Excessive state switching underlies reversal learning deficits in cocaine users

Edward H. Patzelt[1,2], **Zeb Kurth-Nelson**[3], **Kelvin O. Lim**[3], and **Angus W. MacDonald III**[1,3]

[1]Translational Research in Cognitive and Affective Mechanisms Laboratory, Department of Psychology, University of Minnesota

[2]Department of Psychology, Harvard University

[3]Department of Psychiatry, University of Minnesota

[4]University College London, UK

Abstract

Background—Markers of chronic cocaine exposure on neural mechanisms in animals and humans is of great interest. The probabilistic reversal-learning task may be an effective way to examine dysfunction associated with cocaine addiction. However the exact nature of the performance deficits observed in cocaine users has yet to be disambiguated.

Method—Data from a probabilistic reversal-learning task performed by 45 cocaine users and 41 controls was compared and fit to a Bayesian hidden Markov model (HMM).

Results—Cocaine users demonstrated the predicted performance deficit in achieving the reversal criterion relative to controls. The deficit appeared to be due to excessive switching behavior as evidenced by responsivity to false feedback and spontaneous switching. This decision-making behavior could be captured by a single parameter in an HMM and did not require an additional parameter to represent perseverative errors.

Conclusions—Cocaine users are characterized by excessive switching behavior on the reversal-learning task. While there may be a compulsive component to behavior on this task, impulsive decision-making may be more relevant to observed impairment. This is important in building diagnostic tools to quantify the degree to which each type of dysfunction is present in individuals, and may play a role in developing treatments for those dysfunctions.

Keywords

reinforcement learning; Bayesian hidden Markov model; reversal learning; decision making; reward; state switching; cocaine; impulsivity; compulsivity

*To whom correspondence should be addressed: Angus W. MacDonald, III, N218 Elliott Hall, 75 East River Rd., Minneapolis, MN, 55455, angus@umn.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.

1. INTRODUCTION

Cocaine addiction is characterized by maladaptive behaviors related to an inability to arrest drug use despite increasingly harmful consequences. An emerging view suggests that addiction can be viewed as a failure of neural decision-making systems to adapt to changing reward feedback in the organism and the environment (Redish et al., 2008; Heyman, 2009). A number of behavioral tasks have been used in animals and humans to identify the specific cognitive dysfunctions underlying addiction (Liu et al., 2009; Porter et al., 2011; Woicik et al., 2011). One task that is particularly promising for studying cognitive dysfunction in cocaine addiction is the probabilistic reversal-learning task (Jentsch et al., 2002; Fillmore and Rush, 2006; Calu et al., 2007; Lee et al., 2007; Ersche et al., 2011; Lucantonio et al., 2012). The current paper examines the nature of cocaine addicts' decision-making biases on a probabilistic reversal-learning task and proposes a model to account for these biases.

In the probabilistic reversal-learning task, participants are presented with two neutral stimuli. After selecting one stimulus, they are told their choice was "correct" or "incorrect." Most of the time this feedback reflects the current correct contingency, but 20% of the time it is misleading feedback that should not affect one's strategy. The initial learned stimulus-reward association is called the acquisition stage and establishes task comprehension (Swainson et al., 2000). Following the acquisition stage the contingencies reverse (reversal stage), and this is the reversal-learning component of the task. Participants must achieve 90% accuracy over the course of several trials in order to trigger a reversal; failure to meet 90% accuracy increases the number of trials before a reversal can occur. Cocaine users have exhibited impaired performance in the number of trials required to reach reversal (Ersche et al. 2008, 2011). In principle, there are two basic reasons for impaired performance on the task: 1) perseveration, or continued responding to the old contingency, or 2) excessive switching, where subjects jump between apparent contingency beliefs.

In the past, it has been suggested that stimulant dependent subjects' deficits in achieving reversals on the task resulted from perseveration (Fillmore and Rush, 2006; Ersche et al. 2008,2011). This has in turn been linked to reduced cognitive flexibility (Kehagia et al., 2010;Lucantonio et al., 2012). As stated, there is another underlying cause for impaired ability to achieve reversals; in the case of excessive switching the subject displays a response pattern whereby they repeatedly jump between reward contingencies in the face of misleading feedback or as the result of spontaneous switching. Previous research has found both types of behavioral impairment present in stimulant dependent subjects such that they exhibit perseveration and elevated spontaneous switching relative to controls (Ersche et al., 2011). However, in cocaine dependent subjects, the two underlying possibilities for reduced behavioral performance (excessive switching versus perseveration) have not been directly compared using analysis methods that encompass both types of behavior. Additionally the reversal-learning task has been used to identify the neural correlates of reduced performance, which is important in knowing the type of mechanistic dysfunction that may be present in cocaine addicted individuals.

The oribito-frontal cortex (OFC) has a significant role in response perseveration in reversal-learning such that OFC-lesioned primate and human subjects show increased perseveration (Swainson et al., 2000; Calu et al., 2007). Furthermore orbito-frontal dopamine plays a significant role in cognitive performance such that reductions in D2 receptor availability are associated with reductions in OFC activity (Volkow et al., 1993). Genetic polymorphisms leading to D2 receptor reduction are also associated with vulnerability to cocaine addiction (Volkow et al., 2007, 2009; Shumay et al. 2012), and with exposure to cocaine (Nader et al., 2006). However polymorphisms resulting in reduction of D2 receptor availability, a dopamine profile similar to that resulting from chronic cocaine exposure, did not result in

increased response perseveration; rather, the polymorphism resulted in excessive switching on the reversal-learning task (Jocham et al., 2009). Alternatively in stimulant dependent subjects, previous research found that perseveration, but not spontaneous switching, could be alleviated by administration of the dopamine agonist pramipexole (Ersche et al., 2011). Thus, the role of D2 receptor availability is equivocal with respect to reversal-learning performance despite the reduction in D2 receptors as a result of cocaine exposure. A more comprehensive model of performance that incorporates both perseveration and excessive switching may guide our understanding of the neurobiology underlying the behavior.

Recently, reversal-learning behavior has been analyzed with hidden Markov models (HMMs), which provide a normative framework to describe decision-making in an environment with changing contingencies (Hampton et al., 2006). In previous studies with cocaine users, aberrant performance such as increased *trials to reversal* has contributed to understanding about the dysfunction underlying decision-making in the task. However, these analyses fail to constrain the mechanistic dysfunction resulting in degraded performance to a single parameter. A quantitative approach allows us to model the trial-level decision-making processes in cocaine users to further specify the underlying construct. By fitting a parameter of the HMM to each subject, it is possible to assay directly each subject's propensity for perseveration versus excessive switching, during normal performance or when there are increases in trials needed to obtain reversals. The single parameter of the HMM represents that abstract state space of the task as a trait that is allowed to vary across subjects. Thus, the HMM allows for the formalization and direct comparison of competing hypotheses about the underlying mechanisms of decision-making within the task. The current study, therefore, used the reversal-learning task to compare the hypotheses that: 1) cocaine users' are primarily struggling with a proclivity to perseverate after feedback contingencies have changed; or, 2) cocaine users' are primarily struggling with excessive switching, which is indicated by a tendency to use little evidence in selecting an anticipated source for rewards. The selected task allowed for errors of either type to occur and the use of a formal HMM model allowed us to test the source of the group differences between cocaine users and healthy controls to determine whether excessive switching or perseveration mechanisms are more closely linked to chronic cocaine usage.

2. METHODS

2.1 Participants

The study consisted of 45 cocaine users and 41 non-using controls who provided informed consent and were compensated for their participation. All participants were matched on sex, age and parental education level. Parental education level was preferable to probands' education level insofar as the disorder has an impact on educational attainment, thereby introducing a demographic confound if used as a criterion. Participants completed a Structured Clinical Interview for DSM IV disorders (SCID-IV; First et al., 1995) and were excluded for serious neurological or endocrine disorders, any medical condition or treatment known to affect the brain, documented loss of consciousness for 30 minutes or longer, loss of consciousness with neurological sequelae, DSM-IV-TR criteria for mental retardation, alcohol use of 10 or more drinks per week for women and 14 or more drinks per week for men, evidence of stroke or lesions observed on clinical MRI, and history of schizophrenia, bipolar or other psychotic disorders. Because participants underwent GABA spectroscopy in another phase of the study, the use of medications known to alter GABA brain levels (e.g., topiramate, baclofen) was also an exclusion criterion. Cocaine users were also excluded for dependence on, but not use of, other psychoactive agents and for self-reported HIV seropositivity. The cocaine users met DSM-IV-TR criteria of cocaine dependence for the past year with a minimum of 6 months of self-reported cocaine use (including intravenous,

nasal, smoking or combinations of methods of use). To qualify, users must have used cocaine at least 6 days in 0the last month, but not used within the past 48 hours. There was no difference between the cocaine users and non-using controls in the rate of non-GABA medication use (SSRI's; analgesics; antihistamines; antidiabetics; antiasthmatics; over-the-counter). All procedures were approved by the University of Minnesota Institutional Review Board.

2.2 Materials and Procedure

As part of a larger experimental battery, the participants completed a probabilistic reversal-learning task implemented in Eprime (Waltz and Gold, 2007). Two disparate gray scale stimuli were displayed simultaneously on the screen for up to 6 seconds and participants were asked to choose which pattern they believed to be correct. At the onset of the task, participants were told that one pattern square would be the "correct" pattern and the other pattern square would be the "incorrect" pattern. After each choice, "correct!" or "incorrect" appeared on the screen. Participants were told that the "correct" pattern square would be correct for a while and that the patterns would eventually switch positions. Additionally, participants were instructed as follows: "There are 2 things that will make this test trickier. Sometimes a pattern square will change; sometimes it will seem like the pattern square has changed because it says your response is incorrect, but it hasn't really changed". False feedback within the reversal-learning task occurred randomly 20% of the time.

At the onset of each block, participants were in an acquisition stage. The participants were required to achieve 90% accuracy within 10–14 consecutive responses for a reversal to occur (reversal stage) and in the absence of this goal the block would end after 50 trials. The number of trials needed to meet 90% accuracy (and corresponding reversal) defined *trials to reversal*. The reversal allotment and correct response requirement were consistent and independent between the three blocks of the task. Additionally at each reversal stage, if the subject did not achieve the 90% accuracy required for a reversal to occur within 50 trials the block ended and they were automatically advanced to the next block.

2.3 Statistical Analyses and Computational Modeling

2.3.1 Hierarchical Mixed Linear Analysis—The acquisition stage is thought to measure a baseline capacity for learning associations, and not a reversal-learning deficit (Swainson et al., 2000; Ersche et al., 2008). Subsequent reversal stages 1 and 2 are identified as the reversal-learning stages of the task; previous research excluded participants who failed to pass the acquisition stage assuming they do not understand the task demands (Swainson et al., 2000; Ersche et al., 2008, 2011). However, exclusion of participants who have not passed the acquisition stage does not allow for the inclusion of all useable observations. We reconciled the ideal of including all useable observations with this constraint by building a hierarchical mixed linear analysis to test the dependent variable *trials to reversal*. We further tested the hierarchical analysis in only those subjects who passed the acquisition stage for convergent information with previous studies. Akaike's Information Criterion and Bayesian Information Criterion were used as the design comparison statistic to assess if each subsequent linear mixed effects analysis provided significantly better information. First, random effects were tested and then fixed effects were tested for significance along with covariates.

2.3.2 Hidden Markov Modeling—To gain further specificity with respect to the dependent variable "*trials to reversal*" and reconcile perseveration versus switching, a Bayesian hidden Markov model was employed at the trial level of the data (Figure 1). This trial level approach incorporates both summary measurements of perseveration and switching. Thus, it distills the inference to a single transition parameter representing state

transition. The one parameter HMM was fit to each subject's trial level choice sequence by varying one free parameter, the symmetric transition probability of the HMM delta (δ). In the following equations the 2 patterns are denoted by "*A*" and "*B*", feedback by "*F*" and choice by "*C*". The HMM begins with the prior probabilities "*A*" is correct P(A) and " *B*" is correct P(B). The priors are set to 0.5 initially as the model has no posterior at the first trial. At each time point the priors P(A) and P(B) are recursively updated in a Markovian fashion using the posterior from the last trial and the current value of δ (equation 1):

$$P(A)=P_{(A-1)}*(1-\delta)+P_{(B-1)}*\delta \quad P(B)=1-P(A) \quad \text{(1)}$$

Then, the likelihood is set using the false feedback ratio and is defined as P(F|A,C), which is the observed feedback given "A is correct" and "A was chosen," in the case where A is the correct pattern. If the subject receives the reward, the value is 0.8; alternatively if they receive false feedback the value is 0.2. Finally the basis of the HMM is the probabilistic choice map created by Bayes rule where each time point has a new posterior defined by equation (2). This is the probability of being in State "A" given the observed feedback and the choice.

$$P(A|F,C)=\frac{P(F|A,C)*P(A)}{P(F|C)} \quad \text{(2)}$$

Enumeration was used to examine all possible values for δ between 0.01 and 0.5 with an increment of .001 to minimize the incorrect number of model predictions. For each value of δ the model was recursively conducted using each subject's data. Each subject's δ was therefore conceptualized as a trait and calculated by minimizing prediction error. The prediction error was defined by the error at each trial when comparing the model choice to the actual subject choice. The model choice on each trial was the result of the probabilistic HMM where the model chose "*A*" if the probability was above 0.5 for pattern "*A*" and "*B*" if the probability fell below 0.5 for pattern "*A*" because P(B) = 1-P(A).The transition parameter for the HMM results in a perseverative response pattern with a low delta, or an excessive switching response pattern with a high delta thereby directly contrasting the two hypothesized deficits.

2.3.3 Logistic Regressions—By design Markov models change over time due only to the trial in the immediate past (*t-1*). In the case where accuracy in the period directly following a reversal showed a pattern of perseveration, an additional mechanism of action would be required in the Markov-based model reflecting the variation in this slower temporal component of learning. A logistic regression examined the nature of the *group × stage* interaction to determine if *accuracy* in the period directly following a reversal was unique relative to trials throughout the task. However, using a trial-by-trial analysis lacks the power to detect an effect that may be present in groups of trials following a reversal. To increase sensitive to detect perseverations affecting the *accuracy* component after a reversal, we reduced the 20 trials subsequent to each reversal into *bins* containing between 3 to 5 trials (i.e., *4 trials * 5 bins = 20 trials*). This resulted in 3 regressions of *group × bin × stage* with the dependent variable of *accuracy*.

3. RESULTS

3.1 Sample Characteristics

The demographic characteristics of the sample can be found in Table 1. Age, gender, handedness, ethnicity, education, father's education, and mother's education were analyzed with education as the only significant difference between users and non-users (t(63.25) =

−3.63, p < .001). This is to be expected, in so far as education is influenced by the choice to use, therefore matching these two groups on educational attainment which continues to accumulate throughout the third decade would have risked examining non-representative portions of the two groups' distribution (Meehl, 1971). Nevertheless, education was also tested as a covariate of interest in the mixed effects hierarchical analysis.

3.2 Stage Completion

Both groups had a 98% pass rate for the acquisition stage. Of those cocaine users that passed the acquisition stage, 34 subjects (75%) passed reversal stage 1 and 30 subjects (88%) passed reversal stage 2. For non-users 40 subjects (98%) passed reversal stage 1, and 39 subjects (98%) passed reversal stage 2. The groups did not differ in their ability to pass the acquisition stage (Fisher's exact p=0.5). As expected they did differ in their ability to pass reversal stage 1 (Fisher's exact p< .001). Furthermore only subjects that passed retained values for reversal stage 2, and this was not significant (Fisher's exact p=0.4).

3.3 Mixed Effects Hierarchical Analysis

When *trials to reversal* was entered as the dependent variable in the mixed effects hierarchical analysis, Tukey HSD contrasts between acquisition stage and reversal stage 1 (z = 8.553, p < .001) and acquisition stage and reversal stage 2 (z = 5.866, p < .001) were significant (figure 2). Because reversal stage 1 and reversal stage 2 did not differ (z = −2.088, p = .092) they were averaged per participant in subsequent behavioral analyses. Random and fixed effects were tested in a hierarchical process (Table 2) with the final mixed analysis showing a significant interaction of group × stage (t(576) = −2.41, p = .016). To further the validity of this model we subsequently removed the 2 participants who did not pass the acquisition stage; again the interaction was significant (t(570) = −2.35, p = .019). The 2 participants who did not pass the acquisition stage were then removed from subsequent analyses. The effects of several covariates were tested in the hierarchical mixed analysis, including education. There was not a significant effect of these covariates, which also included age, ethnicity, gender, mother's education, and father's education.

3.4 Bayesian Hidden Markov Model

The transition parameter delta (δ) was distributed non-normally and a non-parametric rank sum test was used to compare users and non-users showing higher transition in cocaine users (Figure 3a, W = 958, p = 0.013; Mann-Whitney Wilcoxon). Prediction error signaling was used to measure model fit and prediction of each choice was on average 84% for controls and 75% for cocaine users (Sutton and Barto, 1998). Using a binomial distribution to calculate a 95% confidence interval for chance performance, 37 (82%) cocaine users and 39 (95%) controls were fit better than chance. The model fit differed significantly between groups (t(−3.3), p = 0.001). In the absence of equivalent model fit, to establish validity we tested the hierarchical mixed linear analysis using only those subjects for whom the model fit better than chance to examine whether the primary effect was still present following the modeling procedure. This is a conservative assumption, because it includes data from only those who are the least likely to respond randomly or be confused by the procedure. The *trials to reversal* interaction of *group × stage* was significant (t(545) = −3.04, p = 0.0024). For validity in subsequent analyses of *perseveration, probabilistic switching, and spontaneous switching* again only subjects with model fit better than chance were used.

3.5 Perseveration, Probabilistic and Spontaneous Switching

The results of the HMM suggested the possibility that excessive switching acted as a trait throughout the task, rather than a problem that arose simply in response to a reversal event. A logistic regression used the *group × bin × stage* interaction to test whether perseverations

were a particular problem. For 3-trial and 4-trial bins, the *group × bin × stage* interaction was non-significant (3-trial: chi^2 (14) =18.16, *p*=.20; 4-trial: chi^2(8)=7.62, *p*=.47), indicating no group difference in a slower temporal component of learning. For 5 trial bins, the *group × bin × stage* was nearly significant (chi^2 (6)=12.46, *p*=.052), however the two groups were *closest* in the five trials immediately following the reversal where any differences in perseverations should be manifest (cocaine users M(SD)=.44(.50); control =. 46(.50)). The period directly following a true reversal was next used to calculate the number of consecutive preservative errors. This is the number of trials the participant continued to choose the old (previously rewarded) pattern before adapting their behavior to the shift in reward contingencies. Examining perseverations directly in this manner, there was no difference between cocaine users and controls (t(73.58) = −0.08, p = 0.94). Thus there was no evidence from any of these analyses for a slower temporal component of learning.

Probabilistic switching, defined as the subject switching their choice on the first trial following false feedback, was calculated by dividing the total number of switches following misleading feedback by the total number of misleading feedback trials received (Figure 3b). As predicted by the HMM, this parameter did differ significantly between groups (t(59.186), p-value = 0.026). Spontaneous switching – defined as switching to the other pattern in the absence of negative feedback and calculated by dividing the total number of spontaneous switches by the total number of trials, also differed significantly between groups (t(71.889), p-value = 0.002, Figure 3c), which is also consistent with the findings from the HMM modeling.

4. DISCUSSION

In this study, we used a probabilistic reversal learning task to examine the breakdown in adaptive behavior that cocaine users exhibit (Fillmore and Rush, 2006; Ersche et al., 2008, 2011). There were three primary findings. First, consistent with previous results (Fillmore and Rush, 2006; Ersche et al., 2008, 2011; Lucantonio et al. 2012), cocaine users were significantly impaired relative to healthy controls in achieving the reversal criterion. Second, users exhibited probabilistic and spontaneous switching, but not perseveration. Third, our computational model suggested this impairment was explained by the propensity for cocaine users to switch excessively between belief states.

The impairment of cocaine users is consistent with a growing literature and supports the contention that cocaine users are impaired in a way that increases the number of trials required to reach the reversals (Schoenbaum et al., 2006; Ersche et al., 2008, 2011; Lucantonio et al., 2012). In particular, it is not always clear that perseveration is solely driving increased *trials to* reversal in stimulant dependent subjects, because spontaneous switching is present but probabilistic switching is not (Ersche et al., 2011). Our findings mirror the *trials to reversal* finding, but we found excessive switching rather than perseveration to be the cause. Moreover the lack of a perseverative effect and the presence of both spontaneous and probabilistic switching indicate that the *trials to reversal* finding is fully accounted for by the response pattern associated with transition.

The use of the an abstract state-based HMM provides information about the dysfunction in higher-order decision making processes in cocaine users that results in specific behavioral outputs from the task such as perseveration. Moreover, fitting the parameter of an HMM allowed us to directly test whether the reversal learning deficit arose from excessive switching or perseveration, as the single transition parameter adjudicates these two behaviors. Since the delta transition parameter was higher rather than lower in cocaine users, and contrary to previous suggestions (Fillmore and Rush, 2006; Ersche et al., 2008, 2011), the problem was not with perseveration but instead excessive switching between belief

states. Perseverative responding suggests an inability to integrate new information from the environment, and in particular changes in reward contingencies resulting in compulsive response patterns. On the other hand, excessive switching without adequate evidence of a change in contingencies may imply an impulsive decision making bias. It has been stated in previous work that perseveration is compulsive in nature and reflects the progression from abuse to dependence. However in our sample excessive switching as evidenced by the responsiveness to misleading feedback and spontaneous switching is the primary finding and is captured by the transition parameter. Additionally the parameter allowed us to test differences between cocaine users and controls while accommodating varied response profiles.

We observed that the model fit better for controls than for cocaine users. Here, model fit could be degraded by a higher level of guessing, poor task comprehension, poor strategy selection or a number of generalized impairments that rendered some participants' performance unpredictable. Strictly speaking, then, the above observations are relevant to that portion of the cocaine users, 82%, for whom the model predicted performance significantly better than chance. It is possible, however, that the transition parameter in many of the remaining participants was so high the version of the reversal learning task used here did not provide enough trials to calculate an accurate estimate.

The literature has begun to differentiate the underlying neurotransmitters responsible for these varied response profiles. Drugs of abuse act through neuromodulatory systems, particularly dopamine (Barrós-Loscertales et al., 2011), and also produce long-term changes in these systems. The ventrolateral prefrontal cortex has been shown to activate when a reversal occurs and subjects begin responding to the new contingency (Cools et al., 2002). This is particularly relevant in that OFC-lesioned patients show deficits both during acquisition and reversal stages of the task (Tsuchida and Doll, 2010). Furthermore the ventral fronto-striatal system plays a significant role in performance on the reversal learning task such that subjects with low baseline dopamine synthesis learned better from punishments relative to rewards (Cools et al., 2009). In a sample of marmosets, striatal dopaminergic inactivation resulted in impairments in achieving reversal criterion that were non-perseverative in nature (Clarke et al., 2011). Therefore, it is particularly interesting to consider the role of dopamine in reversal learning deficits in addicts.

Jocham and colleagues (2009) showed that subjects with DRD2 polymorphisms, resulting in a reduction in D2 receptor availability (Thompson et al., 1997; Pohjalainen et al., 1998) switched more frequently on the reversal learning task and spontaneously switched more frequently following positive feedback despite evidence that dopamine agonists reduce perseverative responding (Ersche et al., 2011). Recently it has also been shown that the PER2 gene, which plays a role in regulating striatal D2 receptor availability, may also be related to vulnerability in cocaine addiction (Shumay et al., 2012). These findings highlight the possibility that the observed decision-making bias in cocaine users may have been due to use resulting in a reduction in D2 receptor availability, or a genetic predisposition with a similar consequence. Adjudicating between these two possibilities was beyond the scope of the current project.

There are limitations to the current study. As a single task, reversal learning does not capture all aspects of compulsive (perseveration) proclivities and impulsive tendencies (excessive switching). Moreover there is a need for convergence from additional studies on the compulsive and impulsive descriptions of the behavior within the task. Similarly, the HMM model used here comes from a rapidly developing field of computational modeling and may be surpassed by a model with a new level of generality in the future. As noted above the current study was unable to incorporate factors such as genetic liability to cocaine use versus

the effects of current cocaine use. Users were asked to abstain for the 48 hours immediately preceding testing, which was validated through self-report and interviewer assessment. There was no additional biological assay to confirm the subjects were not currently affected by cocaine or other psychoactive substances. Furthermore, if variation in level of recent use before 48 hours previous to testing influenced performance, the current study was unable to examine these effects. We can be relatively assured, however, that the current results were not due to the acute effects of cocaine use directly before testing due to the structure of the procedures. Previous work indicates effect sizes for the difference between groups with cocaine positive versus cocaine negative urinalysis results had a negligible impact on decision-making in at least one report (Woicik et al., 2009). Still, future studies may benefit from quantification of consumption amounts of cocaine and other psychoactive substances to help with interpretation of findings and to determine where these substances can impact performance. Finally, two aspects of the task are important for comment. The task did not include a tangible (e.g. monetary) reinforcer, which is a difference with some other studies in this literature. In addition, a task with more trials (several 100's) would provide more observations for more precisely estimating individual's transition parameters, perhaps allowing the inclusion of participants who were remove from the current analyses. More accurate parameter estimates could then allow for a better opportunity to observe relationships to severity (days per week use) and chronicity (years of use) or other relevant individual difference measures.

A novel application of Bayesian model allowed us to discriminate between cocaine users and controls and may have implications for clinical and applied research. In parallel to the brain mechanisms previously identified in switching the model allows us to identify mechanistic dysfunction in maladaptive decision-making (Hampton et al., 2006). While some users may have compulsive features that would need addressing, others may exhibit impulsive behaviors requiring a different treatment. While it is too early to recommend direct clinical applications without further research, one could imagine measuring users' transition parameters, or proclivities to switch strategies, as a means of guiding treatment or as an outcome to target through treatment. Understanding the model of cognition used in decision-making on reversal learning tasks may therefore some day help in specifying the nature of deficits in cocaine users and approaches to reconcile poor decision making with relevant stimuli in their lives.

Acknowledgments

REFERENCES

Barrós-Loscertales A, Garavan H, Bustamante JC, Ventura-Campos N, Llopis JJ, Belloch V, Parcet MA, Avila C. Reduced striatal volume in cocaine-dependent patients. NeuroImage. 2011; 56:1021–1026. [PubMed: 21338692]

Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. Nat. Neurosci. 2007; 10:1214–1221. [PubMed: 17676057]

Calu DJ, Stalnaker TA, Franz TM, Singh T, Shaham Y, Schoenbaum G. Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. Learn. Mem. 2007; 14:325–328. [PubMed: 17522022]

Clarke HF, Hill GJ, Robbins TW, Roberts AC. Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. J. Neurosci. 2011; 31:4290–4297. [PubMed: 21411670]

Cools R, Clark L, Owen A. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. J. Neurosci. 2002; 22:4563–4567. [PubMed: 12040063]

Cools R, Frank M, Gibbs S. Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J. Neurosci. 2009; 29:1538–1543. [PubMed: 19193900]

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. Neuron. 2011; 69:1204–1215. [PubMed: 21435563]

Ersche KD, Roiser JP, Abbott S, Craig KJ, Müller U, Suckling J, Ooi C, Shabbir SS, Clark L, Sahakian BJ, Fineberg N, Merlo-Pich EV, Robbins TW, Bullmore ET. Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a D2/3 receptor agonist. Biol. Psychiatry. 2011; 70:754–762. [PubMed: 21967987]

Ersche KD, Roiser JP, Robbins TW, Sahakian BJ. Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology. 2008; 197:421–431. [PubMed: 18214445]

Fillmore MTM, Rush CR. Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. J. Psychopharmacol. 2006; 20:24–32. [PubMed: 16174667]

First, MB.; Spitzer, RL.; Williams, JBW.; Gibbon, M. Structured Clinical Interview for DSM-IV SCID. Washington, DC: American Psychiatric Association; 1995.

Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 2010; 66:585–595. [PubMed: 20510862]

Hampton AN, Bossaerts P, O'Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J. Neurosci. 2006; 26:8360–8367. [PubMed: 16899731]

Heyman, G. Addiction: A Disorder of Choice. Cambridge, Massacheusetts: Harvard University Press; 2009.

Jentsch JD, Olausson P, De La Garza R, Taylor JR. Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. Neuropsychopharmacology. 2002; 26:183–190. [PubMed: 11790514]

Jocham G, Klein TA, Neumann J, Von Cramon DY, Reuter M, Ullsperger M. Dopamine DRD2 polymorphism alters reversal learning and associated neural activity. J. Neurosci. 2009; 29:3695–3704. [PubMed: 19321766]

Kehagia A, Murray GK, Robbins TW. Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr. Opin. Neurobiol. 2010; 20:199–204. [PubMed: 20167474]

Lee B, Groman S, London ED, Jentsch JD. Dopamine D2/D3 receptors play a specific role in the reversal of a learned visual discrimination in monkeys. Neuropsychopharmacology. 2007; 32:2125–2134. [PubMed: 17299511]

Liu S, Heitz RP, Bradberry CW. A touch screen based Stop Signal Response Task in rhesus monkeys for studying impulsivity associated with chronic cocaine self-administration. J. Neurosci. Methods. 2009; 177:67–72. [PubMed: 18948136]

Lucantonio F, Stalnaker T, Shaham Y, Niv Y, Schoenbaum G. The impact of orbitofrontal dysfunction on cocaine addiction. Nat. Neurosci. 2012; 15:358–366. [PubMed: 22267164]

Meehl P. High school yearbooks: a reply to Schwarz. J. Abnorm. Psychol. 1971; 77:143–148.

Nader MA, Morgan D, Gage HD, Nader SH, Calhoun TL, Buchheimer N, Ehrenkau R, Mach RH. PET imaging of dopamine D2 receptors during chronic cocaine self-administration in monkeys. Nat. Neurosci. 2006; 9:1050–1056. [PubMed: 16829955]

Pohjalainen T, Rinne JO, Någren K, Lehikoinen P, Anttila K, Syvälahti EK, Hietala J. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol. Psychiatry. 1998; 3:256–260. [PubMed: 9672901]

Porter JN, Olsen AS, Gurnsey K, Dugan BP, Jedema HP, Bradberry CW. Chronic cocaine self-administration in rhesus monkeys: impact on associative learning, cognitive control, and working memory. J. Neurosci. 2011; 31:4926–4934. [PubMed: 21451031]

Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. Behav. Brain Sci. 2008; 31:415–437. discussion 415–437. [PubMed: 18662461]

Schoenbaum G, Roesch MR, Stalnaker TA. Orbitofrontal cortex, decision-making and drug addiction. Trends Neurosci. 2006; 29:116–124. [PubMed: 16406092]

Shumay E, Fowler JS, Wang G-J, Logan J, Alia-Klein N, Goldstein RZ, Maloney T, Wong C, Volkow ND. Repeat variation in the human PER2 gene as a new genetic marker associated with cocaine addiction and brain dopamine D2 receptor availability. Transl. Psychiatry. 2012; 2:e86. [PubMed: 22832851]

Sutton, RS.; Barto, AG. Reinforcement Learning. Cambridge, MA: MIT Press; 1998.

Swainson R, Rogers R, Sahakian B. Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. Neuropsychologia. 2000; 38:596–612. [PubMed: 10689037]

Thompson J, Thomas N, Singleton A, Piggott M, Lloyd S, Perry EK, Morris CM, Perry RH, Ferrier IN, Court JA. D2 dopamine receptor gene DRD2 Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics. 1997; 7:479–484. [PubMed: 9429233]

Tsuchida A, Doll B. Beyond reversal: a critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. J. Neurosci. 2010; 30:16868–16875. [PubMed: 21159958]

Volkow ND, Fowler JS, Wang GJ, Baler R, Telang F. Imaging dopamine's role in drug abuse and addiction. Neuropharmacology. 2009; 56:3–8. [PubMed: 18617195]

Volkow ND, Fowler JS, Wang GJ, Hitzemann R, Logan J, Schlyer DJ, Dewey SL, Wolf AP. Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. Synapse. 1993; 14:169–177. [PubMed: 8101394]

Volkow ND, Fowler JS, Wang G-J, Swanson JM, Telang F. Dopamine in drug abuse and addiction: results of imaging studies and treatment implications. Arch. Neurol. 2007; 64:1575–1579. [PubMed: 17998440]

Waltz JA, Gold JM. Probabilistic reversal learning impairments in schizophrenia: further evidence of orbitofrontal dysfunction. Schizophr. Res. 2007; 93:296–303. [PubMed: 17482797]

Woicik PA, Moeller SJ, Alia-Klein N, Maloney T, Lukasik TM, Yeliosof O, Wang G-J, Volkow ND, Goldstein RZ. The neuropsychology of cocaine addiction: recent cocaine use masks impairment. Neuropsychopharmacol. 2009; 34:1112–1122.

Woicik PA, Urban C, Alia-Klein N, Henry A, Maloney T, Telang F, Wang G-J, Volkow ND, Goldstein RZ. A pattern of perseveration in cocaine addiction may reveal neurocognitive processes implicit in the Wisconsin Card Sorting Test. Neuropsychologia. 2011; 49:1660–1669. [PubMed: 21392517]
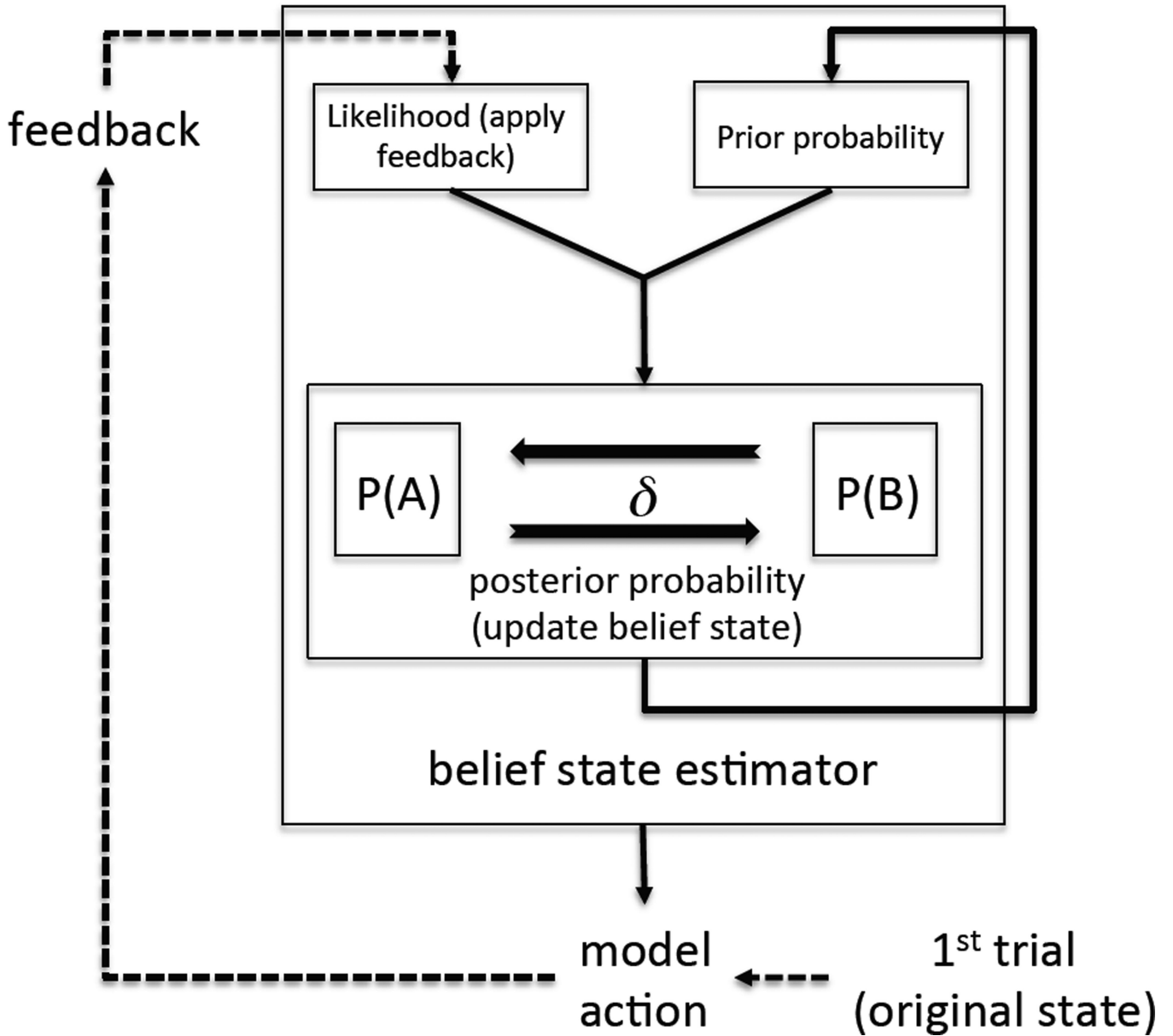
Figure 1. Hidden Markov model schematic
The starting point is the *1st trial*, which is followed by the model action. The feedback is the actual feedback the subject received in the task - on each subsequent trial the belief state is updated by the Bayesian belief state estimator, and the model action follows. We found the transition parameter delta such that the probability of the chosen model was maximal given the observed data.
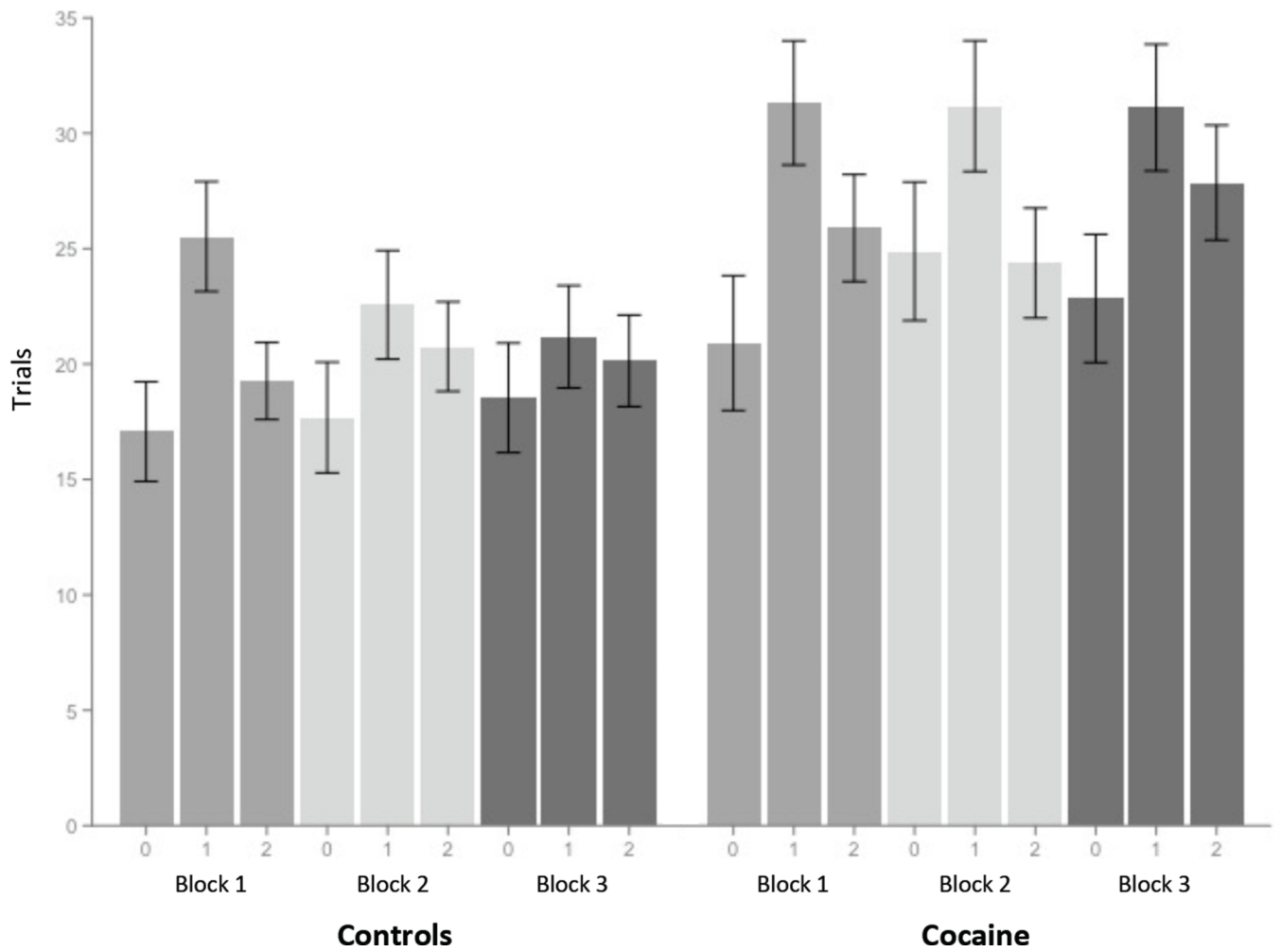
Figure 2. The 1st level of the x-axis in this figure is the reversals within each block
The 0 indicates the acquisition stage, the 1 indicates the number of trials before the 2nd reversal, and the 2 indicates the number of trials before the end of that block. The 2nd level of the x-axis represents the 3 blocks within the task. Participants performed the task with 3 distinct pairs of fractal patterns - 1 pair for each block.
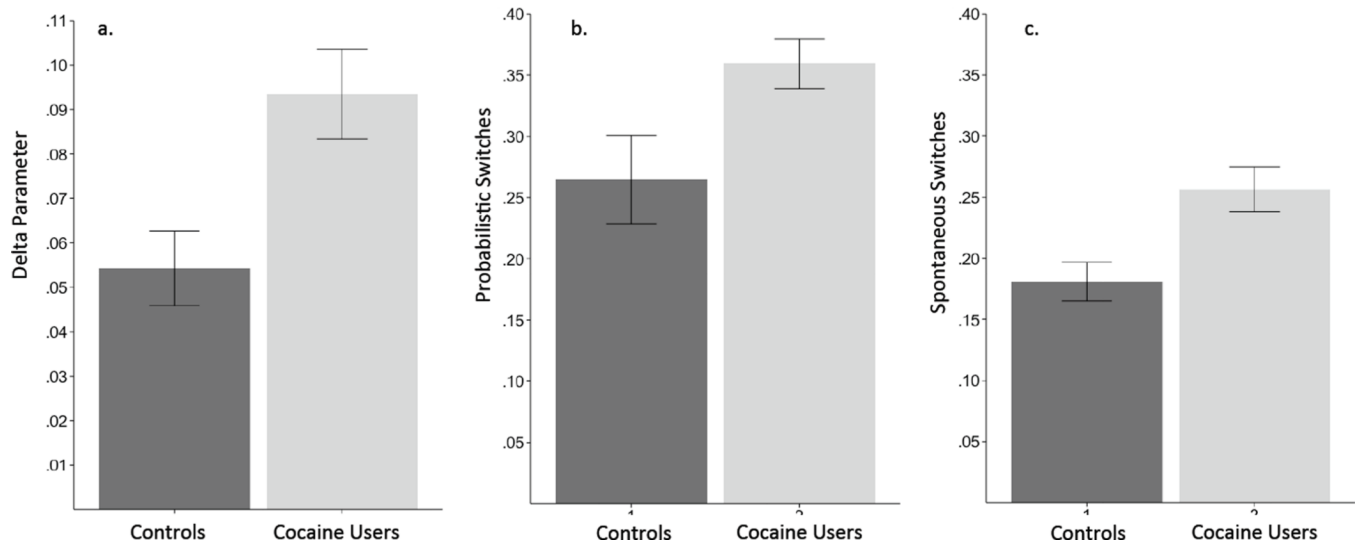
Figure 3. Variables related to switching
3a: transition (delta) parameter estimated from the Hidden Markov model based on each individual's decisions; 3b: number of trials on which subjects switched based on false feedback relative to the number of trials observed; 3c: number of trials switched following correct feedback on non-misleading trials. Note: the scale changes between the delta parameter and the 2 switching measures.

Table 1

Sample Demographics.

	Cocaine	Controls	Significance Testing
Sample Size	45	41	
Age	41.0 (6.9)	40.0 (7.4)	$t(82) = .31$, ns
Gender (% male)	77%	75%	$X^2(1) = \sim 0$, ns
Handedness (% right)	90%	85%	$X^2(1) = \sim 0$, ns
Ethnicity (% Caucasian)	45%	68%	$X^2(6) = 7.4$, ns
Education (in years)	12.9 (1.5)	14.5 (2.3)	**$t(82) = -3.55$, p<.001**
Father Education	2.8 (2.2)	3.3 (1.7)	$t(78) = -1.03$, ns
Mother Education	2.4 (1.7)	2.8 (1.5)	$t(82) = -1.16$, ns
Years of use (mean, range)	14.75 : 2–27	NA	NA
Days per week of use	3.4 (1.6)	NA	NA

Notes: Parental and subject education measured on different scales. Subjects: 12 = 12th grade, 14 = Associate's degree, 16 = Bachelor's degree.

Parental: 2 = 12th grade, 3 = Associate's degree, 4 = Bachelor's degree.

Table 2

Mixed hierarchical analysis

Hierarchical Testing Procedure	AIC	BIC	LogLike	Df	X^2	p-value
RE: Group + Stage (Linear Model)	5269	5287	−2630	1	NA	NA
RE: Added R.E. of Subject (REML)	5146	5168	−2568	5	124.76	p<.001
RE: Added Block (REML)	5148	5175	−2568	6	<.001	0.9940
FE: Group + Stage (ML)	5153	5175	−2571	5	NA	1.0000
FE: Group * Stage (ML)	5149	5176	−2568	6	5.47	**0.0190**

Notes: REML – Restricted Estimation Maximum Likelihood, ML – Maximum Likelihood, AIC – Akaike Information Criterion, BIC – Bayes Information Criterion, LogLike – Log Likelihood Ratio, Df – Degrees of Freedom, X^2 - Chi Squared..