Model-Free Temporal-Difference Learning and Dopamine in Alcohol Dependence: Examining Concepts From Theory and Animals in Human Imaging

Quentin J.M. Huys, Lorenz Deserno, Klaus Obermayer, Florian Schlagenhauf, and Andreas Heinz

ABSTRACT
Dopamine potentially unites two important roles: one in addiction, being involved in most substances of abuse including alcohol, and a second one in a specific type of learning, namely model-free temporal-difference reinforcement learning. Theories of addiction have long suggested that drugs of abuse may usurp dopamine’s role in learning. Here, we briefly review the preclinical literature to motivate specific hypotheses about model-free temporal-difference learning and then review the imaging evidence in the drug of abuse with the most substantial societal consequences: alcohol. Despite the breadth of the literature, only a few studies have examined the predictions directly, and these provide at best inconclusive evidence for the involvement of temporal-difference learning alterations in alcohol dependence. We discuss the difficulties of testing the theory in humans, make specific suggestions, and close with a focus on the interaction with other learning mechanisms.

Keywords: Alcohol, Computational psychiatry, Cue reactivity, Dopamine, Habits, Model-free, Reinforcement learning, Ventral striatum

There is substantial evidence pointing to a role for dopamine in both addiction and learning, which naturally raises the question of whether dopamine’s role in addiction is mediated via its role in learning. Most addictive substances result in dopaminergic release in the ventral striatum (VS) (1–5), where dopamine signals are at the center of biologically increasingly detailed (6) model-free temporal difference (MFTD) accounts of how the brain instantiates iterative learning from reinforcements (7). Given evidence that phasic dopamine signals have a causal role in learning (8), it is reasonable to expect that addictive substances might exert their nefarious effect in part by subverting dopamine’s role in MFTD learning (9,10).

It is unclear to what extent such a theory is supported by existing evidence. Here, we therefore examine findings in one of the societally most important drugs of abuse (11,12) with a huge treatment gap (13): alcohol. We start with a theoretical overview, mapping the valuation of stimuli onto incentive salience and sign-tracking theories, and the valuation of actions onto habituation theories. We then review the relevant imaging literature in humans and close with a discussion of outstanding issues and the limitations of existing tests of MFTD theories in humans.

MODEL-FREE TEMPORAL DIFFERENCE LEARNING

Stimulus Values and Incentive Salience
One influential account of addiction builds on the finding that stimuli paired with dopamine release or stimulation acquire incentive salience (14,15), becoming 1) desirable, 2) reinforcing in their own right, and 3) motivating. These are typical of drug-associated stimuli and might thus contribute to both development and maintenance of addictive states. Anecdotally, patients often report relapsing after encountering alcohol-related stimuli.

Model-free prediction-error learning (16) iteratively updates reward expectation values \( V \) with a prediction error that measures the discrepancy with the actually obtained reward \( r \):

\[
V_{t+1} = V_t + \alpha (r - V_t)
\]

The value \( V \) is a running average of experienced reinforcements that summarizes past reinforcement experience. In MFTD learning, the prediction error incorporates changes in expected rewards induced by reward-predicting stimuli (16). When cues predictive of reward occur unexpectedly, a prediction error proportional to this expectation is elicited.

MFTD valuation of stimuli results in Pavlovian stimulus values \( V(s) \) that capture the three core aspects of incentive salience (17,18). They become desirable in that approaching stimuli with positive value \( V(s) \) is formally optimal (16). Because the specifics of how and when reinforcement happened are discarded, the desirability becomes separated from the details of past experience. Second, temporal difference values capture secondary reinforcement because a positive change in reward expectation can formally substitute for actual rewards.
Third, the motivating aspects (19) are captured by the fact that expectations of reward determine the optimal rate of action (20), although notably this is not specific to MFTD values. Finally, the delay in adapting values to reflect the current rewards provides one account for why wanting the drug (captured by V) is distinct from liking it (the immediate reward r experienced on consumption) and suggests one reason why wanting may persist beyond liking (21).

Stimulus Values and Sign-Tracking
There is substantial individual variation in Pavlovian conditioning. When a light in one corner of a box predicts food delivery in the other corner, sign-tracking rats will come to approach the light conditioned stimulus (CS) and wait there until delivery of the unconditioned stimulus (US), e.g., food. Goal-trackers, in contrast, move to the food delivery site immediately. Phasic dopamine levels in the VS behave like MFTD prediction errors in sign-trackers only, and only in them can learning the CS-US relation be blocked by dopaminergic antagonists (22,23). Hence, sign-trackers rely on dopaminergically mediated MFTD learning, whereas goal-trackers do not. Only in sign-trackers does the CS acquire incentive salience.

The link to addiction comes through animals selectively bred to show high or low responsiveness to novelty (24). The animals selectively bred to show high responsiveness to novelty are preferentially sign-trackers for natural rewards and show a broad range of addiction-like features (25). They respond more to cocaine acutely and show more locomotor sensitization effects (26), show stronger drug-taking acquisition (27), work harder for cocaine (28), seek cocaine when it is no longer available (29), are more impulsive on a range of measures (29), and have reduced dopamine D2 receptor (D2R) availability, also implicated in human addiction (30–32). Cocaine cues lead to escalation and reinstatement after extinction in sign-trackers but not in goal-trackers (33). Alcohol releases dopamine (34), and exposure to sign-tracking paradigms in adolescence increases sign-tracking and ethanol intake in adulthood (35). In humans, effects of Pavlovian stimuli on instrumental behavior have been confirmed, and in one study general Pavlovian-to-instrumental transfer effects predicted to relapse risk in alcohol use disorder (AUD) patients (38), although differences between goal- and sign-trackers have not yet been explored with respect to addiction.

Values, Habits, and Devaluation
Alcohol intake is suggested to have habitual components (37–39) because substance use persists despite obvious harmful consequences. Habits are defined through devaluation insensitivity (40), whereby behavior will continue even though the outcome of the action is no longer consumed if available freely. Habits contrast with goal-directed behavior, where the action will only be performed if the action’s goal is desirable. MFTD values, be they about states or behaviors, capture devaluation insensitivity because they rest entirely on past experiences about how actions lead to outcomes, and they are not updated by information purely about the outcome itself (41) until the association between the state or action and the revalued outcome has been experienced.

Phasic dopaminergic signals are present during instrumental learning (42), and dopamine (43), the dorsolateral striatum, and the infralimbic cortex are required for habit formation both for natural rewards (44,45) and for drugs such as alcohol (46). Similar to habits, sign-tracking itself is resistant to devaluation of the outcome, whereas goal-tracking is not (47), and over extended training goal-tracking for alcohol gives way to sign-tracking (48), suggesting a similarity to a MFTD valuation process. Exposure to stimulants or alcohol speeds up habit formation for drug or natural rewards (46,49). Furthermore, D2R antagonism, putatively modeling the reduction in D2R availability (30–32), further promotes this (50), and dopaminergic signals shift from ventral to dorsal striatum with progression of the habituation (51).

IMAGING MFTD PROCESSES IN ALCOHOLISM
Several features must be satisfied to establish the presence of MFTD learning signals (7,52). Unpredicted rewards and unpredictable changes in reward expectation should result in a positive signal proportional to the difference between reward and expectation or expectation change. Responses to predictable rewards should decrease over repetitions, whereas responses to neutral stimuli predicting rewards reliably should increase over the course of learning. Unexpected omission of an expected reward should result in a negative signal. MFTD signals should not be sensitive to devaluation.

The responses should be visible in dopaminergic target region blood oxygen level–dependent (BOLD) measurements (53,54). For Pavlovian (cue) processes, this should be in the VS (22,55) that receives a strong dopaminergic projection (56). For habitual (action) processes, the signals may arise in the dorsal striatum (51,55,57).

There are two different ways in which MFTD processes might contribute to the development of dependence. Alcohol might specifically affect MFTD learning for stimuli or behaviors associated with alcohol (described in MFTD Processes and Alcohol Cues). Alternatively, a predisposition toward MFTD learning observable also in nondrug scenarios may predispose toward alcohol addiction (described in MFTD Processes and Nonalcoholic Rewards).

MFTD Processes and Alcohol Cues
Alcohol may specifically usurp MFTD processes to engender particularly powerful learning in situations associated with it. Definite evidence for this would require the learning process to be observed longitudinally over the course of the development of addiction. Cross-sectional examination of the end-result of learning, that is, responses to putative CSs, is weaker. Nevertheless, on the basis of the features of MFTD learning noted above, the following criteria should be met to support the involvement of MFTD processes:

- 1a: Responses to drug CSs should be more pronounced among individuals who have developed an addiction than among those who have not.
- 1b: Unexpected presentations of drug CSs should be accompanied by phasic dopaminergic release in the VS for Pavlovian settings and either in the ventral or dorsal striatum for instrumental settings (55).
Learning and Dopamine in Alcohol Dependence

- 1c: Drug CSs should be able to replace rewards and drive further learning in second-order conditioning experiments.
- 1d: Drug CSs should show resistance to experienced or instructed devaluation akin to MFTD values.

Functional Magnetic Resonance Imaging Cue Reactivity Studies. Cue-reactivity functional magnetic resonance (fMRI) paradigms compare the neuroimaging responses evoked by drug-associated stimuli and stimuli not associated with drugs (Table 1). Although indirect, BOLD signals in target areas of dopaminergic projections may still capture dopaminergic transients (63,54,73). In terms of the first criterion, 1a, drug-associated stimuli lead to higher autonomic activation than stimuli not associated with drugs (74), and this is accompanied by recruitment of the VS, with higher BOLD signals in the VS for alcohol-related stimuli compared with control stimuli (75,76). However, a meta-analysis showed that this does not differentiate patients from healthy control subjects (Table 1) (76).

One cause for this lack of group difference could be the inclusion of treatment-seeking patients. Treatment seeking is more likely once alcohol consumption becomes aversive (77). Because different subcomponents of the VS subserve not only MF but also goal-directed decision making (78,79), the goal to reduce drinking may affect ventral striatal responses to drink stimuli by mechanisms other than MFTD mechanisms. Drug cues also evoke counterregulatory processes (80), and both treatment (76) and abstinence (65,81) reduce cue reactivity in the VS.

The studies examining actively drinking, not treatment-seeking, patients, however, paint a similar picture. Among the studies examined by Schacht et al. (76), three reported enhanced ventral striatal cue response in patients versus control groups (58,59,61), and two identified positive correlations between striatal and ventral tegmental area activations with measures of drinking severity (60,62). Seven studies reported no group differences or differences in the opposite direction (66–70,72,82), and the remaining did not examine group comparisons or correlations with severity. Newer studies, although not yet subjected to a formal meta-analysis, paint a similar picture. One study reported a ventral striatal group difference but a dorsal striatal correlation with drinking severity (63), one reported a dorsal striatal group difference (65), and two further studies failed to find ventral striatal effects (64,71). Interestingly, dorsal striatal cue responses are more prominent in subjects who go on to show escalating drinking (64), but earlier relapse after detoxification is accompanied by a lower VS cue reactivity response (83), both without group differences.

Further experimental aspects may have contributed to the null findings. Presentation of visual cues in an unusual, novel environment may only partially tap into the relevant stimulus dimensions (84). Accelerated learning can be prompted by exposure in novel circumstances, and this may have led to response extinction in patients (85,86). Finally, block designs have low power to observe learning effects. Because only the first few stimulus presentations are unexpected and would have resulted in a phasic temporal prediction error, collapsing over the entire block would reduce effect sizes. The current evidence is therefore weak. It suggests that alcohol is apportioned a higher MF value than neutral cues, but also

<table>
<thead>
<tr>
<th>Table 1. fMRI Reactivity to Alcohol Cues in NTS Alcoholics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study, Year</strong></td>
</tr>
<tr>
<td><strong>Overall Positive Results</strong></td>
</tr>
<tr>
<td>Kareken et al. (59), 2004</td>
</tr>
<tr>
<td>Myrick et al. (60), 2008</td>
</tr>
<tr>
<td>Filley et al. (61), 2008</td>
</tr>
<tr>
<td>Ihssen et al. (61), 2011</td>
</tr>
<tr>
<td>Claus et al. (62), 2011</td>
</tr>
<tr>
<td>Sjoerds et al. (63), 2014</td>
</tr>
<tr>
<td>Dager et al. (64), 2014</td>
</tr>
<tr>
<td>Brumback et al. (65), 2015</td>
</tr>
<tr>
<td><strong>Overall Negative Results</strong></td>
</tr>
<tr>
<td>George et al. (66), 2001</td>
</tr>
<tr>
<td>Tapert et al. (67), 2003</td>
</tr>
<tr>
<td>Myrick et al. (68), 2004</td>
</tr>
<tr>
<td>Tapert et al. (69), 2004</td>
</tr>
<tr>
<td>Park et al. (70), 2007</td>
</tr>
<tr>
<td>Fryer et al. (71), 2013</td>
</tr>
<tr>
<td>Vollstädt-Klein et al. (72), 2010</td>
</tr>
</tbody>
</table>

AD, alcohol dependence; ADS, Alcohol Dependence Scale; AUD, alcohol use disorder; AUDIT, Alcohol Use Disorder Identification Test; CD, current drinkers; DS, dorsal striatum; DT, transition from moderate to heavy drinking; fMRI, functional magnetic resonance imaging; HC, healthy control; HD, heavy drinker; HDS, heavy drinker, stable; ICS-FS, failed control subscale of Impaired Control Scale; MDS, moderate drinker, stable; mPFC, medial prefrontal cortex; NTS, non-treatment seeking; OFC, orbitofrontal cortex; SD, social drinker; TS, treatment seeking; VS, ventral striatum; VTA, ventral tegmental area.
that the value does not differ between patients and control subjects. Hence, fMRI cue reactivity studies at present do not support criterion 1a.

**Positron Emission Tomography Studies.** Positron emission tomography (PET) studies may be helpful to examine whether alcohol cues release phasic dopamine (criterion 1b). Displacement studies measure dopamine release by comparing the binding potential of radiolabeled competitive ligands in a session in which neutral cues are presented and in another session with drug-related cues. The displacement of the tracer is a measure of the amount of dopamine released, and the difference between sessions is a measure of cue-evoked dopamine release. Alcohol itself displaces ligands from D2Rs (Table 2) and hence likely releases dopamine [Boileau et al. (87), Yoder et al. (96), Urban et al. (90), but see Yoder et al. (88,89)]. It also increases BOLD signals in the VS (39).

The impact of alcohol cues on dopamine release has only been examined in nondependent subjects (Table 2). Presentation of alcohol cues followed by (the predicted) alcohol ingestion results in ventral striatal dopamine release (96). However, presentation of the cues without the following outcome is somewhat complex. The cues themselves should elicit dopamine transients, but the absence of the predicted outcome should also elicit dopamine dips. The long timescale of PET would average over the two effects. Their relative preponderance depends on the baseline dopamine level as this will determine the asymmetry between phasic prediction error increases and dips in dopamine and on the details of the probe. Yoder et al. (96) found a reduction in dopamine release, whereas Oberlin et al. (87), using beer taste as a cue, found an increase in dopamine release. Because this cue is a proximal and reliable predictor of alcohol, it may have resulted in both larger positive transients and larger negative prediction errors when alcohol was not received immediately after the cue. Interestingly, the dopamine release signal in the study by Oberlin et al. (87) was driven entirely by those subjects with a positive family history for AUD, raising the possibility of a genetic contribution akin to how sign-tracking can be selectively bred (25). D2Rs are reduced in early detoxification (32,93), as is dopamine release capacity measured via amphetamine-induced raclopride displacement, specifically so in the VS (81), whereas dorsal synthesis capacity is increased (92). Overall, alcohol cues appear to modulate dopaminergic release. It is unclear, however, whether this

**Table 2. PET/SPECT Studies**

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Subjects</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RAC Displacement by Alcohol</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boileau et al. (87), 2003</td>
<td>7 HC</td>
<td>RAC in VS is displaced by oral alcohol ingestion</td>
</tr>
<tr>
<td>Yoder et al. (88), 2005</td>
<td>9 HC</td>
<td>RAC in striatum is not displaced by IV alcohol administration</td>
</tr>
<tr>
<td>Yoder et al. (89), 2007</td>
<td>13 HC</td>
<td>RAC in striatum is not displaced by IV alcohol administration, but correlates with subjective effects</td>
</tr>
<tr>
<td>Urban et al. (90), 2010</td>
<td>29 HC</td>
<td>RAC is displaced by oral alcohol ingestion; displacement in all striatal subregions but most prominent in the VS; displacement correlated positively with acute subjective effects and negatively with drinking history in men</td>
</tr>
<tr>
<td><strong>D2R Availability and DA Synthesis Capacity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tiihonen et al. (82), 1998</td>
<td>10 DeA, 8 HC</td>
<td>DA synthesis availability* is increased in DeA</td>
</tr>
<tr>
<td>Heinz et al. (32), 2004</td>
<td>11 DeA, 13 HC</td>
<td>VS D2R availability* is reduced in DeA</td>
</tr>
<tr>
<td>Heinz et al. (93), 2005</td>
<td>12 DeA, 13 HC</td>
<td>VS DA synthesis capacity* is not reduced in DeA but correlates negatively with craving</td>
</tr>
<tr>
<td>Deserno et al. (84), 2015</td>
<td>13 DeA, 14 HC</td>
<td>VS DA synthesis capacity* is not reduced in DeA and does not correlate with craving; alcohol intake abolishes correlation between DA synthesis capacity and MFTD PE</td>
</tr>
<tr>
<td><strong>PET/SPECT Reactivity to Alcohol Cues</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modell and Mountz (85), 1995</td>
<td>9 AD</td>
<td>Blood flow in caudate increased with alcohol + alcohol imagination vs. appetitive imagination alone</td>
</tr>
<tr>
<td>Lingford-Hughes et al. (63), 2006</td>
<td>6 AD, 6 SD</td>
<td>No whole-brain group differences in blood flow* during alcohol vs. neutral cues</td>
</tr>
<tr>
<td>Yoder et al. (96), 2009</td>
<td>18 HC</td>
<td>Striatal RAC is displaced by alcohol cues + IV alcohol; alcohol cues without alcohol lead to increased RAC binding</td>
</tr>
<tr>
<td>Oberlin et al. (97), 2013</td>
<td>4 AD, 45 SD</td>
<td>Proximal alcohol cues displace RAC in the VS; this is not proportional to desire to drink or measures of drinking extent but is more pronounced in those with a family history of alcoholism</td>
</tr>
<tr>
<td><strong>RAC Displacement by Nonalcohol Reward</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiland et al. (98), 2016</td>
<td>33 FH+, 11 FH−</td>
<td>RAC is displaced by reward task in the FH+ high-risk group more than in other groups</td>
</tr>
</tbody>
</table>

AD, alcohol dependence; BP, binding potential; DA, dopamine; D2R, dopamine 2 receptor; DeA, detoxified alcoholic; FH+, family history for alcohol use disorder positive; FH−, no family history for alcohol use disorder; HC, healthy control; IV, intravenous; MFTD, model-free temporal difference; PE, prediction error; PET/SPECT, positron emission tomography/single photon emission computed tomography; RAC, [11C]-raclopride; SD, social drinker; VS, ventral striatum.

*Measured using 6-[18F]-fluoro-L-dopa.
*Measured using [18F]-desmethoxyfallypride.
*Measured using [99mTc] SPECT.
*Measured using H2[15O].
follows the quantitative predictions of MFTD theory, and whether it is related to addiction.

MFTD PROCESSES AND NONALCOHOLIC REWARDS

The results on sign-tracking suggest a more general tendency toward MFTD learning that may be visible beyond the specific substance an addiction is being developed to. An increased tendency to rely on MFTD learning in situations not related to alcohol and on a timescale shorter than that of the development of the addiction may be a risk factor for the development of an addiction. MFTD learning should then explain a higher proportion of behavior (criterion 2a). In terms of imaging, the measured neural learning rate, that is, the speed at which the neurally measured PE signals adapt, should be more tightly linked to behavioral adaptation in those at risk (criterion 2b). Of note, a faster learning rate could facilitate a more rapid establishment of MFTD behaviors (criterion 2c), but faster learning could also arise through other, for example, model-based, mechanisms. When comparing or assessing learning rates, it is therefore critical either to exclude learning through non-MFTD systems or to measure the MFTD neural learning rate directly.

Because most of the evidence in this domain (Table 3) pertains to instrumental learning, the group differences could emerge in both VS and dorsal striatum, particularly in the dorsolateral striatum or putamen (44,55,72,100).

**Behavioral Data**

At a purely behavioral level (criterion 2a), there have been three studies broadly arguing for a behavioral shift toward MFTD learning in alcoholism (101–103). However, there have also been two negative results in patients (63,104), and young at-risk social drinkers do not show a bias toward MFTD learning (Stephan Nebe, Dipl. psych., et al., unpublished data, 2016). The positive results have shown impairments of model-based cognition, rather than a strengthening of MFTD processes. Four of these studies use the two-step task (78), where most studies have shown impairments in goal-directed control rather than specifically increments in MFTD components (101,103–107). Furthermore, tonic dopamine appears to promote, rather than inhibit, goal-directed components in humans (79,108).

The anticipation of alcohol and acute alcohol intoxication can shift decision making from goal-directed to habitual patterns acutely in healthy control subjects and animals (109–111). However, these effects appear to be due to impairments in model-based components. The initial acquisition and the post-devaluation reacquisition are not affected by alcohol intoxication but should be if alcohol intoxication indeed affected MFTD processes. There is therefore currently little behavioral evidence for a dopaminergically mediated bias toward MFTD learning in alcohol dependence.

**Imaging Data**

**Reversal Learning.** Two studies have used reversal learning in AUD patients (94,112), and a third study used a related task (Andrea Reiter, Ph.D., et al., unpublished data, 2016). These tasks involve a frequent or even continuous change in outcomes that is intended to encourage non-MFTD learning strategies (119,120). In terms of criterion 2a, MFTD learning models fitted to the behavior captured patients’ and control subjects’ behavior in two studies (94,112), whereas control subjects in a third study were characterized by an additional counterfactual process with no difference in the MFTD

**Slips of Action**

Table 3. Studies With Nonalcoholic Rewards

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Subjects</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reversal Learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al. (113), 2010</td>
<td>20 DeA, 16 HC</td>
<td>No group difference in VS PE correlates</td>
</tr>
<tr>
<td>Deserno et al. (79), 2015</td>
<td>13 DeA, 14 HC</td>
<td>No group difference in VS PE correlates</td>
</tr>
<tr>
<td>Reiter et al. (unpublished data), 2016</td>
<td>43 DeA, 35 HC</td>
<td>No group difference in VS PE correlates</td>
</tr>
<tr>
<td>Two-Step Task</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebe et al. (unpublished data), 2016</td>
<td>188 SD</td>
<td>No correlation between VS PE and drinking measures; increased DS PE correlates with earlier drinking onset</td>
</tr>
<tr>
<td>Slips of Action</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snoerds et al. (113), 2013</td>
<td>40 AD, 19 HC</td>
<td>Increased posterior putamen activity during incongruent trials</td>
</tr>
<tr>
<td>Monetary Incentive Delay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrase et al. (114), 2007</td>
<td>16 DeA, 16 HC</td>
<td>VS response to reward anticipation reduced in DeA; responses to actual rewards not reported</td>
</tr>
<tr>
<td>Bjork et al. (110), 2008</td>
<td>23 SDP, 23 HC</td>
<td>No group difference in VS response to reward anticipation; VS response to actual reward increased in DeA</td>
</tr>
<tr>
<td>Beck et al. (116), 2009</td>
<td>19 DeA*, 19 HC</td>
<td>VS response to reward anticipation reduced in DeA; no group differences in response to actual rewards</td>
</tr>
<tr>
<td>Bjork et al. (117), 2012</td>
<td>23 AD, 23 HC</td>
<td>No group difference in VS response to reward anticipation or actual rewards</td>
</tr>
<tr>
<td>Weiland et al. (98), 2016</td>
<td>33 FH−, 11 FH+</td>
<td>No group difference in VS response to reward anticipation or actual rewards</td>
</tr>
<tr>
<td>Becker et al. (118), 2016</td>
<td>32 DeA, 35 HC</td>
<td>VS response to reward anticipation increased in DeA</td>
</tr>
</tbody>
</table>

AD, alcohol dependence; DeA, detoxified alcoholic; DS, dorsal striatum; FH+, family history for alcohol use disorder; FH−, no family history for alcohol use disorder; HC, healthy control; PE, prediction error; SD, social drinker; SDP, substance-dependent patient with alcohol as primary addiction; VS, ventral striatum.

*Three DeA from Wrase et al. (114) included.
component (Andrea Reiter, Ph.D., et al., unpublished data, 2016). In terms of criterion 2b, all three studies found robust BOLD responses correlate with MFTD prediction errors in ventral and dorsal striatal regions in both patients and control subjects, without group differences. In terms of criterion 2c, no differences in MF learning rates were observed.

Two-Step Task. No imaging data on AUD patients have been reported. Among young at-risk social drinkers, correlations with measures of drinking and MFTD PE signals arise not in the VS but tentatively in the putamen with earlier drinking onset (Stephan Nebe, Dipl. psych., et al., unpublished data, 2016).

Slips of Action Task. The slips of action task compares trials on which the instrumental discriminative stimulus and the outcome conflict with trials when there is no such conflict. Habitual processes should not be affected by the incongruence. BOLD signals are more pronounced in alcohol dependence during conflict trials in the posterior putamen (63). To what extent this maps onto a MFTD account of habits is unclear because the analysis did not examine the learning phase and did not involve computing a prediction error. In addition, devaluation did not differ between groups, and the imaging data did not examine the devaluation part.

Monetary Incentive Delay Task. The monetary incentive delay (MID) task was originally developed to distinguish appetitive and consummatory responses to rewards (98). BOLD responses to anticipated rewards (122) or obtaining rewards (99) vary with the degree of reward-induced dopamine release using PET (123). Results in addiction samples have generally been mixed in part probably due to heterogeneity in task design, acute drug effects, and addiction phase (124).

In detoxified patients with AUD, two studies showed reduced ventral and dorsal striatal activity during the anticipation of rewards compared with control subjects (114,116); see Hägele et al. (125) (Table 3), whereas a third reported the converse (118). A fourth study found no effect of risk status on either outcome or anticipation stages using fMRI, but it did find that overall dopamine release was higher in subjects who had started to drink (98). Bjork et al. (115) attempted to disentangle outcome- from anticipation-related activations by inserting longer and more variable delays between trials and found that BOLD responses during anticipation did not differ between groups, whereas reward-elicited responses in the VS were larger for AUD patients than for control subjects. When further modifying the MID paradigm to separate expectations from motor preparation aspects, neither increased reward sensitivity, nor differences in anticipation were found (117).

The results with the MID are heterogeneous, possibly due to factors beyond the task such as variation in disease severity or length of disease. More importantly, learning processes have been examined, and so these studies do not directly assess any of the criteria. The task also does not exclude the influence of other (e.g., model-based) decision processes. Nevertheless, if one were to average over different learning rates, then the overall signals would be affected such that slower learning should result in smaller overall reward anticipation. This would be compatible with the reductions in VS anticipation BOLD responses seen in some of the studies (114,116,125).

DISCUSSION

Animal models make a strong case that a bias toward MFTD learning is a risk factor for developing addictions and that drugs of abuse can induce biases toward MFTD learning. The imaging literature overall does not, so far, provide much support for either process in human alcohol addiction. MFTD learning makes predictions about situations directly rewarded by the drug itself. The evidence speaking to this rests largely on cue-reactivity studies, examining cues that have been associated with the drug. To date no study has examined cue-reactivity over the course of the development of alcohol addiction. They have instead examined the ultimate consequence of learning but have not shown that these end points differ between patients with alcohol addiction and healthy control subjects in the key dopaminergic systems. Although there are experimental caveats for this null result, it is not in keeping with straightforward MFTD theory predictions. PET studies have established that alcohol releases dopamine, but PET cue reactivity has not been examined in addicted populations.

The animal literature suggests that MFTD biases could be a risk factor for the development of addiction. Learning paradigms without direct reference to drugs can be used to test this. Behaviorally, the data so far point to an alteration in a system other than the MFTD system. A bias toward MFTD learning in those with early-onset alcohol use is tantamount to (98); (Stephan Nebe, Dipl. psych., et al., unpublished data, 2016), but the learning paradigms have so far fallen short of establishing a clear alteration in MFTD processes in alcohol dependence. Paradigms not involving learning are more difficult to interpret, and the reasons for their failure to fully replicate have to be better understood.

It is worth reemphasizing that it has also been difficult to establish unambiguous alterations in MFTD learning in other settings (126). It is relatively difficult to avoid engagement of goal-directed, model-based processes in brief experimental sessions, and goal-directed influences have been shown in the VS in human BOLD responses (78,79) and animal studies (127,128). Habitual systems may also learn on a slower timescale than a single experimental session (100). Finally, phasic dopaminergic signals are present in both sign- and goal-trackers and differ mainly in the timescale at which they adapt. Measuring the learning rate at the neuronal level in humans is limited by the slow and noisy nature of BOLD signals.

One omission from the present review concerns Pavlovian-instrumental transfer paradigms. Although promising (36,111,129,130), the exact relation of the variations of this paradigm to MFTD learning is not yet clearly established.

CONCLUSIONS

MFTD learning alone is not a complete theory of (alcohol) addiction and is necessarily accompanied by additional processes. Simple alterations to the prediction-error signal
Learning and Dopamine in Alcohol Dependence

(9) cannot explain how drugs of abuse work (131), and the transformation into compulsions requires extensions and variations (9,10). There is a proliferation of findings in a wide area of brain regions outside the striatum that we have not discussed here but that likely account for critical aspects of the disorder. Expectations will be influenced by model-based interpretations of the experimental situation, instructions, and greater goals such as abstinence or controlled drinking, all of which enrich and confound the findings. Learning about the (possibly homeostatic) control of internal states (132), an important aspect of drug use (133), is yet to be experimentally probed.

MFTD operates at a computational level (134), and details such as differences in the addictive nature that arise from pharmacokinetics or pharmacodynamics elude it. It has yet to be adduced to account for the successes of drug therapies, including substitution, but it would seem particularly pertinent to examine if it might relate to the efficacy of contingency management.

Future Work: Testing MFTD Learning in AUD

We examined the current imaging evidence for MFTD learning in AUD and conclude that it can be ruled neither in nor out. For definitive answers, paradigms need to examine iterative learning in a setting that minimizes interference through other, for example, goal-directed, processes or where the slow transfer from outcome to cue can be shown explicitly or via linkage to striatal learning mechanisms (135). To test the hypothesis that the MFTD system is specifically sensitive to alcohol, alcohol should be used as outcome. To test whether alcohol influences learning driven by other outcomes, it should be examined in the intoxicated state (110,111,136). In addition, however, MF values should be able to substitute for rewards; hence tasks such as secondary conditioning whereby drug-related cues function as rewards or that capture the influence of MFTD values inside model-based decisions (137) should also be informative.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the German Research Foundation Grant Nos. DFG FOR 1617 HE 2597/13-1, HE 2597/14-1, HE 2597/15-1, RA 10472-1 (to AH, FS, and QJMH); the Swiss National Science Foundation project Grant No. 320030_153449/1 (to QJMH); and the BMBF Grant No. 10042034 (to KO).

We thank Miriam Sebold for helpful discussions. All authors report no biomedical financial interests or potential conflicts of interest. Funders had no influence on the paper.

ARTICLE INFORMATION

From the Centre for Addictive Disorders (QJMH), Department of Psychiatry, Psychotherapy and Psychosomatics, Hospital of Psychiatry, University of Zurich; Translational Neuromodeling Unit (QJMH), Institute for Biomedical Engineering, University of Zurich and Swiss Federal Institute of Technology (ETH), Zurich, Switzerland; Max Planck Institute for Human Cognitive and Brain Sciences (LD, FS), Leipzig; Charité Universitätsmedizin (LD, FL, AH), Campus Charité Mitte, Berlin; Department of Neurology (LD), Otto-von-Guericke University, Magdeburg; and Neural Information Processing Group (KO), Institute of Software Engineering and Theoretical Computer Science, Technical University, Berlin, Germany.

Address correspondence to Quentin JM Huys, M.D., Ph.D., UZH & ETH, Translational Neuromodeling Unit, Wilfriedstrasse 6, Zurich 8032, Switzerland; E-mail: qhuys@cantab.net.

Received Feb 2, 2016; revised Jun 9, 2016; accepted Jun 14, 2016.

REFERENCES

50. Nelson AJ, Killcross S (2013): Accelerated habit formation following amphetamine exposure is reversed by D1, but enhanced by D2, receptor antagonists. Front Neurosci 7:76.
Learning and Dopamine in Alcohol Dependence

adventis with alcohol use disorder. Arch Gen Psychiatry 60: 727–735.


