directed and habitual decision-

making found that insular cortex lesions impaired sensitivity only to outcome devaluation and not contingency degradation (Balleine and Dickinson, 1998b). Had only one of these tests been conducted, insular cortex might either have been deemed necessary or unimportant for goal-directed actions. Even within these two tests, multiple processes contribute, and disruption of any of these processes could alter the measurement of goal-directed control. For example, the associative structure of outcome devaluation can be influenced by sensory, motivational, memory, retrieval, performance, and contingency processes. Modern neuroscience has demonstrated distributed neural circuits with cellular and projection specificity contributing to these behavioral concepts (e.g., Parkes and Balleine, 2013). Knowing *what* behavior specifically is disrupted, or *how* a neural circuit contributes would be enormously informative in teasing out specific disruptions in decision-making.

Below we use examples of separable behavioral and neural mechanisms for outcome devaluation and contingency degradation/reversal to highlight how each of these systems could contribute to decision-control phenotypes. However, we want to emphasize that within these two tests, there are still multiple processes contributing, and observed behaviors could and should be understood at a more reduced level. Overall our suggestion, which is not novel, is that specific measurable behaviors in the same animal, occurring at the same or similar time, may show aspects of what has been behaviorally termed goal-directed *and* habitual control processes.

Outcome Devaluation and Contingency Degradation Testing

Testing for goal-directed or habitual control is often done using outcome devaluation and contingency degradation testing. However, there is more than one way to conduct these tests. For example, outcome devaluation can be used to probe goal-directed control using either sensory-specific satiety (Adams and Dickinson, 1981), or via pairing the outcome with an aversive state such as lithium chloride injection (Adams, 1982). Though both manipulations reduce outcome value, they rely on different experimental procedures to achieve this effect. Sensory-specific satiety is achieved through pre-feeding with the outcome previously earned by lever pressing and is usually referred to as a Devalued State. Actions performed in the Devalued State are then compared to actions performed by the same subject in what is termed a Valued State, where the subject has been satiated on an outcome not associated with lever pressing. The Valued State is used to control for the effects of general satiation, thereby allowing for the assessment of if and how outcome value controls decision-making. Goal-directed control by definition should produce reduced responding in the Devalued compared to Valued State. Sensory-specific satiety requires that an animal be sensitive to its hunger or motivational state for a particular outcome, and retrieve and use this reduction in hunger to update the value of the action aimed at procuring that particular outcome. If my goal-directed self just ate a lot of cookies, I would no longer work for cookies, but I would still drink some milk. However, if I habitually ate cookies, I would still work for more cookies.

In contrast, to achieve outcome devaluation via aversive pairing, immediately following exposure to the outcome previously earned by the action, an aversive state is induced generally via a lithium chloride injection (Adams, 1982). Often this outcome-aversive

pairing is repeated until the subject learns the new association between outcome and aversive state. Unlike sensory-specific satiety which usually relies on a within-subject comparison, aversive pairings are generally between groups. Actions performed by the Devalued Group are compared to actions performed by a Control Group that experienced the same aversive state, only not explicitly paired with the outcome. To achieve outcome devaluation with aversive pairings, the subject first has to learn a new association between the outcome and the aversive state. This new aversive association then needs to be retrieved and used to decide whether to direct action towards gaining access to the outcome, or not. After I've been conditioned to associate cookies with intense gastric distress, my goal-directed self will no longer work for cookies, even if I'm hungry. But if I habitually consume cookies, I will still work to obtain more.

Not only are the experimental procedures used to achieve outcome devaluation different, but the behavioral mechanisms are also different. Though both manipulations reduce outcome value, the motivational (hunger) and associative types of devaluation may rely on distinct mechanisms. This should be kept in mind when they are used to test dysfunction in psychiatric disorders such as addiction where drug-seeking actions are often framed as compulsive or insensitive to negative or aversive consequences.

Testing the action-outcome contingency can be done using multiple experimental procedures as well. Contingency degradation, omission, and reversal testing have all historically been used to examine whether subjects can adapt their behavior when there is a contingency change. In contingency degradation (e.g. Balleine and Dickinson, 1998a), non-contingent reward is given in addition to contingent reward (cookies come for free). This erodes the relationship between the action and its outcome. A more extreme variant of this is in extinction testing (work does not produce cookies) or the reversal/omission procedure, where an action performed actually delays the outcome (Dickinson et al., 1998) (I need to *not* work in order for cookies to be delivered). Sensitivity to contingency alteration requires that an animal first recognize a change has occurred and then implement an appropriate change in its behavior. However, the above-mentioned tests do this in different ways. Unexpected outcome deliveries following degradation/reversal may be used to update the animal's model of the environment (Sutton and Barto, 1998), while extinction learning involves novel learning that the lever press no longer produces the outcome (e.g. Bouton, 2002). Any combination of several factors could contribute to sensitivity to contingency, including a loss of flexibility, an inability to remember one's own actions, an inability to represent the contiguity between actions and outcomes, or a limited representation of the environment (e.g. Dutech et al., 2011).

Importantly, although sensitivity to contingency degradation/omission relies on knowledge of the action-outcome contingency (and the ability to update this contingency), it does not seem to be sensitive to the value of the outcome. Devaluation was found to have no effect on sensitivity to an omission test (Dickinson et al., 1998). This highlights the independence of these two tests for goal-directed control. Specifically, that contingency sensitivity is directly affected by the action-outcome association *without* respect to how valued that outcome is.

In addition to these different testing parameters, different training parameters also influence sensitivity to outcome devaluation and contingency alteration. Interestingly, the duration of training can affect the ability to observe goal-directed control, with observed biases towards habitual control given extended training (Adams, 1982). Thus, when you probe behavior may dictate the degree of goal-directed control observed. In addition, the type of schedule used can bias towards a particular control type. Random or variable ratio schedules are often used to bias sensitivity to outcome devaluation and contingency alteration (Dickinson et al., 1983), perhaps due to the underlying relationship between response rate and reward rate (Dickinson, 1985). On the other hand, variable or random interval schedules are often used to bias insensitivity to outcome devaluation and contingency alteration (Dickinson et al., 1983), with the relative degree of temporal uncertainty affecting sensitivity to both outcome devaluation and to contingency reversal or omission testing (De Russo et al., 2010). Importantly, the introduction of choice in the form of multiple response-outcome associations appears to bias away from habitual control (e.g., Colwill and Rescorla, 1985). Thus, two-lever, two-outcome procedures seem to decrease the ability to evaluate habitual processes. Further, recent work suggests that in some scenarios the training of lever press sequences may leave actions sensitive to outcome devaluation, even after extended experience (Garr and Delamater, 2019). In short, *how* something is learned can affect *what* is learned, and therefore, different types of training can engage different neural and behavioral processes.

Sensitivity to the action-outcome contingency and to outcome value can involve many separable behavioral processes. It is therefore likely that the neural mechanisms and circuits responsible for these behaviors may differ. It is outside the scope of this review to cover all the involved neural mechanisms of goal-directed and habitual decision-making. Instead, below we use a cortical area as a case study to highlight the complexity in examining neural mechanisms for contingency and value sensitivity.

Complex Behavioral Mechanisms Underlying Outcome Devaluation and Contingency Degradation: A Prelimbic Case Study

Prelimbic cortex (PLC) is canonically necessary for goal-directed action. However, the literature on what precisely PLC contributes to goal-directed decision-making is surprisingly complex, and serves as a useful case study of and argument for the interrogation of specific behavioral processes. Pre-training lesions of PLC impair sensitivity to outcome devaluation as seen in extinction testing (Balleine and Dickinson, 1998b; Corbit and Balleine, 2003; Killcross and Coutureau, 2003). This finding would support classification of PLC as supporting goal-directed control. However, if the devalued lever presses produced the outcome during a rewarded test, then PLC lesioned rats showed appropriate devaluation (Corbit and Balleine, 2003), suggesting action-outcome encoding was actually intact and could be used as long as the outcome was present. Also unclear is how pre-training PLC lesions affect contingency degradation. Initial studies found that contingency degradation reduced responding for *both* the degraded and non-degraded action, interpreted as insensitivity to contingency degradation (Balleine and Dickinson, 1998b). However, in contrast Corbit and Balleine (2003) find that PLC lesions selectively *increased* responding of

the degraded action during contingency degradation. When the same PLC lesioned animals that had undergone contingency degradation were then tested under extinction conditions, they show similar reductions in both non-degraded and degraded actions supporting the previous finding of an insensitivity to contingency degradation. After additional experiments the authors suggest that action outcome encoding is intact in PLC lesioned rats, but that lesions result in a working memory deficit. Further complicating the story, lesions of medial frontal Prefrontal Cortex (mPFC) that included PLC were found to impair sensitivity to contingency *degradation*, but not to contingency *reversal/omission*, and these animals were still sensitive to action-outcome contiguity (Coutureau et al., 2012). Though it should be noted that these lesions extended into infralimbic cortex, a region canonically involved in habit learning (e.g. Coutureau and Killcross, 2003). Thus, the authors proposed PLC may be required when actions become unrelated (but not inversely related) to their outcome, distinct from a working memory hypothesis. In combination with a modeling paper (Dutech et al., 2011), Coutureau and colleagues (2012) propose that PLC/mPFC may help encode the precise temporal relationships between actions and outcomes in order to assign causal status. This might, in part, be subserved via prediction errors mediated by dopaminergic projections into PLC.

Additional evidence for PLC's complex contribution to goal-directed control comes from more targeted manipulations. Lesions of dopaminergic terminals in PLC impairs sensitivity only to contingency degradation but not to outcome devaluation (Naneix et al., 2009). This same group found that adolescent rats were sensitive to outcome devaluation, but insensitive to contingency degradation, an effect that they attribute to maturation of the mPFC dopaminergic system (Naneix et al., 2012). However, in another study, PLC dopamine lesions impair sensitivity to both outcome devaluation and contingency degradation (Lex and Hauber, 2010). Importantly, this discrepancy could arise from differences in experimental procedures; whereas Naneix and colleagues (2009) used aversive pairing, Lex and Hauber (2010) utilized outcome specific satiety. Thus, here an apparent dissociation in the literature may in fact be due to the use of different methodologies for devaluation and the different behavioral mechanisms they recruit. Recognizing this discrepancy can provide useful insight about the role of involved neural circuits; this pattern of results indicates that prelimbic dopamine may not contribute to the use of aversive information to update value, while it may participate in updating action values in response to changes in internal motivation. It is also important to note that different behavioral and neural mechanisms may operate during acquisition of a goal-directed or habitual action vs. the expression of those actions. As an example, PLC is necessary for the acquisition but not the expression of goal-directed action, as assessed via outcome devaluation (Ostlund and Balleine, 2005; Tran-Tu-Yen et al., 2009). Finally, timing and methodology of inactivation are also important tools that can help resolve discrepancies in the literature. Inactivation can be used to separate effects on acquisition versus expression, since compensatory mechanisms and diaschisis can occur with lesions (e.g., Otchy et al., 2015). Furthermore, timing of inactivation may also be used to reveal a role in encoding vs. retrieval of association (e.g., Parkes and Balleine, 2013).

Other Selective Neural Mechanisms

Aside from PLC, several other neural circuits are selectively involved in either outcome value or contingency sensitivity. Dorsal hippocampus, entorhinal cortex, entorhinal projections to dorsal striatum, parafascicular thalamus, parafascicular projections to dorsal medial striatum, and mediodorsal thalamic projections to dmPFC have been implicated as necessary for sensitivity to action-outcome contingency, but not for sensitivity to outcome devaluation (Corbit and Balleine, 2000; Corbit et al. 2002; Lex and Hauber, 2010; Bradfield et al., 2013a; 2013b; Alcaraz et al., 2018). In contrast, lesions of insular cortex impair or disrupt sensitivity only to outcome devaluation (Balleine and Dickinson, 1998b; 2000). There is growing evidence of dissociations between neural mechanisms supporting sensitivity to outcome devaluation and contingency degradation, as well as differing contributing mechanisms within each test depending on the behavioral mechanisms recruited by the specific experimental procedures. Often, discrepancies found in experimental procedures can provide both limitations and interpretations on what contribution a neural circuit is making. With this in mind, it is important to note that there do seem to be shared neural mechanisms supporting outcome devaluation and contingency degradation independent of which experimental procedure is used.

Parallel Systems

The dorsal striatum, the main input nuclei of the basal ganglia contains two regions that show fairly consistent contributions to decision-making control over actions (for a recent review see Peak et al., 2019). In primates, these regions are largely anatomically distinct, with a caudate and a separable putamen. In rodents, where much functional work has been performed, these correspond to the dorsal medial striatum (DMS), and dorsal lateral striatum (DLS) respectively. Using the experimentally defined definitions of instrumental control, the first studies in rats showed that lesioning or inactivating the DMS resulted in a loss of goal-directed control, while habitual responding remains intact (Yin et al., 2005). Conversely, lesions or inactivation of the DLS disrupted habitual actions and reverted actions to goal-directed control (Yin et al., 2004; 2006). This has now been replicated numerous times in rats and mice (e.g. Corbit and Janak, 2010; Hilario et al., 2012; Gremel and Costa, 2013). These observations following sensory-specific satiation as well as aversive pairing for outcome devaluation, and contingency degradation, as well as omission testing, suggest there may be eventual converging neural mechanisms contributing to goal-directed and habitual control. Of course, the topographical distribution of cortical and thalamic projections into dorsal striatum also suggests there may be more localized pockets of selective computations performed on converging inputs (Klaus et al., 2019), and increasing circuit, projection, and cell-type specificity has been useful in identifying particular circuits of decision-making control (Gremel et al., 2016; Renteria et al., 2018).

However, an important point to make is that goal-directed and habitual action control are two fundamental strategies, either of which can control learning and performance of decision-making. Behavior may be biased more towards goal-directed control in some situations, while in other cases habitual control may come to dominate. Decision-making control develops in parallel with goal-directed and habit circuits concurrently active during

instrumental learning, contributing to the continuum of goal-directed and habitual actions often observed (Thorn et al., 2010; Gremel and Costa, 2013). Though it is often proposed that decision-making progresses from initial goal-directed to eventual habitual control (e.g. Adams, 1982), the DMS is not necessary for the acquisition of instrumental actions, with DLS able to support new action learning (Yin et al., 2005; Hilario et al., 2012; Gremel and Costa, 2013). Hence, evidence to date suggests that action-outcome representations do not need to be transferred from DMS to DLS for the DLS to support habitual control over decision-making. In spirit with the above arguments, this means that impairments in some aspects of goal-directed processes, whether it be sensory-specific satiety or aversive pairings for example, will not prevent other systems from being able to acquire habitual control over decision-making.

Furthermore, it should be noted that the current all-or-nothing treatment often makes it quite difficult to assess if, when, and how control over decision-making may shift from predominately goal-directed to predominately habitual. Methodology often classifies behavior as goal-directed until it suddenly transitions to be habitual. Recent works have taken a stab at this by examining the shift between goal-directed and habitual control in the same animal (Gremel and Costa, 2013; Gremel et al., 2016; Renteria et al., 2018). By training the same mouse on both random ratio and random interval schedules delineated by contextual cues, and then assessing the degree of goal-directedness expressed in each context, gradients of goal-directed control have been observed. Furthermore, this approach removes the reliance on group statistics and allows for a within animal assessment as to the degree of goal-directed control.

Highlighting the behavioral and neural circuit divergence or similarity seen in different experimental procedures for decision-making is particularly important, as these procedures and definitions are being applied in the study of disease. These fundamental processes contribute immensely to the support of ongoing behaviors, and their disruption could produce an impaired decision-making phenotype (Balleine and O'Doherty, 2009). As different experimental procedures may recruit different as well as overlapping behavioral and neural processes, there are potentially many opportunities for these disease phenotypes to manifest as alterations to both goal-directed and habitual control.

Addiction

Disrupted decision-making has been associated with numerous psychiatric diseases, including addiction (Gillan et al., 2016). Greater precision in discussing goal-directed and habitual actions would be tremendously beneficial to help understand how underlying behavioral and neural mechanisms may have gone awry.

One prominent theory is that drug addiction progresses from initial goal-directed use to habitual, and finally compulsive use (Everitt and Robbins, 2005; 2016). Indeed, several studies have shown that chronic passive exposure to drugs or alcohol can lead to a shift from goal-directed to habitual control when examining subsequent instrumental learning in withdrawal (e.g. Nelson, 2006; Nordquist et al., 2007; LeBlanc et al., 2013; Corbit et al., 2014; Renteria et al., 2018). Similarly, self-administration of cocaine, nicotine, or alcohol

can produce habitual control (Zapata et al., 2010; Corbit et al., 2012; Clemens et al., 2014), but not always (e.g. Samson et al., 2004; Halbout et al., 2016). However, a contrasting hypothesis is that addicts may seek drug in a very goal-directed manner, and that drug consumption rather than drug seeking might become habitual (Robinson and Berridge, 2008; Singer et al., 2018). Some of this discrepancy might be explained by interchangeably utilizing either of the operational criteria for habits. For instance, a drug dependent individual could be exquisitely sensitive to the instrumental contingency, and shift their actions in a very goal-directed manner to obtain their drug of choice, yet relatively less sensitive to the value of the outcome (or, less sensitive to the negative consequences associated with drug use). In support of this, prior experimenter-delivered cocaine has been found to either reduce (Corbit et al., 2014) or have no effect (Halbout et al., 2016) on sensitivity to outcome devaluation of food reward, but actually increased sensitivity to contingency degradation (Halbout et al. 2016). Prior self-administered cocaine also increased action-outcome encoding in DLS (Burton et al., 2017), and led to animals that were sensitive to devaluation via aversive pairing when a cocaine discriminative stimulus was present, but insensitive when it was absent (Root et al., 2009). Thus in some cases, prior cocaine seems to make animals both more goal-directed (sensitive to contingency) and more habitual (insensitive to value). Similarly, a recent animal study trained rats to solve different puzzles daily for access to cocaine, and found that these animals were quite sensitive to the changing contingencies, but displayed typical hallmarks of addiction including escalation and (in a subset of animals) resistance to shock-induced reductions in cocaine seeking (Singer et al., 2018). Interestingly, a recent study found cocaine-induced facilitation of inflexible, habitual responding specifically for choice of a non-drug reward, highlighting the complex effect habitual facilitation may have (Vandaele et al., 2019). Drug-induced facilitation of habitual responding is not always observed (Halbout et al, 2016; Singer et al., 2018), and it is important to note that certain forms of instrumental training may prevent the emergence of habitual control. Training schedules that involve multiple instrumental actions and outcomes have been shown to remain goal-directed despite extended training (Colwill and Rescorla, 1985), and this could potentially explain the persistent goal-directed control observed in Halbout et al.'s study (2016). This suggests that the neural circuits that mediate habits are not engaged in the same manner and/or these tasks demand such high executive control as to heavily shift the balance in favor of goal-directed responding.

With more investigation into how these decision-making circuits change in relation to drug dependence and use, greater light will be shed on both the behavioral and neural mechanisms altered. An insensitivity to negative consequences observed in addiction has often been framed as a strengthening or biased used of habitual systems. However, given the parallel nature of action control, disrupted decision-making could arise from strengthening of habit systems and/or disruptions to goal-directed systems. Indeed, recent works have suggested that addiction as well as other psychiatric disorders does involve disruption to goal-directed systems (e.g. Ersche et al., 2016; Gillan et al., 2016), while other works have identified a strengthening of habit systems (Sjoerds et al., 2013; Delorme et al., 2016). Further consideration of the behavioral systems recruited by differing experimental procedures opens the door even wider to behavioral and neural systems that may be disrupted in decision-making and actions.

Discussion

Here we reviewed how using different procedures that recruit different behavioral and neural mechanisms to probe goal-directed and habitual control may result in an incomplete picture of involved processes. It is still unclear how goal-directed and habitual decision-making affect decision-making apart from the confirmatory tests for outcome value and contingency control. The focus on these confirmatory tests and the corresponding negative definition of habits (insensitivity to value and contingency) may contribute to the common current all-or-nothing treatment of goal-directed and habitual decision-making. While perturbations to aspects of goal-directed decision-making can be measured, habitual control is often defined as the null hypothesis that outcome devaluation and contingency manipulation are without effect (see recent reviews: Vandaele and Janak, 2017; Watson & de Wit, 2018).

Neuroscience investigations into mechanisms supporting habitual control in related circuits can shed light on how diseases like addiction may affect these circuits, but do not appear to get us closer to understanding what habits are. Until specific behavioral features of habit are identified and can be probed across species, it seems we are left with this dilemma.

Focusing on the specific behavioral processes will allow for some fractionation of goal-directed and habitual decision-making. For instance, a behavior that is sensitive to outcome devaluation yet insensitive to contingency alterations is both positively and negatively defined. Further pinpointing why that behavior is insensitive to contingency alterations (e.g., an inability to encode the temporal relationship between actions and their outcomes) can provide specific characteristics of behavior, and allow for investigation into how this is instantiated in the brain. Furthermore, brain regions involved in decision-making and action control are composed of heterogeneous cell types, projections, and inputs. Behaviors are often attributed to entire brain regions; however, the dynamics of activity within a region can be critical. While single unit and imaging data show that only subsets of cells show coordinated classically responsive activity, non-classically responsive neurons have also been shown to contribute to behavioral relevance (Insanally et al., 2019). In defining the neural circuits that mediate goal-directed and habitual responding we must take into consideration the need for greater specificity at both the systems and cellular level. Focusing on the behavioral and neural mechanisms at play in goal-directed and habitual decision-making may also open the door to investigating how these processes interact with (or overlap with) fundamental decision variables that mediate action selection and performance (Klaus et al., 2019). A greater understanding of how projection/cell-type specific mechanisms interact with local microcircuitry in shaping neural ensemble activity could resolve discrepancies between the behavioral and neural mechanisms underlying goal-directed and habitual decision-making.

Goal-directed and habitual decision-making have been and continue to be fruitful frameworks to investigate decision-making mechanisms and how they might be disrupted in psychiatric disorders. We have argued that a greater focus on the specific behavioral processes at play may help to resolve and reveal discrepancies, both at the level of mechanistic questions (e.g., what does neural circuit X contribute to decision-making?) and especially at the theoretical level (e.g. how do habits contribute to addiction?). We do want to emphasize that even if the concepts of goal-directed and habitual decision-making are

unsatisfying on some levels, they still hold merit. Indeed, there is a great deal of overlap between the operational definitions of goal-directed behavior, with determinants such as training duration, training schedule, various neural manipulations, and various drug exposure regimens biasing sensitivity or insensitivity to both outcome value and action-outcome contingency. However, treating decision-making as an all-or-nothing, winner-takes-all process can hinder progress. This might be especially true in theoretical attempts to understand psychiatric disorders such as addiction, where specific aspects of goal-directed decision-making may be selectively disrupted.

Other Acknowledgements

The authors would like to thank Ege A. Yalcinbas and Emily T. Baltz for constructive comments on the manuscript.

Grant Information: This project was funded by the NIH (4R00AA021780-02-C.M.G., AA026077-01A1-C.M.G., and F32AA026776-R.R.).

References

- Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. Taylor & Francis; 1982;34(2):77–98.
- Adams CD, Dickinson A. Instrumental Responding Following Reinforcer Devaluation. *Q J Exp Psychol-B*. 1981;33:109–21.
- Alcaraz F, Fresno V, Marchand AR, Kremer EJ, Coutureau E, Wolff M. Thalamocortical and corticothalamic pathways differentially contribute to goal-directed behaviors in the rat. *eLife*. 2018 2 6;7.
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998a 4;37(4–5):407–19. [PubMed: 9704982]
- Balleine BW, Dickinson A. The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. *Journal of Neuroscience*. 2000 12 1;20(23):8954–64. [PubMed: 11102506]
- Balleine BW, O'Doherty JP. Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology*. 2009 9 23;35(1):48–69.
- Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW. The Thalamostriatal Pathway and Cholinergic Control of Goal-Directed Action: Interlacing New with Existing Learning in the Striatum. *Neuron*. 2013a 6.
- Bradfield LA, Hart G, Balleine BW. The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. *Front Syst Neurosci*. 2013b;7. [PubMed: 23596400]
- Bouton ME. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry* 2002; 52(10) 976–986. [PubMed: 12437938]
- Burton AC, Bissonette GB, Zhao AC, Patel PK, Roesch MR. Prior Cocaine Self-Administration Increases Response-Outcome Encoding That Is Divorced from Actions Selected in Dorsal Lateral Striatum. *Journal of Neuroscience*. 2017 8 9;37(32):7737–47. [PubMed: 28694335]
- Clemens KJ, Castino MR, Cornish JL, Goodchild AK, Holmes NM. Behavioral and neural substrates of habit formation in rats intravenously self-administering nicotine. *Neuropsychopharmacology*. 2014 10;39(11):2584–93. [PubMed: 24823947]
- Colwill RM, Rescorla RA. Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology*. 1985 Vol.11(1): 120–132.
- Corbit L, Balleine B. The Role of the Hippocampus in Instrumental Conditioning. *Journal of Neuroscience*. 2000 6 1;20(11):4233. [PubMed: 10818159]
- Corbit LH, Balleine BW. The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res*. 2003 11 30;146(1–2):145–57. [PubMed: 14643467]

- Corbit LH, Chieng BC, Balleine BW. Effects of repeated cocaine exposure on habit learning and reversal by N-acetylcysteine. *Neuropsychopharmacology*. 2014 7;39(8):1893–901. [PubMed: 24531561]
- Corbit LH, Janak PH. Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur J Neurosci*. 2010 4;31(7):1312–21. [PubMed: 20345912]
- Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. *Biological Psychiatry*. 2012 9 1;72(5):389–95. [PubMed: 22440617]
- Corbit LH, Ostlund SB, Balleine BW. Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *Journal of Neuroscience*. 2002 12 15;22(24):10976–84. [PubMed: 12486193]
- Coutureau E, Esclassan F, Di Scala G, Marchand AR. The role of the rat medial prefrontal cortex in adapting to changes in instrumental contingency. Zhuang X, editor. *PLoS ONE*. 2012;7(4):e33302. [PubMed: 22496747]
- Coutureau E, Killcross S. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav Brain Res*. 2003 11;146(1–2):167–74. [PubMed: 14643469]
- Cushman F, Morris A. Habitual control of goal selection in humans. *Proceedings of the National Academy of Sciences*. 2015 11 10;112(45):13817–22.
- Delorme C, Salvador A, Valabrègue R, Roze E, Palminteri S, Vidailhet M, et al. Enhanced habit formation in Gilles de la Tourette syndrome. *Brain*. 2016 2;139(Pt 2):605–15. [PubMed: 26490329]
- DeRusso AL, Fan D, Gupta J, Shelest O, Costa RM, Yin HH. Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Frontiers in Integrative Neuroscience*. 2010 5 28;4(17):1–8. [PubMed: 20161992]
- Dezfouli A, Balleine B. Habits, action sequences and reinforcement learning. *European Journal of Neuroscience* 2012 4;35(7):1036–51. [PubMed: 22487034]
- Dickinson A, Nicholas DJ, Adams CD. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology*. 1983;35 B:35–51.
- Actions Dickinson A. and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences*. 1985;308(1135):67–78.
- Dickinson A, Balleine B. Motivational control of goal-directed action. *Animal Learning & Behavior*. 1994 3;22(1):1–18.
- Dickinson A, Squire S., Varga Z, Smith JW. Omission Learning after instrumental pretraining. *The Quarterly Journal of Experimental Psychology: Section B*. 1998; 51(3), 271–286.
- Dutech A, Coutureau E, Marchand AR. A reinforcement learning approach to instrumental contingency degradation in rats. *Journal of Physiology-Paris*. 2011 1;105(1–3):36–44.
- Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LHE, Luijten M, et al. Carrots and sticks fail to change behavior in cocaine addiction. *Science*. 2016 6 17;352(6292):1468–71. [PubMed: 27313048]
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005 11;8(11):1481–9. [PubMed: 16251991]
- Everitt BJ, Robbins TW. Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. *Annu Rev Psychol*. 2016;67(1):23–50. [PubMed: 26253543]
- Garr E, Delamater AR. Exploring the relationship between actions, habits, and automaticity in an action sequence task. *Learning and Memory* 2019;26(4):128–132 [PubMed: 30898975]
- Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*. 2016;5:e94778.
- Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, et al. Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. *Neuron*. 2016 6 15;90(6):1312–24. [PubMed: 27238866]
- Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun*. 2013;4:2264. [PubMed: 23921250]

- Gremel CM, Lovinger DM. Associative and sensorimotor cortico-basal ganglia circuit roles in effects of abused drugs. *Genes Brain Behav.* 1st ed. 2017 1;16(1):71–85. [PubMed: 27457495]
- Halbout B, Liu AT, Ostlund SB. A Closer Look at the Effects of Repeated Cocaine Exposure on Adaptive Decision-Making under Conditions That Promote Goal-Directed Control. *Front Psychiatry.* 2016;7:44. [PubMed: 27047400]
- Hilario M, Holloway T, Jin X, Costa RM. Different dorsal striatum circuits mediate action discrimination and action generalization. *Eur J Neurosci.* 2012 4;35(7):1105–14. [PubMed: 22487040]
- Hogarth I, Balleine BW, Corbit LH, Killcross S. Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Annals of the New York Academy of Sciences.* 2012 11; 1282: 12–24. [PubMed: 23126270]
- Hogarth I, Lam-Cassettari C, Pacitti H, Currah T, Mahlberg J, Hartley L, Moustafa A. Intact goal-directed control in treatment-seeking drug users indexed by outcome-devaluation and Pavlovian to instrumental transfer; critique of habit theory. *European Journal of Neuroscience.* 2018 5; 35: 1–13.
- Insanally MN, Carcea I, Field RE, Rodgers CC, DePasquale B, Rajan K, et al. Spike-timing-dependent ensemble encoding by non-classically responsive cortical neurons. *eLife. eLife Sciences Publications Limited;* 2019 1 28;8:693.
- Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex.* 2003 4;13(4):400–8. [PubMed: 12631569]
- Klaus A, Alves da Silva J, Costa RM. What, If, and When to Move: Basal Ganglia Circuits and Self-Paced Action Initiation. *Annu Rev Neurosci.* 2019 7 8;42:459–83. [PubMed: 31018098]
- LeBlanc KH, Maidment NT, Ostlund SB. Repeated Cocaine Exposure Facilitates the Expression of Incentive Motivation and Induces Habitual Control in Rats. *PLoS ONE.* 2013;8(4).
- Lex B, Hauber W. The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. *Cerebral Cortex.* 2010 4;20(4):873–83. [PubMed: 19605519]
- Naneix F, Marchand AR, Di Scala G, Pape J-R, Coutureau E. A role for medial prefrontal dopaminergic innervation in instrumental conditioning. *Journal of Neuroscience.* 2009 5 20;29(20):6599–606. [PubMed: 19458230]
- Naneix F, Marchand AR, Di Scala G, Pape J-R, Coutureau E. Parallel maturation of goal-directed behavior and dopaminergic systems during adolescence. *Journal of Neuroscience.* 2012 11 14;32(46):16223–32. [PubMed: 23152606]
- Nelson A Amphetamine Exposure Enhances Habit Formation. *Journal of Neuroscience.* 2006 4 5;26(14):3805–12. [PubMed: 16597734]
- Nordquist RE, Voorn P, Malsen JG de M-V, Joosten RNJMA, Pennartz CMA, Vanderschuren LJMJ. Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *Eur Neuropsychopharmacol.* 2007 7;17(8):532–40. [PubMed: 17275266]
- Ostlund SB, Balleine BW. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *Journal of Neuroscience.* 2005 8 24;25(34):7763–70. [PubMed: 16120777]
- Otchy TM, Wolff SBE, Rhee JY, Pehlevan C, Kawai R, Kempf A, Gobes SMH, Ölveczky BP. Acute off-target effects of neural circuit manipulations. *Nature.* 2015 12 17;528: 358–363. [PubMed: 26649821]
- Parks SL, Balleine BW. Incentive memory: evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. *Journal of Neuroscience.* 2013 5 15;33(20): 8753–63. [PubMed: 23678118]
- Peak J, Hart G, Balleine BW. From learning to action: the integration of dorsal striatal input and output pathways in instrumental conditioning. *European Journal of Neuroscience.* John Wiley & Sons, Ltd (10.1111); 2019 3 1;49(5):658–71. [PubMed: 29791051]
- Renteria R, Baltz ET, Gremel CM. Chronic alcohol exposure disrupts top-down control over basal ganglia action selection to produce habits. *Nat Commun.* 2018 1 15;9(1):211. [PubMed: 29335427]

- Robinson TE, Berridge KC. The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. 2008;363(1507):3137–46. [PubMed: 18640920]
- Root DH, Fabbriatore AT, Barker DJ, Ma S, Pawlak AP, West MO. Evidence for habitual and goal-directed behavior following devaluation of cocaine: a multifaceted interpretation of relapse. *PLoS ONE*. 2009 9 25;4(9):e7170. [PubMed: 19779607]
- Samson HH, Cunningham CL, Czachowski CL, Chappell A, Legg B, Shannon E. Devaluation of ethanol reinforcement. *Alcohol*. 2004 4;32(3):203–12. [PubMed: 15282114]
- Singer BF, Fadanelli M, Kawa AB, Robinson TE. Are Cocaine-Seeking “Habits” Necessary for the Development of Addiction-Like Behavior in Rats? *Journal of Neuroscience*. 2018 1 3;38(1):60–73. [PubMed: 29158359]
- Sjoerds Z, de Wit S, van den Brink W, Robbins TW, Beekman ATF, Penninx BWJH, et al. Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Transl Psychiatry*. 2013 12 17;3(12):e337. [PubMed: 24346135]
- Sutton RS, Barto AG. Reinforcement learning: An introduction. 1998 MIT Press
- Thorn CA, Atallah H, Howe M, Graybiel AM. Differential Dynamics of Activity Changes in Dorsolateral and Dorsomedial Striatal Loops during Learning. *Neuron*. 2010 6;66(5):781–95. [PubMed: 20547134]
- Tran-Tu-Yen DAS, Marchand AR, Pape J-R, Di Scala G, Coutureau E. Transient role of the rat prelimbic cortex in goal-directed behaviour. *European Journal of Neuroscience*. 2009 8;30(3):464–71. [PubMed: 19614748]
- Vandaele Y, Vouillac-Mendoza C, Ahmed SH. Inflexible habitual decision-making during hchoice between cocaine and a nondrug alternative. *Translational psychiatry*. 2019 3;9(1):1–11. [PubMed: 30664621]
- Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*. 2004;19(1):181–9. [PubMed: 14750976]
- Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res*. 2006 1 30;166(2):189–96. [PubMed: 16153716]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci*. Blackwell Science Ltd; 2005 7;22(2):513–23. [PubMed: 16045504]
- Zapata A, Minney VL, Shippenberg TS. Shift from Goal-Directed to Habitual Cocaine Seeking after Prolonged Experience in Rats. *J Neurosci*. Society for Neuroscience; 2010 11 17;30(46):15457–63.

Significance Statement

Goal-directed and habitual decision-making are widely studied and applied to a variety of psychiatric disorders. This wide application has led to an expanding or often all-or-nothing definition that may at times obscure the actual involved behavioral and neural processes. This all-or-nothing definition holds particular relevance for disorders such as addiction where a growing literature has provided evidence for both habitual and goal-directed control. Some of this discrepancy may have arisen through treating decision-making as either goal-directed or habitual, without respect to the specific behavioral processes at play.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript