Published in final edited form as:

Behav Brain Res. 2012 December 1; 235(2): 136–142. doi:10.1016/j.bbr.2012.07.042.

Enhanced Dorsolateral Striatal Activity in Drug Use: The Role of Outcome in Stimulus-Response Associations

Noam Schneck¹ and Paul Vezina^{2,*}

¹School of Psychology, Fairleigh Dickinson University, Teaneck, NJ, USA 07666

²Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, IL, USA 60637

Abstract

Prolonged stimulant exposure leads to enhanced dorsolateral striatal (DLS) dopaminergic activity in response to the drug and drug-associated cues. This effect has been interpreted in light of evidence that this brain region supports the generation of habitual stimulus-response (S-R) based behaviors to propose the idea that prolonged drug use leads to the development of drug taking and seeking habits that are insensitive to the value of the rewards they procure. In this review, we discuss evidence supporting a continued role for reward value in the performance of S-R based behaviors. We describe how caching of reward value and Pavlovian to instrumental transfer can provide mechanisms for past and current reward values to regulate the performance of S-R habits. The contribution of these constructs is consistent with evidence indicating the continued interaction between ventral incentive processing and dorsal S-R processing striatal regions in the generation of habitual drug seeking behaviors.

Keywords

Addiction; Behavioral control; Conditioning; Drug self-administration; Incentive regulated habits; Learning; Reward

1. Introduction

A large body of research has demonstrated the sensitizing effects of psychostimulant drugs in the striatal regions of the basal ganglia. Two stages have been observed [1, 2]. Following several days of drug use, drugs and associated cues elicit an enhanced dopaminergic response in the ventral aspect of the striatum (VS). As drug use continues, this enhancement radiates to include the dopamine (DA) projections to the dorsolateral aspect of the striatum (DLS) [1–4]. These effects have proven to be fairly reliable and have been demonstrated in rats and primates as well as in humans [5–9].

These findings have served as the basis for a highly influential theory of habitual drug use [10, 11]. Based on research linking the DLS to the formation and expression of stimulus-

© 2012 Elsevier B.V. All rights reserved.

*Correspondence: Paul Vezina, Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, 5841 S. Maryland Avenue, MC 3077, Chicago, IL 60637, TEL: 773-702-2890, FAX: 773-702-0857, pvezina@yoda.bsd.uchicago.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

response associations [12, 13], this theory argues that drug-induced sensitization of DA overflow in the DLS leads to the emergence of stimulus-response associations that mediate pathological drug seeking and taking.

Dual process theories of behavior posit that two types of associations underlie behavior: stimulus-response (S-R) and response-outcome (R-O) associations. S-R associations differ from R-O associations regarding the role played by reward in controlling behavior. Behaviors based on R-O associations are conceptualized as being directly motivated by the value of the outcome procured by the behavior. In contrast, S-R behaviors are conceptualized as habitual or reflexive responses to environmental stimuli that are not driven by the value of the outcome resulting from the behavior [14, 15].

One readily observable difference between S-R versus R-O based behaviors stems from their differential sensitivity to reward, that is the degree to which the intensity of the behavior emitted shifts in response to changes in the value of the contingent reward. R-O associated behaviors tend to be reward-sensitive while S-R associated behaviors generally are not [14, 15]. If pushing a lever produces food, devaluation of that food will either decrease lever pressing or have no effect depending on the association underlying the behavior. Table 1 provides key terms and definitions as they relate to habitual drug use.

According to the theory of habitual drug use of Everitt and Robbins [10, 11], enhancement of stimulant-induced DA activity in the DLS increases the degree to which drug seeking and taking behaviors are under the control of S-R associations. As a result, these behaviors occur in response to environmental stimuli and are relatively unaffected by the value of the rewards subsequently obtained. A constellation of drug procuring behaviors can conceivably be perceived in this way (e.g. reaching for a crack pipe, cigarette, or alcoholic beverage). These habits are cued by environmental stimuli and occur irrespective of the outcomes that have become linked to drug use. Such habitual reward insensitive behavior is argued to be an essential component of addiction [10, 11, 16]. It has also been suggested that reliance on S-R associations increases the cognitive automaticity of behavior so that it may require fewer attentional resources [7, 8, 10, 11, 16, 17].

These arguments suggest that as drug use progresses, the degree to which both positive and negative outcome value plays a role in the behavioral mediation and cognitive experience of drug seeking and taking declines. This theory has significant ramifications for therapies that try to focus drug users on the negative consequences of the drug use lifestyle. If drug pursuit occurs as a habitual response to environmental cues then these therapies are not likely to be very effective.

Here we argue that outcome value can continue to influence the emission of drug procuring behaviors despite the development of drug induced S-R associations. While increased reliance on S-R associations heralds a shift in the role of reward in mediating behavior, it remains that the value of previous rewards as well as competing values of multiple current rewards can continue to influence S-R behaviors. In the first section of this review, we outline a mechanism by which previously learned reward value may continue to influence behaviors that have become habitual. We also describe a mechanism by which different current reward values associated with multiple reinforcers can influence S-R responding. As it has not yet been demonstrated unequivocally that S-R based behaviors recruit fewer attentional resources, we review in the second section the respective determinants of S-R and automatic behaviors and discuss overlaps and distinctions between these two constructs. Finally, we review relevant neurobiological findings to support the assertion that reward value continues to play a role in chronic habitual drug use. These findings demonstrate maintained interactions between R-O and S-R substrates in striatum.

2. Outcome as a modulator of S-R based behaviors

There are different ways in which outcome can influence S-R behavior. Two different types of outcome value have been proposed: cached and current value [18, 19]. While definitions of cached value vary, common among them is the idea that it indexes a memory of some general aspect of reward associated with a given reinforcer experienced in the past [18–20]. Current value refers to the significance of the actual reinforcer based on a representation of reward as it is relevant to the organism at the present time [21]. The following results suggest that both cached and current reward values may be able to influence the performance of S-R behaviors.

2.1. Role of cached outcome values in S-R responding

Theoretical arguments have been made to suggest that cached outcome value can influence the generation of S-R behavior [18–20]. While these arguments did not provide unequivocal evidence for outcome representation during S-R performance, they did introduce the possibility that memories of the value of reinforcers could come to influence the expression of the respective S-R behaviors that developed in their pursuit. As reviewed below, there is evidence to suggest that a representation of reinforcement outcome may be evoked during S-R behavior and that it can influence its expression.

2.1.1. Cached outcome during S-R performance—In a series of studies, Balleine, Dickinson and their colleagues (e.g., [22, 23]) assessed the effect of exposing rats to a food reinforcer either in a sated or food-deprived state on subsequent responding during extinction. In these experiments, rats were food-deprived and trained to press a lever for food pellets in an operant box. They were then exposed to the food pellets in a feeding cage (with no levers present) when either in a non-deprived or deprived state, thereby establishing a memory of the food pellets as either low or high value respectively. Following this exposure, rats were returned to the operant box and allowed to press a lever in extinction again either in a non-deprived or deprived state. The operant training procedures previously used to train the rats favored the development of S-R behavior in that a random interval schedule was used (Table 1). Consistent with this, the level of deprivation during extinction in and of itself had no effect on performance during this test. Remarkably, however, rats exposed to the food pellets in the feeding cages when sated (leading to lowering of food pellet value) and subsequently tested in extinction when sated showed lower levels of responding on this test compared to the other groups. Rats exposed to the food pellets when deprived (thus maintaining high food pellet value) subsequently showed higher levels of responding during extinction whether deprived or non-deprived on this test. These results suggest that the memory of the low value food pellets learned during exposure in the feeding cage was able to influence responding during the subsequent extinction test. They are consistent with a role for cached outcome value in determining operant S-R responding. In this case, cached low outcome value diminished responding during extinction testing while cached high outcome value maintained responding even in sated rats.

Interestingly, rats tested in a deprived state in extinction following food devaluation did not show lower levels of responding, suggesting that the deprivation state of these animals presented a stimulus sufficient to overcome the effects of the food devaluation in these experiments [23]. This may reflect a competitive balance between the cached value associated with food devaluation on the one hand and the influence of food deprivation on the other. Thus, the ability of the food paired lever to elicit approach and the enhancement of this effect in deprived rats may have overwhelmed cached information about food devaluation in this case. The nature of the factors elicited by food deprivation and the manner in which they interact with cached information remains to be determined.

Nonetheless, the above findings clearly indicate that a representation of the outcome can be evoked during S-R performance, that it can influence S-R performance, and that its effects can be regulated by the state of the animal.

2.1.2. Cached outcome during Pavlovian to Instrumental Transfer—Another mechanism by which outcome relevant information can regulate S-R processes is Pavlovian to instrumental transfer (PIT) [10, 24, 25]. PIT occurs when an S-R behavior is performed in the presence of a conditioned stimulus (CS) that is associated with an affectively meaningful outcome or unconditioned stimulus (UCS) [25]. PIT has at least five components: 1) the operant stimulus (e.g., a lever); 2) the response performed on the operant stimulus (e.g. pressing); 3) the reinforcer produced by the response (e.g. food); 4) a CS in whose presence the behavior is performed; 5) a UCS associated with the CS. The UCS and the reinforcer can be the same stimulus or they can be different depending on the type of PIT studied. Through the CS, the value of the UCS can have an excitatory influence on the instrumental response. In some demonstrations of PIT, the value of the UCS appears to reflect cached rather than current value. For example, it has been shown that a CS linked to a food UCS can continue to enhance an operant response even after the UCS has been devalued [26]. In this study, a CS was paired with a reinforcing UCS that was subsequently devalued. The devaluation reduced the CSs ability to evoke a CR but not its ability to enhance operant responding through PIT. These findings suggest that PIT specifically is mediated by the cached value of the UCS while other types of value may influence other properties of the CS. As a result, it appears that PIT, as measured in this study, provides a mechanism for cached outcome value to influence S-R responding. As described below, this appears to be a characteristic of a type of PIT referred to as general PIT.

2.2. Role of current outcome values in S-R responding

Until this point, instances in which cached outcome value can regulate the expression of S-R behaviors have been described. In this section, we review evidence for additional mechanisms by which S-R responding can be regulated. A number of authors have suggested that current reward value can in fact regulate S-R responding through PIT [21, 26]. This suggestion is particularly important because it implies that S-R behaviors can retain a dynamic interaction with current reward value. According to this possibility, current shifts in reward value can still play an important role in influencing behavior even after S-R associations have been established.

2.2.1. Specific PIT—Two different types of PIT have been proposed (general versus specific) that may differentially engage cached and current reward value. In general PIT, *any* UCS can render a CS capable of generating a PIT signal for any instrumental response for any reinforcer. In specific PIT, the CS can only enhance instrumental behavior that produces a reinforcer identical to the UCS that it was originally paired with. It has been proposed [26] that the two types of PIT encode different aspects of the UCS and thus produce different effects. According to this view, the CS is associated with cached outcome value in general PIT and can thus regulate responding for any reinforcer. In contrast, in specific PIT, the CS is associated with current sensory aspects of the UCS and thus regulates instrumental responding for this specific UCS only. Based on this conceptualization, Holland [26] suggested that specific but not general PIT can encode current reward value and by extension that specific but not general PIT should be sensitive to shifts in this value. This suggestion is important because it describes a potential mechanism by which current reward value can influence S-R behavior.

In an attempt to test this possibility, Holland [26] failed to demonstrate that specific PIT is influenced by UCS devaluation. One interpretation of this finding is that current value

information is linked to the sensory aspects of the UCS and these were indeed encoded by the CS, but that a demonstration of an effect of UCS devaluation on specific PIT requires certain conditions that were not met in these experiments. Such conditions are described in the following section.

2.2.2. Single versus multiple reinforcers—One situation in which S-R behaviors appear to fall under the influence of current reward value occurs when multiple stimuli associated with multiple reinforcers are presented in an operant chamber [27]. An example would be that of a single chamber that contains a chain and a lever the pulling or pressing of which delivers food or sucrose respectively. Table 2 lists a number of studies that reported sensitivity or insensitivity to reward value depending on the number of stimuli and reinforcers that were presented in the operant chamber. In the presence of multiple stimuli and reinforcers, animals remain sensitive to reward value even when subjected to training paradigms that produce S-R associations, such as overtraining and interval training schedules [18, 26].

Specific PIT may provide a mechanism by which S-R behaviors can occur in a reward sensitive manner. This suggestion is based on a number of factors that inform how PIT may be produced in a multiple reinforcer setting. 1) Jonkman et al. [28] demonstrated that the operant chamber in such instrumental studies can become a Pavlovian CS that exerts PIT effects on behavior by virtue of its association with the UCS reinforcers provided in the chamber. 2) Holland [26] showed that the presence of multiple UCS reinforcers produces specific PIT. 3) There are currently no theories that fully account for the maintenance of reward sensitivity of S-R behaviors in the multiple stimuli/reinforcer paradigm. It is conceivable that behavior in this paradigm reflects an interaction between S-R associations and specific PIT. The behaviors performed are trained and developed as S-R associations. As per the suggestion of Holland [26], the reward sensitivity evidenced in this paradigm may be the product of reward sensitive specific PIT.

One question that remains is why the effect of specific PIT on behavior in the multiple stimuli/reinforcer paradigm integrates current reward value [Table 2] when specific PIT in the study of Holland [26] did not. The answer may lie in the nature of the associations formed between the multiple UCS s and CS s used in these studies. Interestingly, Holland [26] arranged for the association of each of two different discriminative stimuli with a unique UCS while studies observing reward sensitivity did not. Rather in these latter studies, multiple UCS reinforcers were associated with one CS, the test chamber. Thus it may be that association between multiple UCS reinforcers and one CS is necessary for the PIT signal to encode current information value. Sensitivity to UCS reinforcer devaluation may reflect the limited capacity of one CS to process the sensory characteristics of multiple UCS s so that momentary changes in one UCS would be encoded by the CS. This limited capacity of a single CS to process multiple associations may render it more sensitive to changes in the current value of a UCS. With the individual CS-UCS associations used in the study of Holland [26], devaluation of the UCS may have diminished processing of the sensory characteristics of the UCS by the CS in favor of its general cached value characteristics and increased the prevalence of general PIT that is insensitive to changes in reward value. In the latter case, spared capacity, permitted by an individual CS-UCS association, would allow for the encoding of general motivational in addition to sensory characteristics of the UCS.

Based on this explanation, it would be predicted that the multiple stimuli/reinforcer paradigm will produce reward sensitive responding only when multiple UCS reinforcers are associated with the same CS, such as the operant chamber, but not when they are associated with separate CS s, such as unique discriminative stimuli. A similar failure to detect reward sensitivity, despite the presence of multiple reinforcers, was reported by Tricomi et al.[29].

As in the Holland study [26], different discriminative stimuli signaled specific UCS reinforcers and these remained insensitive to UCS devaluation.

In summary, we have described conditions in which S-R responding can be performed in a reward sensitive way. S-R trained operant behaviors in the multiple stimuli/reinforcer paradigm can be regulated by specific PIT and this is observed when a single CS is associated with multiple UCS s. An alternative explanation of the multiple stimuli/reinforcer effect has been suggested in which the presence of multiple response options inhibits S-R development and maintains reliance on R-O associations [18, 30]. In the case where two responses are emitted, for example, each response detracts from the degree that the other response is performed in a consistent and repetitive way thereby slowing the development of S-R associations. However, studies using discriminative stimuli have demonstrated reward insensitivity despite the potential for response interference [26, 29]. Hence, it appears that the number of responses alone that are available in the testing chamber are not a determining factor in the development of reward insensitivity.

The suggestion that specific PIT may integrate real-time reward information into S-R based behaviors allows for the modification of S-R behaviors in response to real-time shifts in reward value. According to this view, these S-R behaviors would remain distinct from R-O behaviors that are similarly regulated by current value. 1) At least two behaviors are necessary for the regulation of S-R behavior by shifts in current value. 2) These behaviors must be developed and expressed in the presence of the same Pavlovian CS; 3) If the relationship between the CS and the UCS reinforcers is extinguished, the operant responses will become reward insensitive.

2.3. Relevance for the treatment of addiction

We have described how both cached and current reward values can influence S-R responding. Thus, any characterizations of the effect of increased S-R responding on drug use must account for the role of reward value in mediating S-R behavior.

Regarding cached reward value, it will be important to determine what factors play into the caching process and to determine when in the drug use process caching occurs. Interventions and prevention programs could aim to manipulate the caching process so that lower reward values are cached into behavior, or to intervene at points that may prevent caching from occurring.

Regarding current value, it is clear that the degree to which a behavior is performed in a habitual outcome insensitive way is related to the surrounding stimuli present during drug pursuit. Despite the development of S-R behaviors, sensitivity to current shifts in outcome can still be observed in the presence of the appropriate surrounding stimuli. The application of disincentives for drug use by therapeutic and legal means may be more effective if contexts are exploited to promote the direct competition of these alternative reinforcers with drug use. Therefore, careful assessment of drug pursuit contexts could guide therapeutic intervention and inform ways to reintroduce sensitivity to reward value into habitual behavior. In terms of prevention, education about the types of contexts that promote outcome insensitive versus outcome sensitive responding could be used to reduce liability for the development of addictive behaviors.

3. Automaticity and S-R associations

A number of reports have extended the idea that S-R behaviors are reward insensitive to the field of cognition by suggesting that they are more likely to be performed automatically [7, 10, 16, 17]. From a cognitive perspective, automatic behavior is performed without the

requirement of attentional resources [31]. This in turn makes it more difficult to inhibit [32]. To our knowledge, there has been no systematic investigation of the possibility that S-R processes recruit less attentional resources than R-O processes (see [33] for a neurobiological perspective). Here we evaluate this possibility by reviewing the respective behavioral determinants of S-R associations and automaticity.

Tiffany [32] identified two primary requirements for the development of automaticity: overtraining and a consistent relationship between behavior and reward. In addition, Baker et al. [34] suggested that unconscious behavioral processes enter consciousness and demand attention when they are linked with strong emotion. For example, when denied cigarettes, a smoker's previously automatic smoking behaviors will revert to conscious processes that recruit increasing attention as craving develops [34]. While strong emotion may inhibit automaticity there have been no systematic investigations of the effect of emotionality on the performance of S-R versus R-O based behaviors.

The role of overtraining in S-R development remains unclear. Some studies have differentiated between S-R and R-O development based on training time alone [26] while others have demonstrated S-R behaviors in non-overtrained animals [12, 13]. Yet other studies have reported no difference between moderately and over trained animals [27, 28]. Thus, while overtraining may be one pathway to S-R development, it does not appear to be fundamentally necessary.

The second determinant of automaticity, maintenance of a consistent relationship between behavior and reward, is in direct opposition with the use of interval training schedules, a major determinant of S-R development. The development of S-R versus R-O based behaviors depends primarily on whether an interval or ratio training schedule is used [10, 15, 35]. By definition, an interval training schedule establishes an inconsistent relationship between behavior and reward.

Of the two determinants of automaticity, one can potentially contribute to and one is inconsistent with S-R development. This suggests that automaticity and S-R associations constitute separate phenomena. While the DLS has been implicated in both automaticity and S-R performance [29, 33] it remains unclear whether repeated stimulant use necessarily leads to automaticity.

4. Maintained interactions between R-O and S-R substrates in striatum

Thus far, we have described conditions that support the continued interaction between outcome value and S-R associations. These interactions are particularly informative for understanding how drug use can be regulated especially following prolonged periods of exposure that have been argued to sensitize the neural substrates of S-R processing [10, 11]. We suggest that despite the pharmacological enhancement of S-R substrates, outcome value continues to play an important though altered role in drug seeking and taking. In the following section, we review evidence implicating interactions between striatal sites argued to mediate outcome and S-R processing. This evidence clearly demonstrates continued interaction between habit and reward processing substrates in prolonged stimulant use.

The VS has been linked to processing of outcome value [21, 36] as well as the acquisition [37] and generation [38] of instrumental responses. As such, activation of this area generally indicates the processing of outcome value [16, 39]. By contrast, the DLS contributes to reward insensitive responding and habitual modes of behavior. Both of these areas have been implicated in drug seeking and drug taking [11, 40]. Under some conditions, they are activated concurrently while in others, they appear to operate independently.

Increases in dopaminergic activity in response to cocaine intake have been demonstrated simultaneously in both VS and DLS following prolonged cocaine use [2, 4] suggesting a possible interaction between processing in these two sites. Concurrent activation of VS and DLS during drug use may support the encoding of cached reward values into the stimulus component of S-R associations. Similarly, co-activation of these areas may lead to encoding of those CS-UCS associations that subsequently support the regulation of S-R responding by specific PIT. In fact, the development of S-R behavior is dependent on the serial connectivity of the VS and DLS [41]. In this latter report, the authors argue for parallel R-O and S-R learning processes, but the need of VS for S-R behavior clearly indicates a connection between the anatomical substrates for outcome value and S-R associations.

A progression from VS to DLS DA activity has been demonstrated in a number of studies over the course of development of well established responding to drug cues [3, 42]. DLS but not VSDA activity during the performance of well established drug seeking behaviors has been demonstrated [3, 7, 8, 42]. These findings have served as the basis for the suggestion that control over drug seeking behavior evolves from outcome related VS to habitual DLS processing [10]. It is important to note that DLS-only activity during drug seeking was demonstrated in these studies under conditions associated with drug unavailability. In the studies of Volkow et al. [7] and Wong et al. [8], dependent human drug users were presented with drug cues in the laboratory, a condition not associated with drug availability. Similarly, in the studies of Ito et al. [3, 42], DLS-only activity was demonstrated while rats responded under a second order schedule of reinforcement in the absence of drug. It is conceivable that absence of VS activity in these studies reflects suppression of value related processing by a context paired with drug unavailability. In contrast to these findings, Tricomi et al [29] found VS activation in response to food reward cues in humans over the course of habit development. In this study, food was given in the laboratory and as a result subjects were tested in a food paired context. Notably, VS responding to reward cues did not decrease over the course of habit development but remained constant. Thus, it is possible that findings showing diminished VS activity over the course of habit development may in fact reflect conditioned suppression of VS activity by a reward suppressing context [9]. Under these circumstances, DA transmission in VS during drug seeking may thus be reduced relative to that observed in DLS and thereby diminish the ability of outcome value to influence responding. Indeed, more severe interventions such as lesions of the VS or blockade of AMPA receptors in this site prevent established drug seeking altogether [41, 43]. These findings are consistent with the ability of UCS associated CSs to regulate S-R responding and may indicate the neurobiological bases of these effects. Similarly, it is conceivable that the VS is also required to mediate the effects of cached reward value in S-R responding by maintaining its representation in the DLS via VS to DLS projections [41]. The nature of the interactions between VS and DLS during prolonged drug taking need to be explored further.

DLS involvement clearly increases over the course of drug use [1–4]. Nonetheless, it is clear that the DLS continues to interact with the VS throughout this time [2, 41]. These interactions are consistent with our suggestions above [section 2] that outcome value continues to regulate S-R behavior. Thus, while repeated drug use leads to increased reliance on S-R associations, these associations still maintain a dynamic relationship with outcome considerations.

A number of additional studies also indicate that the characterization of the DLS as a value independent brain region is insufficient. For example, DLS lesions have been shown to diminish reinforced responding in R-O [44, 45] as well as S-R paradigms [46]. It remains unclear what the shared role of the DLS in both R-O and S-R behaviors might be. These studies controlled for motor retardation and spontaneous behavior. Thus, the DLS appears to encode some substrate of motivational processing underlying both S-R and R-O behaviors.

It is tempting to speculate that this shared motivational substrate involves cached value as there is representation of cached reward outcome that occurs during S-R behavior [23] and cached value contributes to the performance of R-O behaviors [27].

5. Conclusion: Interaction of S-R associations with outcome

We have outlined a number of different ways in which outcome may continue to regulate the manifestation of S-R associations. These are summarized in Figure 1. Specifically, the intensity of S-R based behaviors may reflect cached reward values encoded into the eliciting stimulus or into a CS generating a PIT signal via general PIT. Current reward values can influence S-R behaviors through the mechanism of specific PIT. From a neurobiological perspective, continued interaction between habit and outcome related brain regions can support their continued simultaneous contribution. Finally, S-R behaviors can also be influenced by the recruitment or absence of attentional processes that relate to the automaticity of responding.

These points guide interpretation of the development of increased DLS activity over the course of drug use. It has been suggested that this phenomenon underlies the expression of “incentive habits” to connote the evolution of habits from previously reward sensitive behaviors [10]. We suggest that the maintained interplay between VS and DLS as well as the continued role played by outcome value in S-R behaviors reflects more the development of **incentive regulated habits**, that is, the development of habitual behaviors that remain sensitive to outcome value.

In future studies, it will be important to characterize the determinants of caching: when in training caching occurs and what environmental factors influence the values of cached representations. It will also be important to further investigate the contribution of specific PIT to the generation of reward sensitive S-R behaviors. As our understanding of these phenomena increases, so too will our ability to exploit them in the treatment of addictions.

Acknowledgments

The preparation of this manuscript was supported by National Institute on Drug Abuse grant DA09397 to PV. The authors wish to thank Dr. Robert McGrath for help initiating this project.

Abbreviations

CS	conditioned stimulus
DA	dopamine
DLS	dorsolateral striatum
PIT	Pavlovian to instrumental transfer
R-O	response outcome
S-R	stimulus response
UCS	unconditioned stimulus
VS	ventral striatum

References

1. Porrino LJ, Lyons D, Smith HR, Daunais JB, Nader MA. Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J Neurosci*. 2004; 24:3554–62. [PubMed: 15071103]

2. Porrino LJ, Smith HR, Nader MA, Beveridge TJ. The effects of cocaine: a shifting target over the course of addiction. *Prog Neuropsychopharmacol Biol Psychiatry*. 2007; 31:1593–600. [PubMed: 17900777]
3. Ito R, Dalley JW, Robbins TW, Everitt BJ. Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. *J Neurosci*. 2002; 22:6247–53. [PubMed: 12122083]
4. Nader MA, Daunais JB, Moore T, Nader SH, Moore RJ, Smith HR, et al. Effects of cocaine self-administration on striatal dopamine systems in rhesus monkeys: initial and chronic exposure. *Neuropsychopharmacology*. 2002; 27:35–46. [PubMed: 12062905]
5. Boileau I, Dagher A, Leyton M, Gunn RN, Baker GB, Diksic M, et al. Modeling sensitization to stimulants in humans: an [¹¹C]raclopride/positron emission tomography study in healthy men. *Arch Gen Psychiatry*. 2006; 63:1386–95. [PubMed: 17146013]
6. Evans AH, Pavese N, Lawrence AD, Tai YF, Appel S, Doder M, et al. Compulsive drug use linked to sensitized ventral striatal dopamine transmission. *Ann Neurol*. 2006; 59:852–8. [PubMed: 16557571]
7. Volkow ND, Wang GJ, Telang F, Fowler JS, Logan J, Childress AR, et al. Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. *J Neurosci*. 2006; 26:6583–8. [PubMed: 16775146]
8. Wong DF, Kuwabara H, Schretlen DJ, Bonson KR, Zhou Y, Nandi A, et al. Increased occupancy of dopamine receptors in human striatum during cue-elicited cocaine craving. *Neuropsychopharmacology*. 2006; 31:2716–27. [PubMed: 16971900]
9. Vezina P, Leyton M. Conditioned cues and the expression of stimulant sensitization in animals and humans. *Neuropharmacology*. 2009; 56 (Suppl 1):160–8. [PubMed: 18657553]
10. Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behav Brain Res*. 2009; 199:89–102. [PubMed: 18950658]
11. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005; 8:1481–9. [PubMed: 16251991]
12. Faure A, Haberland U, Conde F, El Massioui N. Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J Neurosci*. 2005; 25:2771–80. [PubMed: 15772337]
13. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*. 2004; 19:181–9. [PubMed: 14750976]
14. Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998; 37:407–19. [PubMed: 9704982]
15. Dickinson A, Nicholas Dj, Adams Cd. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Quarterly Journal of Experimental Psychology [B]*. 1983; 35:35–51.
16. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nat Rev Neurosci*. 2006; 7:464–76. [PubMed: 16715055]
17. Pierce RC, Vanderschuren LJ. Kicking the habit: the neural basis of ingrained behaviors in cocaine addiction. *Neurosci Biobehav Rev*. 2010; 35:212–9. [PubMed: 20097224]
18. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005; 8:1704–11. [PubMed: 16286932]
19. Wood W, Neal DT. A new look at habits and the habit-goal interface. *Psychol Rev*. 2007; 114:843–63. [PubMed: 17907866]
20. Redish AD. Implications of the multiple-vulnerabilities theory of addiction for craving and relapse. *Addiction*. 2009; 104:1940–1. [PubMed: 19832790]
21. Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)*. 2007; 191:391–431. [PubMed: 17072591]
22. Balleine B. Instrumental performance following a shift in primary motivation depends on incentive learning. *J Exp Psychol Anim Behav Process*. 1992; 18:236–50. [PubMed: 1619392]
23. Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes R. A Motivational control after extended instrumental training. *Animal Learning & Behavior*. 1995; 23:197–206.

24. Holmes NM, Marchand AR, Coutureau E. Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neurosci Biobehav Rev.* 2010; 34:1277–95. [PubMed: 20385164]
25. Rescorla RA, Solomon RL. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychol Rev.* 1967; 74:151–82. [PubMed: 5342881]
26. Holland PC. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process.* 2004; 30:104–17. [PubMed: 15078120]
27. Colwill R, Rescorla RA. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *J Exp Psychol [Anim Behav].* 1985; 11:520–536.
28. Jonkman S, Kosaki Y, Everitt BJ, Dickinson A. The role of contextual conditioning in the effect of reinforcer devaluation on instrumental performance by rats. *Behav Processes.* 2010; 83:276–81. [PubMed: 20060882]
29. Tricomi E, Balleine BW, O'doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci.* 2009; 29:2225–32. [PubMed: 19490086]
30. Kosaki Y, Dickinson A. Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *J Exp Psychol Anim Behav Process.* 2010; 36:334–42. [PubMed: 20658864]
31. Poldrack RA, Sabb FW, Foerdes K, Tom SM, Asarnow RF, Bookheimer SY, et al. The neural correlates of motor skill automaticity. *J Neurosci.* 2005; 25:5356–64. [PubMed: 15930384]
32. Tiffany ST, Carter BL. Is craving the source of compulsive drug use? *J Psychopharmacol.* 1998; 12:23–30. [PubMed: 9584965]
33. Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn Sci.* 2010; 14:208–15. [PubMed: 20207189]
34. Baker TB, Piper ME, McCarthy DE, Majeskie MR, Fiore MC. Addiction motivation reformulated: an affective processing model of negative reinforcement. *Psychol Rev.* 2004; 111:33–51. [PubMed: 14756584]
35. Dickinson A, Campos J, Varga ZI, Balleine B. Bidirectional instrumental conditioning. *Q J Exp Psychol B.* 1996; 49:289–306. [PubMed: 8962537]
36. O'doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science.* 2004; 304:452–4. [PubMed: 15087550]
37. Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, De Lecea L, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science.* 2009; 324:1080–4. [PubMed: 19389999]
38. Ito R, Robbins TW, Everitt BJ. Differential control over cocaine-seeking behavior by nucleus accumbens core and shell. *Nat Neurosci.* 2004; 7:389–97. [PubMed: 15034590]
39. Haber SN. The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat.* 2003; 26:317–30. [PubMed: 14729134]
40. Robinson TE, Berridge KC. Review. The incentive sensitization theory of addiction: some current issues. *Philos Trans R Soc Lond B Biol Sci.* 2008; 363:3137–46. [PubMed: 18640920]
41. Belin D, Everitt BJ. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron.* 2008; 57:432–41. [PubMed: 18255035]
42. Ito R, Dalley JW, Howes SR, Robbins TW, Everitt BJ. Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. *J Neurosci.* 2000; 20:7489–95. [PubMed: 11007908]
43. Di Ciano P, Everitt BJ. Dissociable effects of antagonism of NMDA and AMPA/KA receptors in the nucleus accumbens core and shell on cocaine-seeking behavior. *Neuropsychopharmacology.* 2001; 25:341–60. [PubMed: 11522463]
44. See RE, Elliott JC, Feltenstein MW. The role of dorsal vs ventral striatal pathways in cocaine-seeking behavior after prolonged abstinence in rats. *Psychopharmacology (Berl).* 2007; 194:321–31. [PubMed: 17589830]
45. Suto N, Wise RA, Vezina P. Dorsal as well as ventral striatal lesions affect levels of intravenous cocaine and morphine self-administration in rats. *Neurosci Lett.* 2011; 493:29–32. [PubMed: 21315801]

46. Vanderschuren LJ, Di Ciano P, Everitt BJ. Involvement of the dorsal striatum in cue-controlled cocaine seeking. *J Neurosci.* 2005; 25:8665–70. [PubMed: 16177034]
47. Coutureau E, Killcross S. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav Brain Res.* 2003; 146:167–74. [PubMed: 14643469]
48. Miles FJ, Everitt BJ, Dickinson A. Oral cocaine seeking by rats: action or habit? *Behav Neurosci.* 2003; 117:927–38. [PubMed: 14570543]
49. Dickinson A, Wood N, Smith JW. Alcohol seeking by rats: action or habit? *Q J Exp Psychol B.* 2002; 55:331–48. [PubMed: 12350285]

HIGHLIGHTS

- Dorsolateral striatal activity has been proposed to underlie habitual drug use
- This region supports stimulus-response (S-R) based behaviors
- Behavioral evidence is described showing that reward value can regulate S-R behaviors
- This is consistent with a maintained interaction between ventral and dorsal striatum

Figure 1.

Multiple factors can influence the experience and intensity of S-R behaviors. a) The intensity of an S-R behavior may be related to the value of the initial reinforcer whose memory has been cached into the eliciting stimulus. b) If an S-R behavior is performed in a context paired with multiple reinforcers, it can be modulated by shifts in the current value of any one reinforcer via specific PIT. If it is performed in a context paired with only a single reinforcer, it can be influenced by the cached value of that reinforcer via general PIT. c) If an S-R behavior is overtrained and performed immediately after exposure to the eliciting stimulus, it is likely to occur automatically and not to recruit attentional resources. Alternatively, if it is not overtrained and occurs after a delay following the eliciting cue, it is likely to recruit attention processes. See text for supporting references.

Table 1**Key Terms and Definitions**

<p>R-O. An operant association in which the probability of a response is linked to a rewarding outcome. These associations can be identified by varying the value of the outcome and measuring variations in response amplitude [14].</p> <p>S-R. An operant association in which a stimulus is linked to a response independent of outcome. These associations can be identified by varying the value of the outcome and demonstrating a lack of variation in response amplitude [14].</p> <p>Reward Value Sensitivity. The degree to which operant responses are influenced by changes in the value of reinforcers associated with operant stimuli. Reward-sensitive behaviors are considered reflective of R-O associations while reward-insensitive behaviors are considered reflective of S-R associations [14].</p> <p>DLS/S-R development. The suggestion that over the course of drug use neural activation radiates from ventral to dorsal striatal regions. The latter regions are associated with the performance of habitual S-R behaviors. Hence, DLS activation is suggested as a mechanism by which drug seeking and use become increasingly under the control of S-R modes [11].</p> <p>Interval Training. A reinforcement schedule that rewards operant behavior in timed intervals. This type of schedule distances behavior from reward as behavior is not always involved in making the reward available. This type of schedule has been associated with the development of S-R associations [14].</p> <p>Overtraining. Refers to the extent that a given training paradigm is repeated. The role of overtraining in S-R development is unclear. It appears to be sufficient [26] but not necessary for S-R development [12, 13]. Other studies, however, have found that overtraining produces no effect on the development of S-R associations [27, 28].</p> <p>Ratio Training. Reinforcement schedule that rewards operant behavior according to the amount of behavior produced. This type of schedule establishes a strong relationship between reward and behavior as all responses are related to reward provision. This schedule is more likely to produce R-O associations [14, 15].</p>
--

Table 2

Comparison of Reward Devaluation studies

Number of operant stimuli in performance context	Effect of devaluation	Study [*]
One lever	No effect	Coutureau and Killcross, [47] [†]
One lever	No effect	Yin et al. [13] [†]
Lever and ceiling pole	Lowered responding on devalued stimulus	Miles et al. [48]
Two levers	Lowered responding on devalued stimulus	Dickinson et al. [49]
Bi-directional pole	Lowered responding on devalued stimulus	Dickinson et al. [35]
Lever Press and chain pull	Lowered responding on devalued stimulus	Colwill and Rescorla [27]

^{*} All studies used interval training schedules and reported equivalent numbers of training days and sessions. Number of sessions did not predict devaluation effects.

[†] These studies assessed the effect of lesions on reward sensitivity. The highlighted results are from comparisons between subgroups subjected to devaluation or no devaluation within the non-lesioned control groups.