Title: Dynamic, behavior-dependent interactions between dorsal striatal dopamine and glutamate release predict cognitive flexibility and punishment resistant cocaine use

Abstract:

Cognitive inflexibility covaries with substance use disorder (SUD) risk. To determine if there is a neural relationship between these phenomena, glutamate and dopamine release in the dorsomedial (DMS) and dorsolateral (DLS) striatum were measured as rats performed a discrimination and strategy switching test. Elevations in glutamate release, with reductions in dopamine, at trial initiation (DLS) and prior to choice (DMS and DLS) predicted fast strategy switching and punishment sensitive cocaine seeking. Elevations in DLS and DMS dopamine release at these respective timestamps predicted slow switching and punishment resistance. Orbitofrontal cortex and intralaminar thalamus were significant contributors to DLS and DMS glutamate release, but their relative contributions differed between rats that were fast or slow strategy switchers, and in how they affected behavior. As such, these data describe a neural signature of flexibility and associated circuitry that could be used to predict and treat SUDs in humans.

Introduction: Drug-taking despite threat of punishment or consequences is a hallmark of substance use disorders (SUDs) and may arise from the development of inflexible drug-seeking behaviors that can be triggered by environmental cues and contexts[1]. The presence of general cognitive inflexibility is also a trait marker that can predict the development and severity of substance use in rodents and non-human primates[2-5]. For example, poor reversal learning in rodents[3] is associated with faster acquisition of cocaine self-administration[4], higher cocaine self-administration rates[4], high levels of alcohol self-administration[5], and cue-induced reinstatement of alcohol[5]. Likewise, poor set-shifting in primates is associated with increased preference for and intake of alcohol[6]. In humans it is more difficult to parse out whether cognitive deficits are pre-existent to or caused by drug use. However, family history of substance use is associated with deficits in cognitive flexibility performance, among other cognitive measures (reviewed in[2]). For example, subjects with a family history of alcohol use disorder display less improvement in cognitive flexibility performance across development compared to subjects with no family history[7]. Thus, we hypothesized that a common neural mechanism could contribute to both cognitive inflexibility and risk for punishment resistant substance use.

We and others have shown that glutamate (Glut) and dopamine (DA) release in the dorsomedial striatum (DMS) is involved in the regulation of goal-directed, flexible behavior, and that the dorsolateral striatum (DLS) is involved in the initiation and execution of well-learned, stable behavior. DA release in DMS is necessary for strategy

switching specifically[8] and cognitive flexibility, in general[9,10], and mice self-stimulating dopamine terminals in DMS show improved reversal learning[11]. Glut release in DMS from basolateral amygdala has also been shown to bidirectionally regulate goal-directed, flexible behavior[12], with DMS activity being necessary for flexible adaptation of behavior to changing reward contingencies[13]. In contrast, DA[14] and Gut[15,16] release in the DLS increase prior to the execution of a well-learned action, coinciding with a necessary increase in DLS activity[17]. Furthermore, DA release in the DLS develops and strengthens over time in response to drug-related cues[18]. Our lab demonstrated that the development of outcome-insensitive, punishment resistant cocaine use coincides with a reduction in calcium-indicated neural activity in the DMS, likely due to reduced glutamate input, and increased DLS DA release to cocaine-paired cues across training[19].

Despite the known interaction between substance use phenotypes, cognitive inflexibility, and Glut/DA dynamics in the DMS and DLS, no studies have measured how DMS and DLS DA and Glut interact during tests of cognitive flexibility or determined if those measures predict punishment resistant drug taking within the same animal. Thus, we took advantage of natural variation in rat cognitive flexibility, as assessed by a strategy switching test, to determine if poor cognitive flexibility was associated with the propensity to develop punishment resistant cocaine use across two different self-administration schedules. Furthermore, we used multi-channel fiber photometry to simultaneously measure Glut and DA release in the DMS and DLS of the same rat as they performed a discrimination paradigm and then a strategy switching test. We determined if neural signatures of cognitive flexibility could be identified and if those neural signatures could predict punishment resistant cocaine use. We found that poor performance on the strategy switching test predicted resistance to punishment, independent of self-administration schedule. We also found that coordinated fluctuations in DMS and DLS Glut and DA at key points during discrimination and strategy switching predicted both strategy switching performance and sensitivity or resistance to punishment. Finally, in a separate set of rats, we used pathway-specific designer receptors exclusively activated by designer drugs (DREADDs) to inhibit known Glut inputs to the dorsal striatum to determine the source of the measured Glut signal and whether it was necessary for strategy switching or punished cocaine taking.

Results:

Photometry recordings: Prior studies have linked cognitive inflexibility to increases in cocaine taking[4]. However, it is not known whether cognitive inflexibility can predict punishment resistant cocaine use or whether neural correlates measured during the cognitive flexibility test could link and predict the two behaviors. To test this, male and female rats were first infused with proteins that fluoresce in response to the presence of DA (GRABDA) and Glut (igluSnFR) and implanted with photometry fibers in the DMS

and DLS. Rats featured one DMS infusion/fiber and one DLS infusion/fiber per hemisphere, providing a readout of DMS and DLS DA and Glut simultaneously within the same rat. All rats were verified to have sufficient presence of GRABDA (yellow) and igluSnFR (green) around the photometry fibers in the DMS and DLS (Fig. 1a).

Strategy switching performance was normally distributed: Following recovery from surgery, we first determined whether cognitive inflexibility would predict punishment resistant cocaine use. Naïve male and female rats learned to nose-poke in an operant chamber using a visual cue to receive a sugar pellet reward (pre-training, Fig S1). Following successful completion of the pre-training phase, rats performed a complex discrimination paradigm across 10 days (Fig. 1b). Rats discriminated between 2 active ports using a visual cue (visual strategy), while ignoring the position of the cue (response strategy; Fig. 1c). On the 11th day, the rats performed a strategy switching test requiring them to inhibit responding via the previously rewarded strategy (visual) and learn the newly rewarded strategy (response). We quantified the number of errors and total trials that it took the rats to reach the performance criterion of 10 consecutive trials correct using this newly rewarded response strategy. Performance was normally distributed (mean$\pm$SEM errors: 61.5 $\pm$ 5.1, trials: 167.5 $\pm$ 13.7) and featured a range of responses (Fig. 1e). Males (N=10, errors: 58.2 $\pm$ 6.3) and females (N=12, errors: 64.2 $\pm$ 7.9) were part of the same interval (95% CI males: 44.0-72.4, females: 46.8-81.6), so sexes were combined for these studies. We then performed a median split based on errors such that rats that had greater than 61.5 errors (N=6 males, 5 females) were qualified as “slow switchers” and rats with fewer than 61.5 errors (N=4 males, 7 females) were qualified as “fast switchers” (Fig. 1e). The number of pre-training days or the presence or absence of a side bias did not predict strategy switching performance, as these were evenly distributed across fast and slow switchers (Fig. S1c-d).

Slow strategy switching predicts cocaine punishment resistance: After the strategy switching test had concluded, all rats were implanted with jugular vein catheters and trained to self-administer cocaine. For each lever press (FR1) rats received a light/tone cue and a 0.5 mg/kg intravenous infusion of cocaine (Fig. 1f). Slow switchers reached self-administration criterion (see methods) in fewer days than fast switchers, but this effect did not reach statistical significance (P=0.099; Fig. 1g). Rats with continuous access to cocaine (Fig. 1h) took more cocaine infusions per day (P<0.0001) due to the nature of the continuous access schedule, but intermittent access rats took more cocaine per minute (P<0.0001; Fig. 1i). Slow switchers with continuous access took more cocaine infusions per day than fast switchers (P=0.02); however, fast switchers with intermittent access took more cocaine infusions/minute (P=0.02). On the 11th day all rats were given continuous access to cocaine and they took comparable numbers of cocaine infusions (mean range 50 to 58 infusions across all groups; Fig. 1j). On day 12, all rats performed a punishment resistance test (Fig. 1k), where cocaine infusions were

paired with a foot shock that increased in intensity every 10 minutes on a tenth-$\log_{10}$ scale from 0.13 to 0.79 milliamps[20-22]. Slow switchers took significantly more punished cocaine infusions overall compared to fast switchers ($F(1,48)=11.29$, $P=0.0015$), but this effect was most apparent during the medium shock intensities ($P=0.0011$; Fig. 1l). This effect was independent of self-administration schedule in the slow switchers ($F(3,56)=6.44$, $P=0.0008$). Intermittent access in the fast switcher group did produce an intermediate phenotype, elevating punished cocaine infusions earned relative to the continuous-fast group ($P=0.06$), and was no longer significantly reduced relative to the two slow switcher groups (p's > 0.2), similar to what's been shown previously[20,23,24]. Thus, we determined that poor cognitive flexibility, as demonstrated by slower strategy switching, was associated with greater punishment resistant cocaine use, independent of a cocaine access schedule that has been shown to produce resistance to punishment[20].

Discrimination paradigm: Fast switchers exhibited elevated DLS Glut signaling at trial initiation. To determine if DMS and DLS Glut and DA release measures would be different during performance of the discrimination paradigm in rats that went on to be fast and slow switchers, we performed photometry recordings on discrimination days 1, 5, and 10. We aligned the signal to five separate task-relevant behavioral timestamps: when the rat initiates a trial (Trial Start), a correct or incorrect choice, and pellet retrieval or not (No Pellet; see Fig. 1b). Rats that went on to be fast switchers had significantly different Glut and DA dynamics in DMS and DLS at trial start on day 1 ($F(3,60)=3.6$, $P=0.018$, Fig. 2a) and day 5 ($F(1,19)=4.8$, $P=0.041$, Fig. 2c), compared to slow switchers. By quantifying area under the curve for each release measure, we determined that fast switchers had elevated DLS Glut on Day 1 ($P=0.015$, Fig. 2b) and Day 5 ($P=0.006$, Fig 2d), relative to slow switchers. Another interesting element of the data was not only measuring them individually, but in relation to each other. Thus, we also quantified Glut release in relation to DA release within the DLS and DMS. Interestingly, this relative measure was often more predictive of future behavior than either measure individually (see below). Fast switchers developed a greater difference between Glut and DA in DLS that was present on Day 1 ($P=0.12$) and reached significance by Day 5 ($P=0.016$). Note that these differences in signaling were observed between groups despite no significant differences in behavioral performance (Fig. 1c). By Day 10, the signaling differences were no longer apparent, as the slow switchers also displayed elevated Glut and high Glut-DA ratios in the DLS at trial start (Fig. 2e, f).

Discrimination paradigm: Fast switchers displayed elevated DLS Glut release prior to a correct choice and at pellet retrieval. When we aligned the signal to when the rat got a choice correct and then received a sugar pellet, we also detected differences between rats that went on to be fast and slow switchers. On Day 1, the rats that went on to be fast switchers had significantly different Glut-DA dynamics in DMS

and DLS at pellet retrieval (F(1,20)=9.3, P=0.006, Fig. 3a). Area under the curve analysis found that this was driven by an elevation in DLS Glut in the fast switchers (P=0.026, Fig 3b). On day 5, fast switchers were significantly different from slow switchers prior to a correct choice (F(1,19)=4.9, P=0.039, Fig. 3c), which was driven by elevations in DLS Glut in the fast switchers (P=0.043, Fig. 3d). Again, by day 10, fast and slow switchers showed similar patterns of dorsal striatal signaling (Fig. e-f). There were no differences prior to or following an incorrect choice (no pellet) across any of the 3 discrimination days (Fig. S2).

Strategy Switching Test: Fast and slow switchers differed in DLS Glut and DA release at trial start. We predicted that DMS and DLS DA and Glut signals would change as rats progressed through the strategy switching test, so we took 3 “snap shots” across the test. We separated signals from the first 30 trials (early), during ~trials 60-90 (middle), and during the last 20 trials as the rats were reaching the performance criterion of 10 consecutive correct (late). To determine whether release measures when the rat initiated a trial during the strategy switching test would be different between rats that would go on to be fast and slow switchers, we compared DA and Glut release in the DMS and DLS across the early, middle, and late part of the strategy switching test. Rats that went on to be fast switchers began to show different Glut and DA dynamics in the DMS and DLS at trial start early in the test (Fig 4a-b), but this interaction reached significance during the middle portion of the test (F (3,54) =3.95, P=0.013, Fig 4c). This effect was driven by a dynamic interaction between the release measures in the fast switchers, compared to a relatively quiescent response in the slow switchers. Comparing area under the curve (Fig. 4d), we found an elevation in DLS Glut release in the fast switchers (P=0.034), an elevation in DLS DA release in the slow switchers (P=0.004), and a greater difference in Glut relative to DA release in the DLS of fast switchers (P=0.012). The difference between the dynamic response in fast switchers and the quiescent response in the slow switchers continued late into the test (F (3, 54) = 4.12, P=0.011; Fig. 4e). Again, this was driven by elevated DLS Glut (P = 0.008) and DLS Glut-DA difference (P = 0.013) in fast switchers and elevated DLS DA in slow switchers (P=0.037).

Strategy Switching Test: Fast switchers displayed test epoch-dependent changes in DA and Glut release prior to and after correct choices that were not seen in slow switchers. To determine whether specific neural signatures prior to choice and during pellet retrieval during the strategy switching test would be different between rats that would go on to be fast and slow switchers, we compared DA and Glut release in the DMS and DLS across the early, middle, and late part of the strategy switching test. Rats could make a correct choice in two ways; they could be correct when a response following the previous strategy happened to coincide with the new strategy or they could be correct following the new strategy. For example, if the light cue is in the left port

when the rewarded strategy has switched from “the light is correct” to “left is correct,” the rat will be correct and receive a pellet reward even if they are still following the previous visual strategy (Fig 5a). If the rat chooses the left port when the light cue is on the right, then it is clear that they made a correct choice following the new response strategy. We hypothesized that DMS and DLS release measures could be different depending on whether the rat got a choice correct following the previous or new strategy, so we analyzed these correct choices and their subsequent pellet retrievals separately.

Early strategy switching test: Rats that went on to be fast switchers had significantly different Glut/DA dynamics in DMS and DLS when they made a correct choice following the previous strategy ($F(3, 60) = 4.72$, $P=0.005$; Fig. 5b). Comparing area under the curve, we found an elevation in DMS Glut release in the fast switchers ($P=0.039$), and an elevation in DMS DA release in the slow switchers ($P=0.043$). Fast switchers also had a greater difference in Glut relative to DA release in both the DLS ($P=0.027$) and DMS ($P=0.001$). There were no differences during pellet retrieval and consumption during the early portion of the test. There were only a few instances of rats making a correct choice following the new strategy during the early portion of the test, as most rats continue to respond via their well-learned visual strategy. Thus, even though patterns of release were similarly different between fast and slow switchers, high variability and low power prevented statistically significant effects at “new strategy correct” (Fig 5c).

Middle strategy switching test: Release measures during this portion of the test were similarly different between fast and slow switchers when they made a correct choice following the previous strategy ($F(3, 54) = 4.020$, $P=0.012$; Fig. 5d). Fast switchers featured a greater difference in Glut relative to DA release in both the DLS ($P=0.085$) and DMS ($P=0.029$), and slow switchers displayed elevated DA release in the DLS ($P=0.081$). Different from early in the test, however, fast switchers developed a large elevation in DMS DA release during Pellet Retrieval that was not seen in the slow switchers ($P=0.05$). The patterns in release measures were similar prior to making a correct choice following the new strategy (Fig 5d), with greater DLS Glut release prior to a correct choice ($P=0.05$) and a greater degree of activity during Pellet Retrieval in the fast switchers. Again, however, because there were fewer of these types of responses, there was less power to detect statistical significance.

Late strategy switching test: Different from the two earlier snapshots, these tracings are related to how the rat is performing in the test, rather than being aligned to a specific range of trials. Release measures during this portion of the test still appeared somewhat different; however, the dynamics of the slow switchers prior to a correct choice and during pellet retrieval became similar enough to the fast switchers that there were no

longer any statistical differences. There were no differences prior to or following an incorrect choice (no pellet) across any time epoch (Fig. S3).

DLS and DMS Glut release changes both converge and diverge across the discrimination paradigm and strategy switching test: Prior literature has demonstrated evidence for a competitive balance between activity in the DLS and DMS, with DLS activity ascribed to the performance[18] or initiation[14,17] of a well-learned or habitual behavior and DMS activity ascribed to performance of a goal-directed behavior[13,19,25,26]. Interestingly, however, there is also evidence in the literature of cooperation between these two regions in learning[27-30]. To examine this potential contradiction in the current experiment, we analyzed DLS and DMS Glut release across time at the two behavioral timestamps that produced the effects described above: trial start and correct choice. At trial start there was a main effect of dorsal striatum region ($F(1,20)=6.40$, $P=0.0199$; Fig 6a), such that DLS Glut release was greater than DMS Glut release throughout performance of the discrimination paradigm and into the early portion of the strategy switching test. DLS Glut increased throughout the strategy switching test reaching a peak late in the test. Alternatively, DMS Glut peaked during the early and middle portions of the strategy switching test, resulting in a difference between DLS and DMS Glut on day 1 of the discrimination paradigm ($P=0.0953$) and late in the strategy switching test ($P=0.0015$), but not at other time points. DMS and DLS Glut release remained even in slow switchers across all time points, demonstrating a convergence of the two rather than a DLS dominance that is characteristic of behavioral initiation of a well-learned discrimination behavior. (Fig 6b). Prior to a correct choice, there was a main effect of dorsal striatum region ($F(1,20)=15.3$, $P=0.0009$) and a marginal interaction between region and time ($F(8,150)=1.96$, $p=0.055$; Fig. 6c). DLS Glut, but not DMS Glut, increased prior to a correct choice across performance of the discrimination paradigm, such that DLS Glut was greater than DMS Glut by day 10 ($P=0.024$). This difference was no longer present early in the strategy switching test when making a correct choice using the previous strategy as Glut release increased in both regions, but particularly in the DMS. Uniform release in both the DMS and DLS was maintained during the middle portion of the test at this timestamp but then separated again late in the test ($P=0.0245$) as DLS Glut increased and DMS Glut decreased. When making a correct choice following the new strategy, there was a different pattern. During the early portion of the test, the difference between DLS and DMS Glut seen on day 10 went away, and this was again due to an increase in DMS Glut. However, there was a quick re-emergence of the separation of DMS and DLS Glut release as DLS Glut increased and DMS Glut decreased during the middle of the test ($P=0.0027$) and then became further pronounced late in the test ($p<0.0001$). These patterns were absent in the slow switchers, who despite having similar release measures across the discrimination paradigm to fast switchers, featured a surprising reduction in both DMS and DLS Glut early in the strategy switching test relative to their

day 10 release levels (Fig. 6d). Thus, it appears that in fast switchers, DMS Glut release is observed when contingencies change and the action-outcome relationship becomes uncertain, while DLS Glut dominates when actions are well-learned and outcomes are certain, whereas slow switchers fail to engage these striatal signaling dynamics under changing conditions.

Specific differences in DLS and DMS Glut and DA release measures predict strategy switching performance and punishment resistant cocaine use: To determine if any of the observed differences in release measures described above would be predictive of future strategy switching performance (strategy switching errors) or cocaine punishment resistance (punished cocaine infusions during the medium bin), we first computed the correlation matrix for release measures across all rats. We analyzed correlations on any release measure that was different between fast and slow switchers as defined by $P < 0.1$. Similar to what we showed in Fig 1k, we found that the number of errors a rat committed during the strategy switching test was indeed correlated with the number of punished cocaine infusions they took ($r=0.51$, $P=0.041$). We also found that the majority of the measured differences in DMS and DLS DA and Glut release between fast and slow switchers were significantly correlated with the number of strategy switching errors the rat would go on to commit or the number of punished cocaine infusions the rat would go on to take (Table 1). We also found that several measures taken during the middle portion of the strategy switching test were significantly correlated with both strategy switching performance and punishment resistant cocaine use (Table 1). Second, we used a Chi-Square test to identify which release measures were the strongest predictors of strategy switching performance and punishment sensitivity. For this test, we focused on the behavioral timestamps and timepoints that were the most different between fast and slow switchers, namely early in the discrimination paradigm (day 1 and 5) and the early and middle portions of the strategy switching test (see Fig 7a). The Chi-Square test identified 9 measures with a strong prediction strength ($-\log(p)>1.5$) for strategy switching performance and 7 for punishment sensitivity (Fig 7b). Interestingly, many of these measures (metric 1, 3-7, and 9) are top predictors for both strategy switching and punishment sensitivity (Fig 7b), and two metrics (3 and 7) were identified by both the correlation analysis and Chi-Square test as such. To determine if we could achieve 100% predictability by combining neurotransmitter metrics, we created a linear discriminant analysis model. The model examined all possible combinations of the 14 metrics and found that two different combinations of 2 metrics yield 100% predictability for strategy switching performance (Fig 7 c-d) and two different combinations of 5 metrics yield 100% predictability for punished cocaine use (Fig7 e-f). Eight-fold cross-validation was used to evaluate performance of the classification models. Furthermore, the models were able to perfectly categorize the strategy switching performance of all rats with 16 different combinations of 3 metrics, 89 different combinations of 4 metrics, and 229 different

combinations of 5 metrics (not shown). The models were also able to categorize better than 90% of all rats punished cocaine use with 2 different combinations of 2 metrics, 1 combination of 3 metrics, 6 different combinations of 4 metrics, and 27 different combinations of 5 metrics. Thus, the linear discriminant analysis provides strong evidence that neurotransmitter release measures taken during several distinct behavioral timestamps can be combined to describe a neural signature of flexibility that can predict future strategy switching and punished cocaine use.

Orbital frontal and intralaminar thalamic projections to dorsal striatum are responsible for timestamp-specific Glut release and are necessary for fast strategy switching and sensitivity to cocaine punishment: The relationship between Glut and DA release in the dorsal striatum across several behavioral timestamps was a strong predictor of both strategy switching performance and punished cocaine use. Specifically, elevations in Glut, relative to DA, at several behavioral timestamps led to fast switching and punishment sensitivity, whereas elevations in DA, relative to Glut, led to slow switching and punishment resistance. To determine the source(s) of the Glut measured at specific behavioral timestamps and determine whether they are necessary for fast switching and punishment sensitivity, we inhibited two Glut inputs to the DMS and DLS with pathway-specific DREADDs and measured the effects on the photometry signal, discrimination performance, strategy switching, and punished cocaine use. We chose to inhibit the ventral orbital frontal cortex (OFC) and intralaminar thalamus (IN Thal) projections to dorsal striatum because these regions send dense glutamatergic projections to the dorsal striatum that are separate, but parallel to the DMS and DLS. They are also both known to be involved in some version of attention or behavioral switching[15,31-39]. We hypothesized that blocking OFC inputs to DMS and DLS would reduce the Glut release seen prior to a correct choice[33], whereas blocking IN Thal inputs would reduce the Glut release measured at trial initiation[15]. To test this, we infused a Cre-dependent, Gi-coupled DREADD into the OFC, a Cre-dependent KOR DREADD into the IN Thal, and a retrograde CRE virus into the DMS and DLS (Fig. 8a-b). Thus, injection of CNO, the designer drug for the Gi-coupled DREADD, would result in the inhibition of OFC inputs to dorsal striatum, and injection of SalB, the designer drug for the KOR DREADD, would inhibit IN Thal inputs to dorsal striatum. The same photometry viruses and fibers as used in the first experiment were also infused and implanted into the DMS and DLS, allowing the measurement of Glut and DA in response to inhibition of OFC and/or IN Thal inputs. Rats then underwent the discrimination paradigm and strategy switching test as in the first experiment, and we performed photometry recordings as above. Our classification model provided hundreds of different combinations of neurotransmitter metrics that perfectly categorized the strategy switching performance of every rat from the first experiment. We selected three 3-metric models that combined metrics that were taken from these rats when they were unmanipulated (vehicle conditions): discrimination day 1, day 5, and early in the

strategy switching test (see Fig S4). If 2 out of 3 models identified the rat as a fast switcher then we categorized them as a “putative” fast switcher, and vice versa. The model identified 7 putative fast (3 male/ 4 females) and 7 putative slow (5 male/ 2 female) switchers (Fig S4). We analyzed the putative fast and slow switchers separately as we did in the first experiment. In support of this method and our initial findings, the resulting mean photometry tracings from the “vehicle” conditions recapitulated our findings from the first experiment (see below).

Discrimination paradigm: We found the greatest difference between eventual fast and slow switchers on discrimination day 5, so between days 4-6 rats underwent 1 day with a vehicle injection, one day with a CNO injection, and one day with a SalB injection (order counterbalanced across rats) to compare the effects on Glut release both within and across rats. Discrimination performance was impaired in the putative slow switchers, such that there was a significant reduction in the number of correct responses on the day after the CNO/SalB injection days (day 7: $p=0.022$; Fig 8c). Interestingly, learning in the putative slow switchers was reset to some degree as their mean percent trials correct dropped from 91% on day 6 to 84% on day 7, similar to what it had been on days 1 and 2 (82%). Discrimination performance was not affected in the putative fast switchers. In accordance with the behavioral data, inhibition of OFC inputs actually increased Glut release in the DLS of putative fast switchers at both trial start ($t(6)=1.59$, $p=0.081$) and prior to a correct choice ($t(6)=2.28$, $p=0.032$; Fig 8d). In putative slow switchers, inhibiting OFC ($t(6)=1.76$, $p=0.064$) and IN Thal ($t(6)=4.84$, $p=0.001$) reduced glutamate release in DLS prior to a correct choice (Fig 8e). The only similarity between putative fast and slow switchers was that inhibition of IN Thal inputs reduced glutamate release in DMS prior to a correct choice in both fast ($t(6)=3.28$, $p=0.008$) and slow switchers ($t(6)=2.23$, $p=0.034$; Fig. S5).

Strategy switching test: We found the greatest difference between eventual fast and slow switchers during the early (trials ~1-30) and middle (trials ~60-90) portions of the test. So, we split the first 90 trials of the strategy switching test into 3 sets (1-30, 31-60, 61-90) and each rat underwent 1 set with a vehicle injection, 1 set with a CNO injection, and 1 set with a SalB injection (order counterbalanced) to compare the effects on Glut release both within and across rats. Behavior was stopped for the day after 30-trial sets with injections of CNO or SalB to allow drug washout before the next set of 30 trials on the following day. In putative fast switchers, inhibition of both IN Thal ($t(5)=3.05$, $p=0.014$) and OFC ($t(4)=2.79$, $p=0.025$) inputs reduced Glut release in the DLS at trial start and prior to a correct choice following the previous strategy (OFC: $t(4)=2.25$, $p=0.044$; IN Thal: $t(5)=3.02$, $p=0.015$; Fig 9a). The effects in DMS were more specific, with OFC inhibition reducing Glut release prior to a correct choice following the previous strategy ($t(4)=6.23$, $p=0.002$) and IN Thal inhibition reducing Glut release at trial start ($t(6)=1.85$, $p=0.057$; Fig. 9b). Interestingly, we also measured several changes in DA

release in DMS and DLS following inhibition of OFC and IN Thal inputs. The most consistent were an increase in DMS DA at trial start ($t(4)=4.20$, $p=0.007$) and a decrease in DMS DA during pellet retrieval ($t(4)=2.95$, $p=0.021$), both following inhibition of OFC inputs to DMS (Fig. 9b). Neither inhibition of OFC or IN Thal inputs had any consistent effect on the Glut measured prior to a correct choice following the new strategy, although fewer rats sampled the new strategy during this portion of the test (Fig S6). There were, however, reductions in DMS DA during pellet retrieval (Fig S6b). Neither inhibition of the OFC or IN Thal inputs had any consistent effect on Glut release at any behavioral timestamp in the putative slow switchers (Fig S7), which was expected given the general lack of Glut release in slow switchers. After the first 90 trials, all rats finished the strategy switching test without further manipulation. Putative fast switchers were impaired in their strategy switching performance relative to the fast switchers from the first experiment ($P=0.0067$), whereas putative slow switchers were improved ($P=0.0009$; Fig 9c, see discussion).

Cocaine Self-administration: Following the strategy switching test, all rats were implanted with jugular vein catheters and trained to self-administer cocaine as before. Putative slow switchers again took fewer days to acquire cocaine self-administration, but this effect did not reach statistical significance ($t(10)=1.24$, $p=0.121$, not shown). After acquisition, all rats were given continuous access to cocaine for 2 hours across 10 days, as this revealed the largest difference between fast and slow switchers in the prior experiment. Putative slow switchers took similar numbers of cocaine infusions as fast switchers across the first 7 days. Slow switchers took more cocaine infusions across the last 3 days, which reached significance on day 10 ($P=0.039$; Fig. 9d). Rats then underwent a first punishment resistance test, followed by 4 further days of self-administration and then a second punishment resistance test. Each rat underwent 1 punishment resistance test following a vehicle injection and one test following an injection of both CNO and SalB (order counterbalanced). Putative slow switchers took more punished cocaine infusions overall ($F(1,9)=8.94$, $p=0.015$), as in the first experiment; however, injection of CNO and SalB increased the number of punished cocaine infusions taken in both putative fast and slow switchers (Fig 9e). CNO injection was given prior to the start of the punishment resistance test, but, because of SalB's short half-life, it was given just prior to the medium shock intensities. Thus, we can observe an effect of inhibiting OFC inputs across the whole test but can only compare the effect of inhibiting both OFC and IN Thal projections during the medium shock intensities. Treatment with CNO and SalB increased the number of punished cocaine infusions across all rats ($F(1,9)=43.64$, $p<0.0001$) and within both putative fast ($p=0.005$) and slow ($p=0.0003$) switchers (Fig 9f).

Discussion:

We showed that rats that were inherently slow at strategy switching were also the most likely to develop punishment resistant cocaine use, similar to what others have shown with poor reversal learning predicting higher cocaine use[4]. Photometry recordings during discrimination and the strategy switching test showed that fast switchers had elevated dorsal striatal Glut release and reduced DA release at trial initiation (DLS) and prior to a correct choice (DMS and DLS). They also featured dynamic fluctuations in release measures to the positive feedback of receiving the pellet reward, particularly an increase in DA release in the DMS. Slow switchers showed the opposite pattern, with elevated DMS and DLS DA release prior to a correct choice and a quiescent response at trial initiation and to positive feedback. Many of these release measures were significant predictors of either strategy switching performance, punishment resistant cocaine use, or both as indicated by our correlation matrices, Chi-Square test, and classification model. Interestingly, quantifying the dynamic relationship of Glut release relative DA release was often a better individual predictor than measuring either DA or Glut individually. Furthermore, combining 2-5 release metrics allowed our classification model to categorize switching speed and punished cocaine use with 100% accuracy. This emphasizes the importance of measuring the dynamic relationship between Glut and DA release across several behavioral measures. In a separate experiment, we showed that rats categorized by our model as fast and slow switchers showed differential recruitment of Glut inputs to the dorsal striatum during discrimination and strategy switching. Putative fast switchers showed significant contributions from both the OFC and the IN Thal in DLS and DMS across several strategy switching timestamps but showed little to no recruitment of these regions during performance of the discrimination paradigm. Putative slow switchers showed little to no contribution from the OFC and IN Thal during the strategy switching test but showed significant contribution of both regions to the DLS during the discrimination paradigm. There was also a differential effect on behavior across these rats that aligned with the reductions in Glut release, with OFC and IN Thal inhibition impairing putative fast switchers on the strategy switching test only and putative slow switchers on the discrimination paradigm only. Interestingly, inhibition of OFC and IN Thal inputs to dorsal striatum increased the number of punished cocaine infusions taken across all rats. This emphasizes our correlation matrix data showing that Glut/DA interactions during both the discrimination paradigm and strategy switching can predict punished cocaine taking but also demonstrates that the source of the Glut across discrimination performance and strategy switching is likely different. Based on these data, we propose a model where the combination of projection-specific Glut release and a DA dip at behavioral initiation (DLS) or at a decision point (DMS and DLS), with a subsequent increase in DA and Glut during the feedback period, represents a neurochemical “signature” of flexibility that predicts a greater ability to adapt to changing contingencies (e.g. a strategy switch or punishment addition, Fig. 10). In contrast, dorsal striatal DA release leading to a

decision point, with a quiescent response at behavioral initiation and during the feedback period, represents an inflexible signature (Fig. 10).

Glut-DA dynamics for flexible behavior: Our data suggest that elevated Glut release, with a concomitant reduction in DA release, at trial initiation or prior to choice predicts both fast strategy switching and punishment sensitivity, whereas elevated DA release at these timepoints predicts the opposite. This suggests that the timing of the interaction of DA and Glut release at these timepoints is critical. We have two suggestions for why this might be. First, an increase in Glut paired with a decrease in DA could prime the striatum to maximally attend to trial-by-trial feedback. The striatum is primarily made up of medium spiny projection neurons (MSN) that have a hyperpolarized resting potential[40]. Prior studies have shown that Glut inputs to striatum can reduce this resting potential, putting them in an “upstate” and making it easier for them to fire[41]. A dip in DA levels can further favor Glut release from these inputs by releasing them from presynaptic D2-mediated inhibition[42,43], which specifically favors subsequent input from corticostriatal projections[44]. Because of this, feedback-related elevations in phasic Glut and DA release can facilitate LTP and learning[45,46]. A quiescent Glut signal with small increases in DA at trial start or prior to a choice would instead result in further inhibition of Glut release via presynaptic D2’s[42,43], leaving MSN’s in a hyperpolarized “downstate”. We see evidence of this notion in our own data. For example, fast switchers, which feature the increase in Glut and decrease in DA at trial start and prior to a correct choice, also feature dynamic fluctuations in Glut and DA release during pellet retrieval, including a significant elevation in DMS DA. Furthermore, slow switchers, which feature elevated DA and no increase in Glut at trial start and prior to a correct choice, also feature no feedback-related activity during pellet retrieval. Lastly, when the Glut inputs from OFC and IN Thal were removed from putative fast switchers, we showed a decrease in Glut release at trial start and prior to choice which coincided with a reduction in feedback-related DMS DA release at pellet retrieval. This occurred without manipulating DA inputs directly, suggesting that it was a result of the Glut inhibition (See Fig 9b, S6b).

A second possibility is that the increase in Glut and decrease in DA could facilitate the indirect pathway, leading to a greater inhibition of the previously learned strategy. MSN’s are categorized by the presence of D1 or D2 receptors. D1 receptor-containing MSN’s facilitate movement and action selection via activation of the direct pathway and D2 receptor-containing MSN’s inhibit movement and refine action selection via activation of the indirect pathway[47,48]. A dip in DA release has been shown to facilitate LTP in D2 MSN’s and LTD in D1 MSN’s[49], thus favoring the indirect pathway and motor inhibition. Furthermore, thalamic inputs to the dorsal striatum synapse onto both MSNs and cholinergic interneurons (CIN’s)[50]. Thalamic input-mediated excitation of CIN’s leads to a burst-pause firing pattern, which mimics the physiological response of CINs following

presentation of salient stimuli[50]. CIN activation favors corticostriatal signaling, ultimately leading to activation of the indirect over the direct pathway[50]. This has been suggested by others as a mechanism for salient information to lead to action suppression to facilitate shifts in attention and behavior[50,51]. Again, there is evidence of this notion in our data. We showed that the Glut we measured at trial start and prior to a correct choice following the previous strategy was from cortex (OFC) and thalamus (IN Thal), both of which would favor activation of the indirect pathway. Importantly, these inputs were not responsible for the Glut measured when fast switchers were making a correct choice following the new strategy (Fig S6). Thus, the increase in thalamocortical Glut inputs with a DA dip in the fast switchers could lead to activation of CINs and D2 MSNs, potentially improving the fast switcher's ability to inhibit responding via the previous, well-learned strategy despite getting that trial correct. The slow switchers, which did not feature Glut increases from thalamus or cortex during the strategy switching test (Fig S7), had increased DA release prior to choice. This would be expected to activate D1 MSN's leading to direct pathway activation[47,48], inhibition of corticostriatal input[44], and perseveration on the previous strategy[50].

Differential effect of OFC and IN Thal inhibition on Glut release in putative fast and slow switchers: We inhibited OFC and IN Thal inputs to dorsal striatum and measured how these manipulations affected Glut release across several key behavioral timestamps during the discrimination paradigm and strategy switching. We chose to inhibit OFC and IN Thal because our data suggested that the Glut we measured correlated highly with strategy switching performance and these Glut inputs have been implicated in cognitive flexibility or some version of task switching. The OFC sends dense and sparse projections across the dorsal striatum[52] and the ventral and lateral subregions have been consistently implicated in flexibly adapting behavior in uncertain situations or when contingencies change (reviewed in[34]). For example, OFC has been shown to track changes in action value and be necessary for shifts in behavioral strategies[33] and probabilistic reversal learning[32]. The OFC also supports state representations as a means to promote new learning when outcomes diverge from expectations[53], as would happen when contingencies change during strategy switching. The intralaminar nuclei of the thalamus (IN Thal) is the primary source of excitatory inputs to the striatum[54], are necessary for initiation[15] and execution[16] of well-learned behaviors, and have been implicated in cognitive flexibility[31,37-39]. In particular, both the anterior (centromedian, centrolateral) and the posterior (parafascicular) subregions send parallel projections to DMS and DLS[55,56] and are necessary for redirecting attention[35,36] and integrating new and old information[31] in response to contingency shifts.

Inhibition of OFC inputs resulted in diminished Glut release in both the DMS and DLS prior to a correct choice following the previous strategy, as well as at trial start in the

DLS. Inhibition of the IN Thal resulted in diminished Glut release at trial start in both the DMS and DLS, as well as prior to a correct choice following the previous strategy in the DLS. However, this was only the case in putative fast switchers. Inhibition of the OFC and IN Thal had little to no effect on Glut release during the strategy switching test in putative slow switchers. Slow switchers showed significantly less Glut released during the strategy switching test timestamps in the first experiment, suggesting that Glut levels may not have been able to be reduced due to a floor effect. That said, rats in the putative fast switcher group had a range of Glut release magnitudes, including some that were quite low, which were reduced by chemogenetic inhibition of the OFC and/or IN Thal inputs. Furthermore, in putative slow switchers, inhibition of OFC and IN Thal inputs reduced DLS Glut during the discrimination paradigm, and this did not occur in the putative fast switchers. In fact, inhibition of OFC inputs actually increased Glut release in the DLS at trial start and prior to a correct choice in fast switchers. This suggests that there is a fundamental difference between the dorsal striatum input architecture of fast and slow switchers, with opposite activation patterns of OFC and IN Thal inputs during discrimination and strategy switching in these two rat phenotypes. Based on our correlational analyses and classification modeling, this difference in input architecture and utilization may be at the heart of the difference between “flexible” and “inflexible” rats.

Importantly, inhibition of OFC and IN Thal projections did not reduce Glut at several key points in both fast and slow switchers. Notably, inhibition of OFC and IN Thal had no effect on the Glut measured prior to a correct choice following the new strategy in either group of rats. Glut release measured during this behavioral timestamp could come from three possible sources. First is the motor cortex. The motor cortex sends dense glutamatergic projections to the DLS[52,57], is necessary for learning complex movement patterns, but is not necessary for executing the behaviors once they are well-learned[16,57]. Thus, the new learning required to acquire the new strategy may involve Glut from the motor cortex. A second possibility is the prelimbic prefrontal cortex. The prelimbic cortex is necessary for strategy switching[58-60] and tracks values changes of different actions to facilitate flexible decision making[61]. However, the prelimbic only projects to the DMS[52] and likely could not explain release in the DLS. A third possibility is the basolateral amygdala (BLA). Glutamate inputs from the BLA[52,62], although sparser than those from cortex or thalamus[52], have been implicated in the shifting of behavior from outcome-insensitive to goal-directed[12]. Furthermore, the amygdala shares reciprocal connectivity with the OFC, and this projection has also been implicated in adaptive decision making[63]. Thus, the amygdala could regulate dorsal striatal-mediated flexibility either directly or via the OFC. Further study will be needed to fully understand the dynamic nature of the Glut inputs to the dorsal striatum at play during discrimination performance and the strategy switching test.

Differential effect of OFC and IN Thal inhibition on behavior in putative fast and slow switchers: A curious finding was that rats predicted by our model to be slow switchers were better at the strategy switching test than the slow switchers from our first experiment. Is it possible that our model was incorrect in its categorization of these rats or could there be another explanation for this counterintuitive phenomenon? It is possible that the model miscategorized a few of the rats. However, we used 3 different 3-metric models, which perfectly categorized the rats from the first experiment, to categorize these rats to minimize this risk. Furthermore, 10 of the 14 rats were categorized as fast or slow by all 3 models, suggesting a high level of confidence in the classifier of at least those 10. Furthermore, the mean tracings for the putative fast and slow switchers look remarkably similar to the respective fast and slow switchers from the first experiment. The putative fast and slow switchers also demonstrated interesting differences in the contributions of OFC and IN Thal inputs during the discrimination paradigm and strategy switching test, with putative fast switchers showing significant contributions of these regions during strategy switching but not discrimination, and putative slow switchers showing the opposite. These effects were consistent across rats within each group, suggesting a difference in input architecture and utilization that aligns with their behavior. Thus, although it is still possible, we do not believe that the putative fast and slow switchers were miscategorized. Instead, we believe the improved strategy switching in the putative slow switchers may be because their performance in the discrimination paradigm was destabilized by the chemogenetic inhibition. We found that inhibition of both OFC and IN Thal inputs during the discrimination paradigm reduced Glut release in the DLS of putative slow switchers. This led to a dip in performance from day 6 (91%), such they got a similar percentage of trials correct on day 7 (84%) as they did on days 1 and 2 (82%). Given that activity in the DLS increases across learning[19], at the initiation of a well-learned behavior[17], and across performance of the discrimination paradigm (Fig 6), and that DLS inhibition increases rats' sensitivity to changes in action-outcome contingency[64], it may be that the reduction in DLS Glut could have reset how "well-learned" the visual strategy was, increasing the putative slow switchers' sensitivity to the strategy switch. Importantly, the putative fast switchers did not show a reduction in DLS Glut at any timestamp during the discrimination paradigm, nor any dip in performance on any discrimination day.

Dual effect of DMS DA release: DA release in the striatum is necessary for both learning and cognitive flexibility by strengthening the association of actions with rewarding outcomes[65], which is separate from movement-related signaling[66]. We see evidence of this as well, as the rats that showed the greatest elevations in DMS DA release during pellet retrieval switched strategies the fastest (Table 1, Fig 7), and the reductions in strategy switching performance after OFC or IN Thal inhibition were accompanied by significant reductions in DMS DA release during pellet retrieval (Fig 9b, S6b). The more paradoxical finding may be that DMS DA released at trial start or prior

to choice predicted both slow strategy switching and punishment resistance. This finding may seem to be at odds with prominent theories in the literature; however, two recent studies lend credence to the idea that DMS DA release can predict inflexibility[67] and resistance to punishment[68]. For example, van Elzelingen et al (2022) showed that the combination of elevated DMS DA release to a distal seeking cue (seeking lever extension) and decreased DMS DA levels at pellet delivery significantly predicted which rats would show habitual behavior during a habit test[67]. The opposite pattern predicted which rats would be goal-directed, and these patterns increased in magnitude across training[67]. Furthermore, optogenetic stimulation of DMS DA release during the distal cue seeking accelerated habit formation[67]. Moreover, Seiler et al (2023) found that DMS DA release during a rewarded nosepoke (RI30 schedule) early in training predicted which rats would be punishment resistant on a first probe test[68], and that optogenetic stimulation of DMS DA release accelerated the development of punishment resistant sucrose taking[68]. Interestingly, all mice showed an increase in DMS DA release by the end of training, indicating that the timing of the DMS DA release, i.e. when in training it developed, was critical for whether a mouse would be resistant or sensitive to punishment[68]. Thus, in both cases and similar to our results, the timing of the DMS DA release predicts flexibility. When it occurs early in training[68], early in a behavioral progression[67], or prior to a choice (Fig 5b), it is indicative of an inflexible, punishment resistant phenotype (Fig 7). When it develops late in training[68] and occurs to provide feedback about an action[65-67] (Fig. 5d), it is indicative of sound learning and a flexible, goal-directed phenotype (Fig 7).

Are the DMS and DLS cooperative or competitive? Many studies over the years have highlighted a competitive balance between the DLS and DMS in habitual versus goal-directed behaviors, respectively. However, other studies have highlighted the possibility of more cooperative actions of these two dorsal striatum regions in action selection (reviewed in[30]). Interestingly, we find evidence of both across discrimination performance and strategy switching, and we posit that the difference may lie in the degree of uncertainty in the rat's actions. During the discrimination paradigm in fast switchers, we see evidence of a competitive nature between the DLS and DMS, with DLS Glut release being significantly higher than DMS Glut at trial start and correct choice across days (main effect of region). Because the rats learn to nose-poke using a visual cue throughout pre-training, they often get the majority of their trials correct, even on the first day of the discrimination paradigm. As such, observing DLS Glut primacy fits what we see in the literature, that DLS activity is necessary for initiating or executing a well-learned behavior. DLS Glut primacy is also seen late in the strategy switching test. The late test photometry measurements are recorded as the rat reaches the performance criterion of 10 consecutive correct, i.e. again when they are getting few if any trials incorrect. The alternative to this is seen during the early portion of the strategy switching test, where DMS Glut reaches its peak. Interestingly, however, we don't see a

reversal of the pattern from above, where DMS Glut increases and DLS Glut decreases. Instead, we see a situation where both DMS and DLS Glut release increase, particularly just prior to making a correct choice following the previous strategy. This suggests that when uncertainty is high early in the strategy switching test, where the rat is likely perceiving conflicting information about whether the previous strategy is still rewarded or not, that the DMS and DLS are working cooperatively instead of competitively.

The final experiment may also help in disentangling these views to some extent. We showed that OFC inhibition reduced DMS and DLS Glut prior to a correct choice using the previous strategy but only reduced DLS Glut at trial start. Alternatively, IN Thal inhibition reduced both DLS and DMS Glut at trial start, but only DLS Glut prior to a correct choice using the previous strategy. Therefore, one possibility is that DLS and DMS Glut inputs from OFC may be working cooperatively at choice points with high uncertainty, but competitively at behavioral initiation with low uncertainty. Likewise, DLS and DMS Glut inputs from IN Thal may be working cooperatively at behavioral initiation when uncertainty is high and competitively at choice points when uncertainty is low.

Potential cause of natural variation in cognitive flexibility and related release measures: The rats in our study were unmanipulated, outside of uniformly being infused with viruses and fibers for photometry. Performance in the strategy switching test followed a normal distribution, such that slow vs fast switching was self-selective. Interestingly, once rats were split into groups, differences in dorsal striatal Glut release were present from the very first day of the discrimination paradigm. Specifically, the rats that went on to be fast switchers had elevated Glut release in the DLS at trial initiation and pellet retrieval, the former of which predicted switching performance. In fact, Glut release at trial start on day 5, which occurred days before the strategy switching test and weeks before the punishment resistance test, was the best individual predictor of both (>70%, Fig 7). As such, we hypothesize that the differences we observed across these rats represent a trait feature in the way these rats learn and adapt to changing environmental contingencies. The interesting question here is, why are there differences in rats' striatal signaling? While there certainly may be a genetic component given that this is an outbred strain of rats, unknown environmental factors may also play a role. For example, dorsal striatal signaling may be programmed during development through maternal care or other factors[69-71]. Stress is also known to impair measures of cognitive flexibility[72,73], and it is possible the rats had differential responses to or experiences with stress. Stress also increases dopamine activity[74] and release[75] in rodents and humans[76], and impairs excitatory transmission in the dorsal striatum[73]. Interestingly, however, while stress increases dorsal striatal DA release, human studies have also noted a reduction in error signaling, specifically signaling in response to positive reward prediction error[77]. We similarly observed that slow switchers had elevated DA responses prior to choice, overall reduced Glut release, and blunted responses to positive

feedback. Importantly, all of these effects significantly predicted strategy switching performance. It is important to note that there were no known differences in stressful experiences in our rats; however, anecdotally, we noticed that fast switchers and slow switchers almost universally came from the same cages (all rats pair-housed). Thus, it is possible that some rats had coincidental stress experiences during transportation to our facility, during a cage change, or due to a wet box or aggressive cagemate. Although speculative, this is one possibility worth further investigation.

Translational relevance: Poor cognitive flexibility predicts the onset and severity of SUD-relevant phenotypes in rodents and primates[2-5]. There is also some evidence of this in humans[7]. Our data suggest that one common link between these could be inherent differences in dynamic fluctuations in DA and Glut release in the dorsal striatum in situations where changing contingencies require behavioral adaptation. In fact, the presence or absence of the above described "flexibility signature" perfectly predicted whether a rat would be a fast or slow switcher. It also perfectly predicted whether a rat would go on to display a prominent SUD-like phenotype in punished cocaine use (Fig 7), despite this punishment resistance test being performed nearly a month later. Thus, monitoring striatal dynamics of *both* Glut and DA signaling across multiple adaptive choice behaviors is more informative than measuring either in isolation, and has several implications for translation. First, the neural signature of flexibility we demonstrated provides a metric that could be measured in high-risk individuals to predict the likelihood of them developing a SUD. If detected, these individuals could represent a population that might benefit from preventive therapy or treatments. Second, the difference between the flexibility and inflexibility signature provides targets for treatments such as transcranial stimulation. Specifically, we show that Glut inputs from OFC and IN Thal are responsible for the Glut we measured at different behavioral timestamps, are differentially recruited during discrimination and strategy switching across flexible and inflexible rats and are necessary for strategy switching and sensitivity to punishment. Thus, these projections could represent treatment targets that may need to be tuned based on each individuals' trait flexibility.

Summary: Across learning and strategy switching, elevations in Glut, with reductions in DA, at trial initiation (DLS) and prior to choice (DMS and DLS), with dynamic fluctuations in DMS DA and DLS Glut to positive reinforcement, represented a "signature of flexibility" that predicted both fast strategy switching and cocaine seeking that was sensitive to punishment. Elevations in DLS and DMS DA at these respective points, with a quiescent response to feedback, represented a "signature of inflexibility" that predicted slow switching and resistance to punishment. The differences in neurochemical signatures were accompanied by differences in input architecture and utilization across discrimination and strategy switching and represent inherent differences between a group of unmanipulated male and female rats. As such, these

data describe a pattern that could be measured and/or manipulated to predict and treat SUDs in humans.

Methods:

Subjects: Adult (>PND70) male and female Sprague-Dawley rats (Envigo, Frederick, MD) were used for all experiments. Confidence intervals were used to compare between sexes (see results). Both sexes were determined to be part of the same interval, as we have shown previously[78,79], and were combined. We also did not observe sex differences in cocaine-motivated behaviors under the experimental conditions we used, similar to our prior work[28]. All rats were pair-housed unless otherwise noted and were maintained on a 12-12hr light-dark cycle, with unlimited access to food chow and water, except during behavioral testing, as described below. Experiments were performed during the dark cycle. Experimental procedures were approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh according to National Institute of Health Guide for the Care and Use of Laboratory Animals.

Virus infusion surgery: All viral constructs were purchased from Addgene. Prior to any behavioral training, Rats underwent an initial virus infusion surgery using our standard protocols[20,80-82] where they received infusions of viral vectors that contain photo-emitting biosensors to detect DA (red- AAV_9-hsyn-GRAB_rDA1m, 140556-AAV9; 100 µL at titer ≥ 1×10^{13} vg/mL, deposited by Yulong Li)[83] and Glut (green- AAV_9-hSynapsin-SF-igluSnFR.A184S, 106174-AAV9; 100 µL at titer ≥ 1×10^{13} vg/mL, deposited by Loren Looger)[84] into the DMS of one hemisphere (AP: -0.2, ML: ±2.2, DV: -4.5) and the DLS of the other hemisphere (AP: +0.8, ML: ±2.8, DV: -5.0). In the same surgery, optic fibers (200µm diameter 0.37 NA, Doric) were implanted into the DMS and DLS (one fiber/ region, Fig. 1a).

For the final experiment only, rats were also infused with AAV_8-hSyn-dF-HA-KORD-IRES-mCitrine (65417-AAV8; 100 µL at titer ≥ 7×10^{12} vg/mL, deposited by Bryan Roth)[85] into the Intralaminar thalamus (IN thal, AP: -4.1, ML: ±1.2, DV: -6.2; 0.5 µL/ hemisphere) and AAV_8-hSyn-DIO-hM4D(Gi)-mCherry (44362-AAV8; 100 µL at titer ≥ 1×10^{13} vg/mL, deposited by Bryan Roth)[86] into the ventral orbitofrontal cortex (OFC, AP: +4.2, ML: ±1.3, DV: -4.8; 1.0 µL/ hemisphere). GRABDA, iGluSnFr, and Cre (pENN.AAV_{rg}-hSyn-HI-eGFP-Cre.WPRE.SV40, 105540-AAVrg; 100 µL at titer ≥ 7×10^{12} vg/mL, deposited by James M Wilson) were combined and infused into the DMS of one hemisphere and the DLS of the other hemisphere (same coordinates as above). After viral infusion, optical fibers (Doric) were implanted over the DMS and DLS at the same location and secured with skull screws and cement, as above. This viral strategy allowed us to individually or jointly inhibit projections from the OFC (CNO, 3 mg/kg i.p.,

Hello Bio) and/or the IN Thal (salvinorin B, 15 mg/kg s.c., Hello Bio) to the DMS or DLS and perform photometry recordings from these same regions.

Discrimination paradigm and strategy switching: Following the virus infusion surgery, rats first learned to generally nose-poke for a reward. They were trained to nose poke in either of two lit, active ports in an operant chamber to receive a sugar pellet reward (nose-poke pre-training, Fig S1a). As such, they all learned to use a simple visual cue to identify an active port. Once rats earned 50 sugar pellets in a 30-minute period, they went on to a second pre-training phase. For this phase, rats had to initiate each trial with a nose-poke on a central port, prior to nose-poking either of the two lit, active ports (Fig S1b). Once rats earned 50 sugar pellets in a 30-minute period in this second pre-training phase, the pre-training phases were complete. Rats that went on to be fast and slow switchers took a similar number of days to complete the two pre-training phases (Fig S1c). A side bias rating was taken on the final pre-training day, which was the number of nose-pokes on the preferred side divided by the total number of nose-pokes. Most rats had a side bias, but there were no differences between rats that went on to be fast or slow switchers (Fig S1d). Ten rats showed a right-side bias and 12 rats showed a left side bias. Rats then performed a complex discrimination paradigm where they continued to discriminate between the 2 active ports using a visual strategy (Fig. 1b). However, now they had to ignore the position of the visual cue, as the light cue appeared in either the left or right port, pseudorandomly. A trial began with a central port being illuminated. The rat had to nose-poke into this port to initiate the trial, similar to the second pre-training phase. Once the trial was initiated, rats were to respond into the port with the light regardless of where it was located. If they chose correctly, they again received a sugar pellet reward, if they chose incorrectly, they did not receive a pellet but there were no other programmed consequences. There was a brief inter-trial interval (4 seconds) and then the central port illuminated, indicating that a new trial was available. Rats did this for 100 trials/ day for 10 days. On the 11th day the rats performed a strategy switching test. The test began with 20 “reminder” trials (pre-switch) that rewarded the same visual strategy that had been rewarded for the previous 10 days (Fig. 1c). On trial 21, the rewarded strategy switched to a response strategy (Fig. 1d). Thus, the rats had to learn to ignore the visual cue and respond either to the left or to the right port (response strategy). The rewarded port side was opposite each rat’s natural side bias as determined at the end of the pretraining phase. We quantified the number of trials and errors it took the rats to reach the performance criterion of 10 consecutive trials correct following the strategy switch (Fig. 1e).

Photometry: DMS and DLS DA and Glut recordings were performed as we have reported in our prior studies[80]. For the first experiment, recordings occurred on discrimination days 1, 5, and 10, and during the strategy switching test. Data from the strategy switching test were divided into the first ~30 trials after the strategy switch

(early strategy switch), trials ~60-90 (middle strategy switch), and the final ~20 trials as rats reached the performance criterion (late strategy switch). For the second experiment, recordings occurred on discrimination training day 1,4,5, and 6. For the strategy switching test, we recorded the first 90 trials, split into 3 30-trial sets. During recording sessions rats were connected to a patch cable (200 µm diameter, Doric) attached to the fiber photometer (Plexon). The igluSnFR sensor (green) was excited at 465 nm with the minimum power that ensures a robust signal to noise ratio (80-150 µW) and that reduces photobleaching to prolong the duration of recording sessions (up to 45 min). The emission fluorescence of the sensor (525 nm) was acquired at a frequency of 30 Hz to allow sub-second time resolution using Plexon software. The GRABDA sensor (red) was excited with light at 560 nm and the emission (600 nm) was acquired with the same frequency. Emissions were collected simultaneously with a high quantum-efficiency CMOS camera. Glut and DA fluctuations were measured as the deflection in fluorescence relative to baseline levels (GuPPy[87]). Frames were interleaved with excitation at the isosbestic point (410 nm), which was then subtracted to control for movement artifacts. Relevant behaviors were time stamped to align to photometry signals. For the discrimination days, we time-stamped photometry measures to trial initiation, correct and incorrect choices. Pellet retrieval and no pellet were quantified relative to a correct and incorrect choice, respectively. For the strategy switching test we time-stamped the photometry signal to trial initiation, correct choices following the previous strategy, correct choices following the new strategy, and incorrect choices. Pellet retrieval and no pellet were quantified relative to a correct and incorrect choice. The photometry signal for each behavior was analyzed in different windows. For trial initiation, correct, and incorrect choices, we analyzed one second prior to the timestamp, for no pellet we analyzed the 1 second after choice, and for pellet retrieval we analyzed the 2 seconds while the rat is consuming the pellet.

Cocaine self-administration (SA): At least two days after the strategy switching test concluded, all rats were implanted with an indwelling intravenous catheter as we have previously published[20,80,82]. Following a 1-week recovery from catheter implant surgery, rats were trained to self-administer cocaine (0.5 mg/kg) paired with a tone/light cue such that each lever press on the active lever delivered both the cue and cocaine (FR1 schedule, Fig. 1f). The lever presses on the inactive lever had no programmed consequences. The sessions lasted 125 minutes. Once the rats took at least 15 infusions of cocaine with twice as many active as inactive lever presses for 2 consecutive days, the training criterion was reached, and rats began 10 days of cocaine self-administration (Fig. 1h). During these 10 days, rats either had access to cocaine continuously on an FR1 schedule for the entire 125-minute session (continuous access, first and second experiment), or they had access to cocaine in 5-minute bins every 30 minutes (intermittent access, first experiment only). Intermittent access schedules have been shown to precipitate a more punishment resistant phenotype[20,23,24]; thus, this

study design allowed us to examine how strategy switching performance interacts with cocaine access schedule. On the 11th day, all rats were given continuous access to cocaine for the entire 125-minute session (continuous access test) to more directly compare cocaine intake between the two access schedules. On day 12, all rats performed a punishment resistance test (Fig. 1k), where cocaine infusions were first earned on an FR1 schedule for 20 minutes with no shock, followed by 90 minutes with shock. During the shock phase, cocaine infusions were paired with a foot shock that increased in intensity every 10 minutes on a tenth-$\log_{10}$ scale from 0.13 to 0.79 milliamps[20-22]. Cocaine infusions across the test were binned into four groups, no shock (20 minutes), low shock (0.13-0.2 mA, 30 minutes), medium shock (0.25-0.4 mA, 30 minutes), and high shock (0.5-0.79 mA, 30 minutes).

Histology: Rats were euthanized, perfused transcardially with 4% paraformaldahyde, and decapitated. Brains slices were harvested and mounted for immunofluorescent confirmation of viral placement and spread and photometry fiber placement. An anti-mCherry-Rabbit monoclonal recombinant IgG (Sysy, Catalog no. 409008, 1:500) and a goat anti-rabbit secondary (abcam150077, 1:500) were used for confirmation of the GRABDA. An anti-GFP-chicken monoclonal recombinant IgG (abcam 13970, 1:2000) and a donkey anti-chicken secondary (Jackson Immuno Research Labs; Catalog no. NC0456003 via Fisher, 1:5000) were used for confirmation of the igluSnFR.

Data Analysis: To examine how photometry signaling in the DMS and DLS relates to strategy switching performance and subsequent cocaine intake and punishment resistance, we analyzed our data both as a continuum and using a median split to parse groups into “fast” and “slow” switchers. Behavioral data (e.g., trials, errors, cocaine infusions, etc.) was analyzed by multi-factor ANOVA with repeated measures as appropriate. When analyzing the data categorically, we used strategy switching errors to perform a median split into “fast” and “slow” strategy switchers. Significant interactions were followed by Tukey’s post-hoc test with $p<0.05$ considered significant. Fiber photometry data were processed in real-time to collect z-scores and a heat map. Behavioral data were collected by MedPC and TTL signals were sent to Plexon to ensure data were within the identical “time-space.” We compared areas under the curve (AUC) across a 1-2 second time window for each behavioral time stamp using the same analysis strategy as above. Because we measured DMS and DLS DA and Glut in every rat, we also analyzed how these signals affected each other within and between rats using repeated measures ANOVA, multiple linear regression models, and correlation matrices. Note that p values in correlation matrices are not corrected for multiple comparisons. We performed our first experiment with 23 rats. One rat was removed from all analyses due to a dislodged headcap. Five rats were removed from the cocaine analysis due to catheter patency issues, and one was removed for errors made administering the punishment sensitivity/ resistance test. The second experiment was

performed with 15 rats. One rat was removed from all analyses due to injection errors. Two rats were removed from the cocaine analysis because they did not meet the self-administration acquisition criteria, and one was removed because the shock floor malfunctioned during the punishment sensitivity/ resistance test.

Linear Discriminant Analysis (LDA) was used to predict two behavioral metrics: switching speed and punishment sensitivity, based on neurotransmitter activity metrics, which served as predictor variables. To identify the most effective combinations of predictors, separate binary classifiers were trained for each behavioral metric using all possible combinations of 1 to 14 predictor variables. 8-fold cross-validation was used to measure the classification accuracy of each model. Models with an average classification accuracy of 100% were then examined to identify the specific predictor variables that were used. The predictive strength of each individual variable was further assessed by ranking them according to the negative logarithm of the p-value obtained from a Chi-square test. This test evaluated each variable's ability to discriminate between the two categories of each behavioral metric.

Strategy switching categorization for experiment 2: We aimed to measure the effects of inhibiting OFC and IN Thal inputs to the dorsal striatum on Glut and DA release and behavior. However, based on the results from our first experiment, we hypothesized that inhibiting these projections may result in different effects in rats that would be fast or slow strategy switchers. However, because we manipulated projections during the strategy switching test, we could not categorize them based on their actual performance. Our linear discriminant analysis model identified hundreds of combinations of the neurotransmitter metrics that yielded perfect classification of switching speed. Thus, we selected 3 3-metric models that used metrics from early in the discrimination paradigm (day 1 and 5) and early in the strategy switching test, i.e. when the rats were unmanipulated (vehicle condition). Each model categorized rats as fast if they featured 2 out of 3 of the following: Glut AUC measures >2.5, DA AUC measures <0, and Glut/DA diff AUC measures > 2.5. They were categorized as slow if they were the opposite of that for 2 out of 3 measures. For rats where the different models disagreed, the final categorization was based on the majority opinion, 2 out of 3 models (see Fig S4).

Figure legends:

Figure 1 Experiment timeline, histology, and behavioral results: a) Male (N=10) and female (N=12) rats were infused with viral vectors that contain photo-emitting biosensors to detect dopamine (DA, yellow) and glutamate (Glut, green) into the dorsomedial striatum (DMS) of one hemisphere and the dorsolateral striatum (DLS) of

the other hemisphere, within the same rat. Representative images with closeup insets (white rectangle) of the recording zones in the DMS (left) and DLS (right) show yellow (DA) and green (Glut) immunofluorescence around the fiber track. **b)** All rats performed a visual discrimination strategy for 100 trials per day across 10 days. **c)** Photometry recordings were performed on days 1, 5, and 10 (red circles). Post hoc analyses after splitting the rats into two groups revealed that although fast and slow switchers both averaged 87% correct on day 10, they displayed significantly different learning curves. Fast switchers displayed a typical learning curve such that their performance peaked on day 6 (91% correct) and remained steady throughout the remaining days. Slow switchers peaked on day 2 (92%) but then declined after that and remained steady for the last few days. Slow switchers got more trials correct on days 2 (P= 0.04) and 3 (P=0.04). Fast switchers got more trials correct during the 20 “pre switch” trials (89%) compared to slow switchers (81%; P=0.03). **d)** On day 11, all rats performed a strategy switch, such that the visual strategy was no longer rewarded and a response strategy was rewarded. **e**) All rats (gray bar) performed the strategy switch, but there was a natural distribution of performance. Male (squares) and female (circles) rats were split into two groups using a median split of errors. Fast switchers (orange bar) took fewer trials (114.6±5.9) and committed fewer errors (41.6±2.1), compared to slow switchers (purple bar, trials: 220.4±14.0, errors: 81.4±5.0; both P’s < 0.0001). **f)** Following the conclusion of the strategy switching test, all rats were implanted with jugular vein catheters and trained to press a lever (FR1) for an infusion of cocaine (0.5 mg/kg). Infusions of cocaine were paired with a light/tone cue. **g)** Slow switchers (2.9±0.5 days) reached self-administration criterion in fewer days than fast switchers (3.7±0.4 days; t(18)=1.34, P=0.099). **h)** Once rats were reliably self-administering cocaine, rats were given access to cocaine continuously for a 125-minute session (continuous access), or intermittently (5-minutes on, 25 minutes off) for a 125-minute session (intermittent access) across 10 days. **i)** Continuous access rats took more cocaine infusions per day (fast switchers: 44.9±2.7, slow switchers: 51.6±*2.*3) relative to intermittent access rats (fast switchers: 23.1±1.2, slow switchers: 20.1±1.0; P<0.0001) due to their continuous access, but intermittent access rats took more cocaine per minute (fast switchers: 0.9±0.05, slow switchers: 0.8±0.04) relative to continuous access rats (fast switchers: 0.4±0.02, slow switchers: 0.4±0.02; P<0.0001). Similar to previous studies, fast switchers with continuous access took fewer cocaine infusions per day than slow switchers (P=0.02), but fast switchers with intermittent access took more cocaine infusions/ minute (P=0.02). **j)** On the 11th day all rats were given continuous access to cocaine (All Cont) and they took comparable numbers of cocaine infusions (mean range 50 to 58 infusions across all groups). **k)** On day 12, all rats performed a punishment resistance test, where cocaine infusions were paired with a foot shock that increased in intensity every 10 minutes on a tenth-$\log_{10}$ scale from 0.13 to 0.79 milliamps. **l)** Slow switchers took more punished cocaine infusions overall (no shock: 12.0±1.4, low shock:

12.9±1.2, medium shock: 11.1±0.8, High shock: 3.1±0.6) compared to fast switchers (no shock: 10.8±0.9, low shock: 10.9±0.5, medium shock: 6.7±0.9, High shock: 2.8±0.7). This effect was most apparent at the medium shock intensity where intermittent access-slow (10.8±1.4, P=0.046) and continuous access slow (11.7±0.7, P=0.0006) both took significantly more punished cocaine infusions than the continuous access-fast rats (5.3±1.1). Intermittent access produced an intermediate phenotype in the fast switchers (8.4±1.0).

Figure 2 Fast switchers show elevated DLS Glut at trial start: This figure depicts mean tracings aligned to trial initiation (Trial Start) of DLS Glut (red), DLS DA (blue), DMS Glut (green), and DMS DA (yellow) release across fast (left, N=11) and slow (right, N=11) switchers. Area under the curve (Mean ± SEM) for 1 second prior to trial start (as indicated by the green box) for each trace was quantified and compared across fast (orange) and slow (purple) switchers, as well as the difference between Glut and DA release in the DLS and DMS. **a,b**) On day 1 of the discrimination paradigm, rats that went on to be fast switchers showed elevated DLS Glut (4.3 ± 1.1) as they initiated the trial, compared to slow switchers (1.0 ± 0.5). There were no differences in DLS DA (fast: 0.1 ± 0.4, slow: 1.0 ± 0.5), DMS Glut (fast: 0.9 ± 1.0, slow: 0.5 ± 0.8) or DMS DA (fast: 0.1 ± 0.7, slow: 1.1 ± 0.7) on day 1. **c,d**) This effect was also present on day 5 (DLS Glut- fast: 3.2 ± 0.6, slow: -1.1 ± 1.6). While there were no differences between groups in DLS DA (fast: -1.1 ± 0.8, slow: -1.4 ± 1.1), fast switchers demonstrated a significant elevation in DLS Glut relative to DLS DA (fast: 4.2 ± 1.2, slow: -0.04 ± 0.6). There were no differences in DMS Glut (fast: 1.7 ± 1.4, slow: -0.4 ± 1.1) or DMS DA (fast: 0.2 ± 0.8, slow: -1.0 ± 1.0). **e,f)** By day 10 all measures normalized across rats (DLS Glut-fast: 4.3 ± 1.4, slow: 3.2 ± 1.1; DLS DA -fast: -0.5 ± 0.9, slow: -1.4 ± 1.2; DMS Glut-fast: 1.5 ± 0.8, slow: 3.4 ± 1.6; DMS DA-fast: -1.4 ± 0.8, slow: 0.8 ± 2.1).

Figure 3 Fast switchers show elevated DLS Glut at correct choice and pellet retrieval: This figure depicts mean tracings aligned to correct choices (black box) and subsequent pellet retrieval (red box) of DLS Glut (red), DLS DA (blue), DMS Glut (green), and DMS DA (yellow) release across fast (left, N=11) and slow (right, N=11) switchers. Area under the curve (Mean ± SEM) was again quantified for 1 or 2 seconds as indicated by the colored boxes. **a,b)** Fast (orange) and slow (purple) switchers showed no differences in Glut or DA prior to a correct choice on day 1 (DLS Glut- fast: 4.0 ± 1.0, slow: 2.7 ± 0.8; DLS DA- fast: 0.9 ± 0.6, slow: 1.8 ± 1.4; DMS Glut- fast: 1.8 ± 1.4, slow: 2.3 ± 0.8; DMS DA- fast: 0.9 ± 0.9, slow: 2.0 ± 1.2). Fast switchers had elevated DLS Glut release during pellet retrieval (8.1 ± 5.4) compared to slow switchers (-5.8 ± 2.9), but no differences across the other 3 release measures (DLS DA-fast: -1.0 ± 1.2, slow: -11.9 ± 5.7; DMS Glut- fast: -11.9 ± 5.7, slow: -8.9 ± 5.2; DMS DA- fast: -0.9 ± 4.3, slow: -1.9 ± 3.9). **c,d)** On day 5, there was a significant difference in DLS Glut prior to a correct choice with an elevation in fast switchers (fast: 4.0 ± 1.0) compared to

slow switchers (slow: 0.8 ± 1.1); however, there were no differences in the other release measures (DLS DA - fast: 1.0 ± 0.8, slow: -.2 ± 0.8; DMS Glut - fast: 2.2 ± 0.8, slow: 0.3 ± 0.8; DMS DA - fast: 0.2 ± 0.8, slow: 0.8 ± 1.1). There were also no significant differences during pellet retrieval (DLS Glut- fast: 2.0 ± 3.8, slow: -4.0 ± 3.5; DLS DA- fast: 2.9 ± 4.3, slow: 2.5 ± 3.7; DMS Glut- fast: -13.7 ± 6.7, slow: -23.4 ± 8.0; DMS DA- fast: -1.1 ± 5.1, slow: -9.4 ± 4.1). **e,f)** On the final training day there were no significant differences between DA or Glut release within fast and slow switchers for correct lever pressing (DLS Glut - fast: 5.8 ± 1.6, slow: 3.6 ± 1.0; DLS DA - fast: -0.1 ± 0.5, slow: 1.1 ± 0.8; DMS Glut - fast: 2.1 ± 0.5, slow: 2.2 ± 1.4; DMS DA - fast: -0.4 ± 1.0, slow: 2.6 ± 1.6) or pellet retrieval (DLS Glut - fast: 1.7 ± 6.4, slow: -3.6 ± 5.2; DLS DA - fast: -3.8 ± 6.2, slow: -5.9 ± 7.8; DMS Glut - fast: -19.0 ± 5.3, slow: -18.6 ± 4.9; DMS DA - fast: 4.9 ± 5.1, slow: -8.6 ± 7.9).

Figure 4 Fast switchers show elevated DLS Glut and a greater difference between Glut and DA at trial start during the strategy switching test: This figure depicts mean tracings aligned to trial start and is organized as prior figures. **a,b)** The differences in Glut and DA levels did not reach statistical significance at trial start during the early portion of the strategy switching test (DLS Glut - fast: 4.3 ± 1.9, slow: 0.3 ± 1.6; DLS DA - fast: -0.7 ± 1.5, slow: 0.6 ± 1.6; DMS Glut - fast: 3.8 ± 1.6, slow: 0.3 ± 1.9; DMS DA - fast: -2.9 ± 1.7, slow: -0.7 ± 1.5). **c,d)** In the middle portion of the test, DLS Glut was elevated in fast switchers (fast: 7.4 ± 2.2, slow: 0.9 ± 1.8) and DLS DA was elevated in slow switchers (fast: -4.2 ± 0.8, slow: 1.4 ± 1.4). There was also a significant difference when comparing Glut relative to DA in the DLS of fast switchers (11.6 ± 2.0) that was not present in slow switchers (-0.5 ± 2.5). There were no differences in the DMS (Glut - fast: 3.8 ± 1.6, slow: 0.3 ± 1.9; DA - fast: -2.9 ± 1.7, slow: -0.7 ± 1.5). **e,f)** Late in the test, DLS Glut remained elevated in fast switchers both alone (DLS Glut - fast: 11.2 ± 3.0, slow: 0.8 ± 1.5) and relative to DLS DA (fast: -2.9 ± 1.3, slow: 2.0 ± 1.8). Again, there were no differences within the DMS (DMS Glut - fast: 1.4 ± 2.8, slow: 1.2 ± 2.1; DMS DA - fast: 0.6 ± 2.1, slow: 1.5 ± 2.3).

Figure 5 Fast switchers displayed test stage-dependent changes in DA and Glut release prior to and after correct choices that were not seen in slow switchers: This figure depicts mean tracings aligned to previous strategy and new strategy correct and is organized as prior figures. a) Rats could be correct by following the previous strategy when the new strategy happened to coincide with the previous strategy, such as when the strategy switches from light is correct to left is correct. Choosing left when the light is on the right would indicate a correct choice following the new strategy. **b)** Early in the strategy switching test, fast switchers showed elevated DMS Glut (5.2 ± 1.7) compared to slow switchers (-0.8 ± 2.1), but reduced DMS DA release (-2.2 ± 2.1) compared to slow switchers (5.2 ± 2.7). DLS Glut (fast: 7.9 ± 2.4, slow: 2.1 ± 2.4) and DA (fast: 1.6 ± 0.7, slow: 5.0 ± 2.9) followed a similar pattern but did not reach

significance. However, when quantifying Glut release, relative to DA release, between fast and slow switchers, we found that fast switchers had significantly higher Glut to DA ratios in both the DLS (6.4 ± 2.2) and the DMS (7.4 ± 2.6), while the slow switchers had elevated DA release, relative to Glut in both the DLS (-2.9 ± 2.3) and DMS (-6.1 ± 3.0). There were no differences in release measures during pellet retrieval during this epoch of the strategy switching test (DLS Glut - fast: 10.6 ± 8.9, slow: 12.1 ± 5.9; DLS DA - fast: 6.5 ± 6.2, slow: 5.3± 9.3; DMS Glut - fast: -21.9 ± 5.9, slow: -16.4 ± 5.4; DMS DA - fast: 6.8 ± 6.1, slow: 3.1 ± 4.9). **c)** Early in the strategy switching test, there were no significant differences when making a correct choice following the new strategy, although the patterns of release were very similar to those when making a correct choice following the previous strategy (DLS Glut - fast: 5.5 ± 1.8, slow: 1.0 ± 3.1; DLS DA - fast: -0.3 ± 2.1, slow: -0.7 ± 2.4; DMS Glut - fast: 3.2 ± 1.9, slow: 1.6 ± 2.4; DMS DA - fast: -1.9 ± 2.1, slow: 4.9 ± 3.8). This was also true during pellet retrieval (DLS Glut - fast: 0.5 ± 9.1, slow: -4.9 ± 8.6; DLS DA - fast: 12.5 ± 9.3, slow: -4.6 ± 11.1; DMS Glut - fast: -15.2 ± 5.6, slow: -9.2 ± 9.3; DMS DA - fast: 12.0 ± 5.7, slow: 2.2 ± 9.4). **d)** In the middle portion of the strategy switching test, the fast switchers had decreased DLS DA release (0.5 ± 1.1), while the slow switchers had elevated DLS DA release (4.9 ± 2.1) prior to making a correct choice following the previous strategy. Fast switchers also featured elevated Glut release in DLS (fast: 3.9 ± 1.1, slow: 0.9 ± 1.8) and DMS (Glut - fast: 3.6 ± 1.3, slow: -0.3 ± 2.4) and decreased DA release in DMS (fast: -2.3 ± 1.7, slow: 2.7 ± 2.5), but these effects did not reach statistical significance. Again however, fast switchers had elevated Glut in both the DMS and DLS when compared against DA release (DLS: 3.4 ± 4.5; DMS: 5.9 ± 6.4) compared to slow switchers (DLS -3.9 ± 10.2; DMS: -3.0 ± 7.9). When looking at pellet retrieval, fast switchers had significantly elevated DMS DA release (27.4 ± 10.1) compared to slow switchers (0.2 ± 5.8). No other release measures within the pellet retrieval period were significantly different (DLS Glut - fast: 9.2 ± 8.2, slow: 0.5 ± 10.6; DLS DA - fast: 13.2 ± 8.6, slow: -5.8 ± 12.0; DMS Glut - fast: -21.3 ± 12.6, slow: -22.1 ± 7.0). **e)** In fast switchers, DLS Glut release was elevated prior to a correct choice following the new strategy (fast: 9.1 ± 4.0, slow: 2.1 ± 1.3), but there were no other significant differences (DLS DA - fast: 0.3 ± 3.6, slow: 1.5 ± 1.9; DMS Glut - fast: -0.6 ± 2.9, slow: 0.6 ± 1.7; DMS DA - fast: 3.1 ± 1.3, slow: -1.0 ± 2.7). The differences during the pellet retrieval did not reach statistical significance due to high variability (DLS Glut - fast: 19.3 ± 10.2, slow: -0.4 ± 9.1; DLS DA - fast: 15.1 ± 16.2, slow: -8.4 ± 7.2; DMS Glut - fast: -19.2 ± 15.0, slow: -15.9 ± 9.8; DMS DA - fast: 9.2 ± 10.8, slow: 4.0 ± 6.5). **f)** By late in the strategy switching test, dorsal striatal DA and Glut release measures were no longer significantly different prior to a correct choice following the previous strategy (DLS Glut - fast: 7.2 ± 2.3, slow: 1.9 ± 2.4; DLS DA - fast: -1.3 ± 2.6, slow: -0.7 ± 2.2; DMS Glut - fast: 1.3 ± 2.0, slow: 1.4 ± 2.8; DMS DA - fast: -0.5 ± 2.4, slow: 0.4 ± 2.3) and its respective pellet retrieval (DLS Glut - fast: 7.5 ± 6.2, slow: 6.6 ± 17.3; DLS DA - fast: 3.2 ± 10.5, slow: -11.6 ± 10.5; DMS Glut - fast: -

16.1 ± 11.9, slow: -26.8 ± 13.8; DMS DA - fast: 9.8± 8.8, slow: 6.4 ± 11.4). **g)** This was also the case for the new strategy correct (DLS Glut - fast: 8.8 ± 3.0, slow: 2.8 ± 3.0; DLS DA - fast: 0.1 ± 1.9, slow: 0.5 ± 2.0; DMS Glut - fast: -2.3 ± 1.6, slow: 2.0 ± 2.7; DMS DA - fast: -4.2 ± 2.4, slow: 0.3 ± 3.1) and its respective pellet retrieval (DLS Glut - fast: 4.8 ± 7.4, slow: 8.5 ± 11.7; DLS DA - fast: 15.7 ± 9.2, slow: -5.6 ± 9.1; DMS Glut - fast: -19.3 ± 8.4 slow: -18.1 ± 11.4; DMS DA - fast: -5.0 ± 12.7, slow: 5.8 ± 14.6).

Table 1 Dorsal striatal DA and Glut release measures correlate with strategy switching performance and punishment resistant cocaine use.

Figure 6 DLS and DMS Glut release changes both converge and diverge across the discrimination paradigm and strategy switching test: a) DLS Glut release was elevated relative to DMS Glut throughout performance of the discrimination paradigm and late in the strategy switching test. This was particularly apparent on day 1 of the discrimination paradigm (P=0.0953) and late in the strategy switching test (P=0.0015). DMS Glut peaked during the early and middle portions of the strategy switching test. **b)** DMS and DLS Glut release remained low and even in slow switchers across all time points except at day 10. Interestingly, the release ratio of DLS and DMS Glut on day 10 in the slow switchers resembles the ratio of fast switchers early in the strategy switching test. **c)** Prior to a correct choice, there was a main effect of dorsal striatum region (F(1,20)=15.3, P=0.0009) and an interaction between region and time (F(8,150)=1.96, p=0.055; Fig. 6c), such that DLS Glut release was greater than DMS Glut on day 10 (P=0.024), late in the test at previous strategy correct (P=0.0245), and middle (P=0.0027) and late in the test (p<0.0001) at new strategy correct. DMS Glut release peaked early in the strategy switching test. **d)** These patterns were absent in the slow switchers, who despite having similar release measures across the discrimination paradigm to fast switchers, featured a surprising reduction in both DMS and DLS Glut early in the strategy switching test relative to their day 10 release levels.

Figure 7 Specific combinations of DLS and DMS Glut and DA release measures predict strategy switching performance and punishment resistant cocaine use: a) Fourteen neurotransmitter release metrics that were correlated with strategy switching performance and/or punishment sensitivity were tested for their prediction strength using a Chi-Square test. This list of metrics is the legend for the metric numbers across the rest of the figure. **b)** The prediction strength of each in relation to punishment sensitivity (y-axis) and strategy switching (x-axis) were plotted. Interestingly, measures 1, 3-7 and 9 were strong predictors of both punishment sensitivity and strategy switching, with measure 1 being the strongest predictor of both. **c)** Linear Discriminant Analysis (LDA) was used to examine every combination of the 14 neurotransmitter metrics and it identified 2 combinations of 2 metrics that correctly categorized the strategy switching performance of all rats, metrics 2 and 10 and **d)** metrics 4 and 8. **e)** The model identified 2 different combinations of 5 metrics

that correctly categorized the punishment sensitivity of all rats, metrics 3, 6, 11, 12, and 13 and **f)** metrics 3, 9, 11, 12, and 14. These metrics were combined into LDA component 1 and 2 in order to display the separation of the 5 metrics in a 2-D plane.

Figure 8 Inhibition of OFC and IN Thal inputs produce opposite effects on Glut release in DLS in putative fast and slow switchers: a) Each rat received bilateral infusions of a Gi-coupled DREADD into OFC and a KORD DREADD into IN Thal, and unilateral infusions of CRE, iGluSnFr, GRABDA, and photometry fibers into the DMS and DLS. This allowed simultaneous recording of DMS and DLS while individually or jointly inhibiting OFC and IN Thal inputs to the recording region. **b)** Representative images showing green staining from IN Thal and yellow staining from OFC into both the DMS and DLS (insets). Arrows indicate fiber placements in the DMS and DLS. **c)** Inhibition of OFC and IN Thal inputs produced different effects in putative fast and slow switchers ($F(9, 117) = 2.531$, $p=0.011$) that was specifically seen as an impairment in slow switchers on day 7 (fast-92.9, slow-84.1; $p=0.022$). **d)** Mean tracings are aligned to trial start (green box), correct choices (black box), and subsequent pellet retrieval (red box) of DLS Glut (red) and DLS DA (blue). Change in Z-score from no manipulation (vehicle) to inhibition of OFC (CNO) and IN Thal (SalB) inputs is quantified using area under the curve (Mean $\pm$ SEM) for the 1 or 2 seconds in the colored box, as indicated. Inhibition of OFC inputs resulted in an increase in DLS Glut in putative fast switchers (N=7) at both trial start (Veh: 1.84 $\pm$ 0.55, CNO: 3.51 $\pm$ 1.14) and correct choices (Veh: 0.89 $\pm$ 0.61, CNO: 2.51 $\pm$ 0.91), but did not affect DA release (trial start- Veh: -1.05 $\pm$ 1.12, CNO: -0.99 $\pm$ 0.57; correct- Veh: 0.18 $\pm$ 0.32, CNO: 0.06 $\pm$ 0.24). Inhibition of IN Thal inputs had no effect on Glut release (trial start- SalB: 1.29 $\pm$ 1.01; correct-SalB: 0.57 $\pm$ 1.02) or DA release (trial start- SalB: 0.26 $\pm$ 1.01; correct- SalB: -0.24 $\pm$ 1.09). **e)** Inhibition of OFC (Veh: 1.49 $\pm$ 0.49, CNO: 0.08 $\pm$ 1.15) and IN Thal (SalB: 0.78 $\pm$ 0.49) inputs reduced glut release at correct choice in putative slow switchers (N=7) but had no effect on DA release (Veh: 0.50 $\pm$ 0.80, CNO: 1.85 $\pm$ 1.93, SalB: 0.21 $\pm$ 0.80). There was no effect on Glut (Veh: 1.26 $\pm$ 0.78, CNO: 1.55 $\pm$ 1.02, SalB: 1.69 $\pm$ 1.13) or DA (Veh: -0.55 $\pm$ 0.46, CNO: -0.08 $\pm$ 0.74, SalB: -1.19 $\pm$ 0.71) at trial start.

Figure 9 OFC and IN Thal projections to dorsal striatum are responsible for behavioral timestamp-specific Glut release in putative fast switchers and are necessary for fast strategy switching and sensitivity to cocaine punishment: This figure depicts mean tracings aligned to trial start (green box), correct choices following the previous strategy (black box), and subsequent pellet retrieval (red box) of Glut (red) and DA (blue) in only putative fast switchers. Data across rows are within subjects and represents the change in signal (Z-score) from no manipulation (vehicle, left) to inhibition of OFC (CNO, middle) and then inhibition of IN Thal (SalB, right) inputs to dorsal striatum. The tracings are quantified using area under the curve (Mean $\pm$ SEM) for the 1 or 2 seconds in the colored box, as indicated. **a)** Inhibition of OFC inputs

decreased **DLS** Glut (Veh: 4.45 ± 1.99, CNO: 2.93 ± 1.51) and increased DA (Veh: -1.36 ± 2.36, CNO: 2.26 ± 1.02) at both trial initiation and at correct choice following the previous strategy (Glut-Veh: 8.36 ± 2.30, CNO: 3.69 ± 1.93; DA- Veh: -4.46 ± 3.25, CNO: 1.24 ± 1.40). Two of seven rats are not shown here because 1 rat finished the strategy switching test prior to the CNO injection and 1 rat had a medial placement in the OFC. Inhibition of IN Thal inputs also decreased DLS_Glut at both trial initiation (Veh: 5.46 ± 1.91, SalB: 2.12 ± 1.46) and a correct choice following the previous strategy (Veh: 7.42 ± 2.10, SalB: 3.48 ± 2.77). However, inhibition of IN Thal inputs did not affect DA release (trial start- Veh: -1.34 ± 1.92, SalB: -1.62 ± 2.26; Prev Strategy correct- Veh: -3.64 ± 2.31, SalB: -0.32 ± 1.15). One rat is not shown here because there was not adequate staining to confirm the thalamus to DLS inhibition. **b)** Inhibition of OFC inputs did not reduce **DMS** Glut at trial initiation (Veh: 2.75 ± 2.48, CNO: 2.19 ± 1.70) but did result in an increase DA release (Veh: -5.08 ± 1.89, CNO: -1.45 ± 1.44). Inhibition of OFC inputs decreased DMS Glut prior to a correct choice following the previous strategy (Veh: 1.40 ± 2.03, CNO: -1.36 ± 1.74) and decreased DMS DA during the pellet retrieval period (Veh: 20.88 ± 5.96, CNO: 9.25 ± 4.15). DA release at previous strategy correct was not affected by OFC inhibition (not quantified on graphic). Inhibition of IN Thal inputs decreased DMS Glut at trial initiation (Veh: 2.31 ± 1.96, SalB: -0.26 ± 1.86) but did not significantly affect DA release (Veh: -4.46 ± 1.43, SalB: -0.87 ± 2.04). Inhibition of IN Thal inputs did not significantly reduce DMS Glut prior to a correct choice following the previous strategy (Veh: 1.90 ± 1.44, SalB: 0.94 ± 1.66) but did still reduce DMS DA during the pellet retrieval period (Veh: 21.79 ± 5.10, SalB: 4.27 ± 9.70). DA release at previous strategy correct was not affected by IN Thal inhibition (not quantified on graphic). **c)** Putative fast switchers that finished the strategy switching test after treatment with CNO and SalB (errors: 67.8 ± 7.96) were significantly impaired compared to fast switchers from the first experiment (orange bar: 41.6±2.1). Putative slow switchers (50.0 ± 6.03) were improved relative to slow switchers from the first experiment (purple bar, 81.4±5.0). **d)** Putative fast and slow switchers took similar numbers of cocaine infusions across the 10 self-administration days; however, slow switchers (67.2 ± 6.31) took more infusion on day 10 than fast switchers (48.57± 3.41; p=0.039). **e)** Putative slow switchers took more punished cocaine infusions overall (Putative slow switchers: No shock-12.0±0.9, Low-13.4±0.9, Med-10.4±1.7, High-6.8±3.6; Slow + CNO/SalB: No Shock-11.8±0.7, Low-15.6±2.7, Med-15.4±2.2, High-8.2±3.3), relative to fast switchers (Putative fast switchers: No shock-11.7±0.8, Low-10.7±1.9, Med-5.3±1.4, High-2.7±1.0; Fast + CNO/SalB: No Shock-10.2±0.6, Low-12.8±0.5, Med-8.3±1.5, High-4.0±1.2). The medium shock intensities (highlighted by rectangle) are expanded in part f. **f)** Due to the short half life of SalB, the effect of inhibiting both IN thal and OFC inputs to dorsal striatum was limited to the medium shock bin (30 minutes total). Inhibition of OFC and IN Thal inputs increased the number of punished cocaine infusions that putative fast (P=0.005) and slow switchers

(P=0.0003) relative to their to vehicle injection. Slow switchers took more punished cocaine infusions than fast switchers under both vehicle (P=0.046) and SalB+CNO (P=0.008) conditions.

Figure 10 Working model: A dynamic interaction between DA (blue) and Glut (red) represents a general signature of cognitive flexibility (left) or inflexibility (right). The flexible signature features increased Glut release at behavioral initiation (DLS) and at a decision point (DLS and DMS), with a concurrent dip in DA. It also features increased DA and Glut release during the feedback period when a correct choice is rewarded. The inflexible signature features increased DA release at behavioral initiation and a decision point, but not during the feedback period. It also features no change in Glut release at behavioral initiation, prior to a choice, or after a choice. These signatures predicted which rats went on to be both fast or slow strategy switchers, as well as which were sensitive or resistant to punishment.

References:

1 Ostlund, S. B. & Balleine, B. W. On habits and addiction: An associative analysis of compulsive drug seeking. *Drug Discov Today Dis Models* **5**, 235-245 (2008). https://doi.org:10.1016/j.ddmod.2009.07.004

2 Melugin, P. R., Nolan, S. O. & Siciliano, C. A. Bidirectional causality between addiction and cognitive deficits. *Int Rev Neurobiol* **157**, 371-407 (2021). https://doi.org:10.1016/bs.irn.2020.11.001

3 Izquierdo, A. & Jentsch, J. D. Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology (Berl)* **219**, 607-620 (2012). https://doi.org:10.1007/s00213-011-2579-7

4 Cervantes, M. C., Laughlin, R. E. & Jentsch, J. D. Cocaine self-administration behavior in inbred mouse lines segregating different capacities for inhibitory control. *Psychopharmacology (Berl)* **229**, 515-525 (2013). https://doi.org:10.1007/s00213-013-3135-4

5 Loos, M., Staal, J., Smit, A. B., De Vries, T. J. & Spijker, S. Enhanced alcohol self-administration and reinstatement in a highly impulsive, inattentive recombinant inbred mouse strain. *Front Behav Neurosci* **7**, 151 (2013). https://doi.org:10.3389/fnbeh.2013.00151

6 Shnitko, T. A., Gonzales, S. W. & Grant, K. A. Low cognitive flexibility as a risk for heavy alcohol drinking in non-human primates. *Alcohol* **74**, 95-104 (2019). https://doi.org:10.1016/j.alcohol.2018.04.007

7 Corral, M., Holguin, S. R. & Cadaveira, F. Neuropsychological characteristics of young children from high-density alcoholism families: a three-year follow-up. *J Stud Alcohol* **64**, 195-199 (2003). https://doi.org:10.15288/jsa.2003.64.195

8 Darvas, M. & Palmiter, R. D. Contributions of striatal dopamine signaling to the modulation of cognitive flexibility. *Biol Psychiatry* **69**, 704-707 (2011). https://doi.org:10.1016/j.biopsych.2010.09.033

9 Lhost, J. *et al.* Interplay Between Inhibitory Control and Behavioural Flexibility: Impact of Dorsomedial Striatal Dopamine Denervation in Mice. *Neuroscience* **477**, 25-39 (2021). https://doi.org:10.1016/j.neuroscience.2021.09.026

10 Grospe, G. M., Baker, P. M. & Ragozzino, M. E. Cognitive Flexibility Deficits Following 6-OHDA Lesions of the Rat Dorsomedial Striatum. *Neuroscience* **374**, 80-90 (2018). https://doi.org:10.1016/j.neuroscience.2018.01.032

11 van der Merwe, R. K. *et al.* Characterization of striatal dopamine projections across striatal subregions in behavioral flexibility. *Eur J Neurosci* **58**, 4466-4486 (2023). https://doi.org:10.1111/ejn.15910

12 Giovanniello, J. R. *et al.* A dual-pathway architecture enables chronic stress to promote habit formation. *bioRxiv* (2023). https://doi.org:10.1101/2023.10.03.560731

13 Schumacher, J. D., van Holstein, M., Bagrodia, V., Le Bouder, H. B. & Floresco, S. B. Dorsomedial striatal contributions to different forms of risk/reward decision making. *Neurobiol Learn Mem* **178**, 107369 (2021). https://doi.org:10.1016/j.nlm.2020.107369

14 Jin, X. & Costa, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457-462 (2010). https://doi.org:10.1038/nature09263

15 Cover, K. K., Lieberman, A. G., Heckman, M. M. & Mathur, B. N. The rostral intralaminar nuclear complex of the thalamus supports striatally mediated action reinforcement. *Elife* **12** (2023). https://doi.org:10.7554/eLife.83627

16 Wolff, S. B. E., Ko, R. & Olveczky, B. P. Distinct roles for motor cortical and thalamic inputs to striatum during motor skill learning and execution. *Sci Adv* **8**, eabk0231 (2022). https://doi.org:10.1126/sciadv.abk0231

17 Crego, A. C. G. *et al.* Complementary Control over Habits and Behavioral Vigor by Phasic Activity in the Dorsolateral Striatum. *J Neurosci* **40**, 2139-2153 (2020). https://doi.org:10.1523/JNEUROSCI.1313-19.2019

18 Willuhn, I., Burgeno, L. M., Everitt, B. J. & Phillips, P. E. Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proc Natl Acad Sci U S A* **109**, 20703-20708 (2012). https://doi.org:10.1073/pnas.1213460109

19 Bender, B. N., Stringfield, S. J. & Torregrossa, M. M. Changes in dorsomedial striatum activity during expression of goal-directed vs. habit-like cue-induced cocaine seeking. *Addict Neurosci* **11** (2024). https://doi.org:10.1016/j.addicn.2024.100149

20 Bender, B. N. & Torregrossa, M. M. Intermittent cocaine self-administration has sex-specific effects on addiction-like behaviors in rats. *Neuropharmacology* **230**, 109490 (2023). https://doi.org:10.1016/j.neuropharm.2023.109490

21 Bentzley, B. S., Jhou, T. C. & Aston-Jones, G. Economic demand predicts addiction-like behavior and therapeutic efficacy of oxytocin in the rat. *Proc Natl Acad Sci U S A* **111**, 11822-11827 (2014). https://doi.org:10.1073/pnas.1406324111

22 James, M. H. *et al.* Increased Number and Activity of a Lateral Subpopulation of Hypothalamic Orexin/Hypocretin Neurons Underlies the Expression of an Addicted State in Rats. *Biol Psychiatry* **85**, 925-935 (2019). https://doi.org:10.1016/j.biopsych.2018.07.022

23 Allain, F. & Samaha, A. N. Revisiting long-access versus short-access cocaine self-administration in rats: intermittent intake promotes addiction symptoms independent of session length. *Addict Biol* **24**, 641-651 (2019). https://doi.org:10.1111/adb.12629

24 Garcia, A. F., Webb, I. G., Yager, L. M., Seo, M. B. & Ferguson, S. M. Intermittent but not continuous access to cocaine produces individual variability in addiction susceptibility in rats. *Psychopharmacology (Berl)* **237**, 2929-2941 (2020). https://doi.org:10.1007/s00213-020-05581-1

25 Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22**, 513-523 (2005). https://doi.org:10.1111/j.1460-9568.2005.04218.x

26 Corbit, L. H. & Janak, P. H. Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur J Neurosci* **31**, 1312-1321 (2010). https://doi.org:10.1111/j.1460-9568.2010.07153.x

27 Featherstone, R. E. & McDonald, R. J. Lesions of the dorsolateral striatum impair the acquisition of a simplified stimulus-response dependent conditional discrimination task. *Neuroscience* **136**, 387-395 (2005). https://doi.org:10.1016/j.neuroscience.2005.08.021

28 Featherstone, R. E. & McDonald, R. J. Lesions of the dorsolateral or dorsomedial striatum impair performance of a previously acquired simple discrimination task. *Neurobiol Learn Mem* **84**, 159-167 (2005). https://doi.org:10.1016/j.nlm.2005.08.003

29 Balleine, B. W. & Dezfouli, A. Hierarchical Action Control: Adaptive Collaboration Between Actions and Habits. *Front Psychol* **10**, 2735 (2019). https://doi.org:10.3389/fpsyg.2019.02735

30 Balleine, B. W., Liljeholm, M. & Ostlund, S. B. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res* **199**, 43-52 (2009). https://doi.org:10.1016/j.bbr.2008.10.034

31 Bradfield, L. A., Bertran-Gonzalez, J., Chieng, B. & Balleine, B. W. The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. *Neuron* **79**, 153-166 (2013). https://doi.org:10.1016/j.neuron.2013.04.039

32 Dalton, G. L., Wang, N. Y., Phillips, A. G. & Floresco, S. B. Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *J Neurosci* **36**, 1996-2006 (2016). https://doi.org:10.1523/JNEUROSCI.3366-15.2016

33 Gremel, C. M. & Costa, R. M. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* **4**, 2264 (2013). https://doi.org:10.1038/ncomms3264

34 Izquierdo, A. Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *J Neurosci* **37**, 10529-10540 (2017). https://doi.org:10.1523/JNEUROSCI.1678-17.2017

35 Minamimoto, T., Hori, Y. & Kimura, M. Roles of the thalamic CM-PF complex-Basal ganglia circuit in externally driven rebias of action. *Brain Res Bull* **78**, 75-79 (2009). https://doi.org:10.1016/j.brainresbull.2008.08.013

36 Minamimoto, T. & Kimura, M. Participation of the thalamic CM-Pf complex in attentional orienting. *J Neurophysiol* **87**, 3090-3101 (2002). https://doi.org:10.1152/jn.2002.87.6.3090

37 Smith, Y., Surmeier, D. J., Redgrave, P. & Kimura, M. Thalamic contributions to Basal Ganglia-related behavioral switching and reinforcement. *J Neurosci* **31**, 16102-16106 (2011). https://doi.org:10.1523/JNEUROSCI.4634-11.2011

38 Smith, Y. *et al.* The thalamostriatal system in normal and diseased states. *Front Syst Neurosci* **8**, 5 (2014). https://doi.org:10.3389/fnsys.2014.00005

39 Kato, S. *et al.* Action Selection and Flexible Switching Controlled by the Intralaminar Thalamic Neurons. *Cell Rep* **22**, 2370-2382 (2018). https://doi.org:10.1016/j.celrep.2018.02.016

40 Nisenbaum, E. S. & Wilson, C. J. Potassium currents responsible for inward and outward rectification in rat neostriatal spiny projection neurons. *J Neurosci* **15**, 4449-4463 (1995). https://doi.org:10.1523/JNEUROSCI.15-06-04449.1995

41 Plenz, D. & Kitai, S. T. Up and down states in striatal medium spiny neurons simultaneously recorded with spontaneous activity in fast-spiking interneurons studied in cortex-striatum-substantia nigra organotypic cultures. *J Neurosci* **18**, 266-283 (1998). https://doi.org:10.1523/JNEUROSCI.18-01-00266.1998

42 Kornhuber, J. & Kornhuber, M. E. Presynaptic dopaminergic modulation of cortical input to the striatum. *Life Sci* **39**, 699-674 (1986). https://doi.org:10.1016/0024-3205(86)90013-5

43 Godukhin, O. V., Zharikova, A. D. & Budantsev, A. Role of presynaptic dopamine receptors in regulation of the glutamatergic neurotransmission in rat neostriatum. *Neuroscience* **12**, 377-383 (1984). https://doi.org:10.1016/0306-4522(84)90059-9

44 Goto, Y. & Grace, A. A. Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* **8**, 805-812 (2005). https://doi.org:10.1038/nn1471

45 Centonze, D. *et al.* Distinct roles of D1 and D5 dopamine receptors in motor activity and striatal synaptic plasticity. *J Neurosci* **23**, 8506-8512 (2003). https://doi.org:10.1523/JNEUROSCI.23-24-08506.2003

46 Kerr, J. N. & Wickens, J. R. Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *J Neurophysiol* **85**, 117-124 (2001). https://doi.org:10.1152/jn.2001.85.1.117

47 Gerfen, C. R. The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. *Nature* **311**, 461-464 (1984). https://doi.org:10.1038/311461a0

48 Lee, H. J. *et al.* Activation of Direct and Indirect Pathway Medium Spiny Neurons Drives Distinct Brain-wide Responses. *Neuron* **91**, 412-424 (2016). https://doi.org:10.1016/j.neuron.2016.06.010

49 Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848-851 (2008). https://doi.org:10.1126/science.1160575

50 Ding, J. B., Guzman, J. N., Peterson, J. D., Goldberg, J. A. & Surmeier, D. J. Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* **67**, 294-307 (2010). https://doi.org:10.1016/j.neuron.2010.06.017

51 Wichmann, T. & DeLong, M. R. Functional and pathophysiological models of the basal ganglia. *Curr Opin Neurobiol* **6**, 751-758 (1996). https://doi.org:10.1016/s0959-4388(96)80024-9

52 Hunnicutt, B. J. *et al.* A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife* **5** (2016). https://doi.org:10.7554/eLife.19103

53 Takahashi, Y. K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269-280 (2009). https://doi.org:10.1016/j.neuron.2009.03.005

54 Elena Erro, M., Lanciego, J. L. & Gimenez-Amaya, J. M. Re-examination of the thalamostriatal projections in the rat with retrograde tracers. *Neurosci Res* **42**, 45-55 (2002). https://doi.org:10.1016/s0168-0102(01)00302-9

55 Groenewegen, H. J. & Berendse, H. W. The specificity of the 'nonspecific' midline and intralaminar thalamic nuclei. *Trends Neurosci* **17**, 52-57 (1994). https://doi.org:10.1016/0166-2236(94)90074-4

56 Parent, M. & Parent, A. Single-axon tracing and three-dimensional reconstruction of centre median-parafascicular thalamic neurons in primates. *J Comp Neurol* **481**, 127-144 (2005). https://doi.org:10.1002/cne.20348

57 Peters, A. J., Liu, H. & Komiyama, T. Learning in the Rodent Motor Cortex. *Annu Rev Neurosci* **40**, 77-97 (2017). https://doi.org:10.1146/annurev-neuro-072116-031407

58 Floresco, S. B., Block, A. E. & Tse, M. T. Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav Brain Res* **190**, 85-96 (2008). https://doi.org:10.1016/j.bbr.2008.02.008

59 Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C. & Tse, M. T. Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology* **31**, 297-309 (2006). https://doi.org:10.1038/sj.npp.1300825

60 Floresco, S. B., Zhang, Y. & Enomoto, T. Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. *Behav Brain Res* **204**, 396-409 (2009). https://doi.org:10.1016/j.bbr.2008.12.001

61 St Onge, J. R. & Floresco, S. B. Prefrontal cortical contribution to risk-based decision making. *Cereb Cortex* **20**, 1816-1828 (2010). https://doi.org:10.1093/cercor/bhp250

62 Hart, G., Leung, B. K. & Balleine, B. W. Dorsal and ventral streams: the distinct role of striatal subregions in the acquisition and performance of goal-directed actions. *Neurobiol Learn Mem* **108**, 104-118 (2014). https://doi.org:10.1016/j.nlm.2013.11.003

63 Lichtenberg, N. T. *et al.* The Medial Orbitofrontal Cortex-Basolateral Amygdala Circuit Regulates the Influence of Reward Cues on Adaptive Behavior and Choice. *J Neurosci* **41**, 7267-7277 (2021). https://doi.org:10.1523/JNEUROSCI.0901-21.2021

64 Yin, H. H., Knowlton, B. J. & Balleine, B. W. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res* **166**, 189-196 (2006). https://doi.org:10.1016/j.bbr.2005.07.012
65 Hamid, A. A., Frank, M. J. & Moore, C. I. Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell* **184**, 2733-2749 e2716 (2021). https://doi.org:10.1016/j.cell.2021.03.046
66 Hart, G., Burton, T. J., Nolan, C. R. & Balleine, B. W. Striatal dopamine release tracks the relationship between actions and their consequences. *Cell Rep* **43**, 113828 (2024). https://doi.org:10.1016/j.celrep.2024.113828
67 van Elzelingen, W. *et al.* Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. *Curr Biol* **32**, 1163-1174 e1166 (2022). https://doi.org:10.1016/j.cub.2021.12.027
68 Seiler, J. L. *et al.* Dopamine signaling in the dorsomedial striatum promotes compulsive behavior. *Curr Biol* **32**, 1175-1188 e1175 (2022). https://doi.org:10.1016/j.cub.2022.01.055
69 He, Y. *et al.* Early-life adversity selectively interrupts the dendritic differentiation of dorsolateral striatal neurons in male mice. *Brain Struct Funct* **226**, 397-414 (2021). https://doi.org:10.1007/s00429-020-02183-7
70 Somerville, L. H., Hare, T. & Casey, B. J. Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *J Cogn Neurosci* **23**, 2123-2134 (2011). https://doi.org:10.1162/jocn.2010.21572
71 Xu, B. *et al.* The Impacts of Early-life Adversity on Striatal and Hippocampal Memory Functions. *Neuroscience* **490**, 11-24 (2022). https://doi.org:10.1016/j.neuroscience.2022.02.029
72 Goldfarb, E. V., Frobose, M. I., Cools, R. & Phelps, E. A. Stress and Cognitive Flexibility: Cortisol Increases Are Associated with Enhanced Updating but Impaired Switching. *J Cogn Neurosci* **29**, 14-24 (2017). https://doi.org:10.1162/jocn_a_01029
73 de Carvalho, G., Khoja, S., Haile, M. T. & Chen, L. Y. Early life adversity impaired dorsal striatal synaptic transmission and behavioral adaptability to appropriate action selection in a sex-dependent manner. *Front Synaptic Neurosci* **15**, 1128640 (2023). https://doi.org:10.3389/fnsyn.2023.1128640
74 Gomes, F. V., Zhu, X. & Grace, A. A. The pathophysiological impact of stress on the dopamine system is dependent on the state of the critical period of vulnerability. *Mol Psychiatry* **25**, 3278-3291 (2020). https://doi.org:10.1038/s41380-019-0514-1
75 Abercrombie, E. D., Keefe, K. A., DiFrischia, D. S. & Zigmond, M. J. Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J Neurochem* **52**, 1655-1658 (1989). https://doi.org:10.1111/j.1471-4159.1989.tb09224.x
76 Vaessen, T., Hernaus, D., Myin-Germeys, I. & van Amelsvoort, T. The dopaminergic response to acute stress in health and psychopathology: A systematic review. *Neurosci Biobehav Rev* **56**, 241-251 (2015). https://doi.org:10.1016/j.neubiorev.2015.07.008

77 Carvalheiro, J., Conceicao, V. A., Mesquita, A. & Seara-Cardoso, A. Acute stress blunts prediction error signals in the dorsal striatum during reinforcement learning. *Neurobiol Stress* **15**, 100412 (2021). https://doi.org:10.1016/j.ynstr.2021.100412

78 Bortz, D. M., Feistritzer, C. M. & Grace, A. A. Medial prefrontal cortex to medial septum pathway activation improves cognitive flexibility in rats. *Int J Neuropsychopharmacol* (2023). https://doi.org:10.1093/ijnp/pyad019

79 Bortz, D. M., Feistritzer, C. M., Power, C. C. & Grace, A. A. Medial septum activation improves strategy switching once strategies are well-learned via bidirectional regulation of dopamine neuron population activity. *Neuropsychopharmacology* (2022). https://doi.org:10.1038/s41386-022-01387-1

80 Bender, B. N., Stringfield, S. J. & Torregrossa, M. M. Changes in dorsomedial striatum activity mediate expression of goal-directed vs. habit-like cue-induced cocaine seeking. *bioRxiv* (2023). https://doi.org:10.1101/2023.07.24.550364

81 Bender, B. N. & Torregrossa, M. M. Molecular and circuit mechanisms regulating cocaine memory. *Cell Mol Life Sci* **77**, 3745-3768 (2020). https://doi.org:10.1007/s00018-020-03498-8

82 Bender, B. N. & Torregrossa, M. M. Dorsolateral striatum dopamine-dependent cocaine seeking is resistant to pavlovian cue extinction in male and female rats. *Neuropharmacology* **182**, 108403 (2021). https://doi.org:10.1016/j.neuropharm.2020.108403

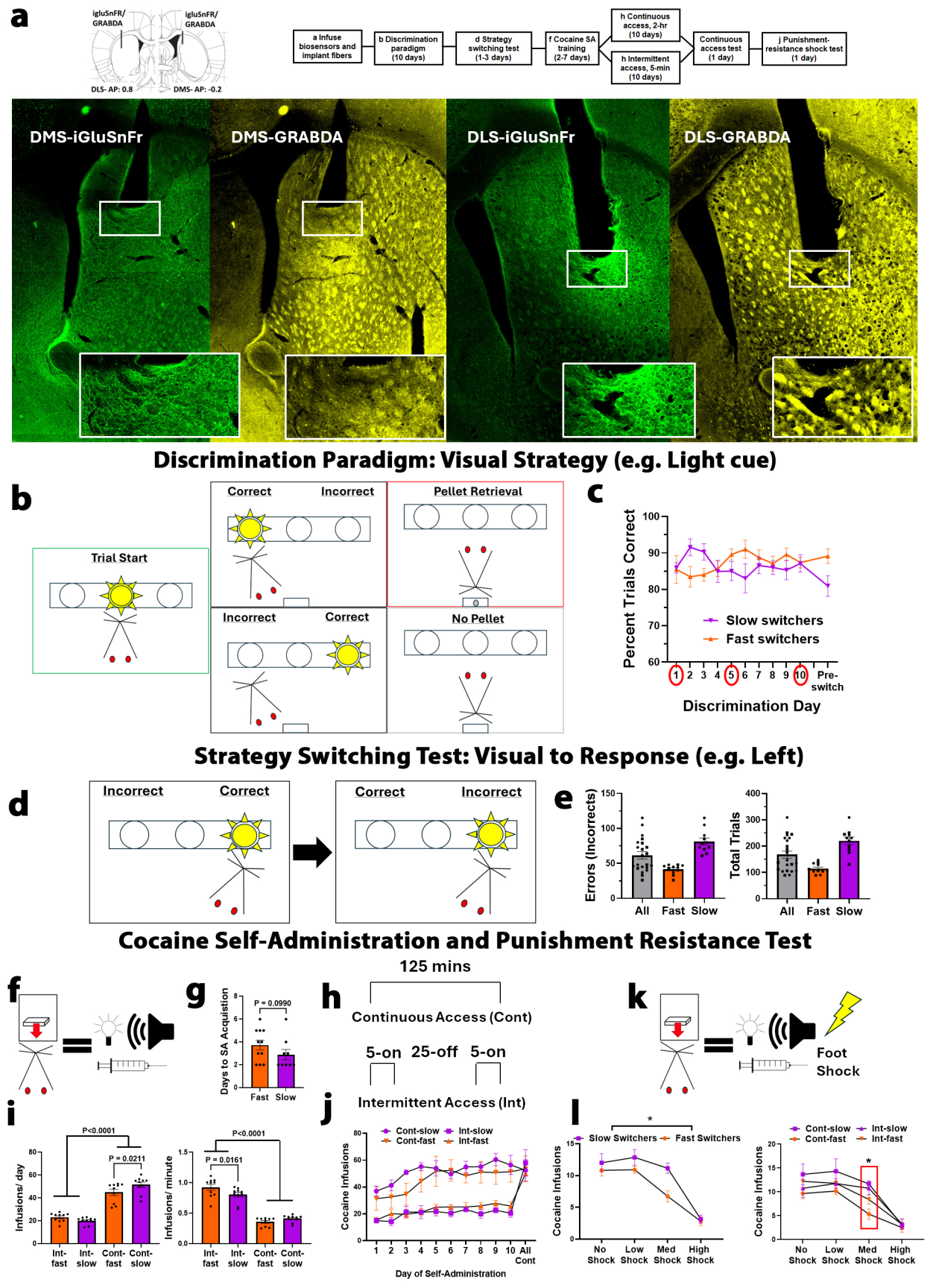
83 Sun, F. *et al.* Next-generation GRAB sensors for monitoring dopaminergic activity in vivo. *Nat Methods* **17**, 1156-1166 (2020). https://doi.org:10.1038/s41592-020-00981-9

84 Marvin, J. S. *et al.* Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. *Nat Methods* **15**, 936-939 (2018). https://doi.org:10.1038/s41592-018-0171-3

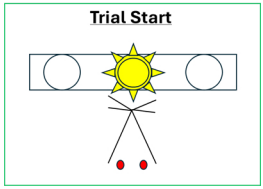
85 Vardy, E. *et al.* A New DREADD Facilitates the Multiplexed Chemogenetic Interrogation of Behavior. *Neuron* **86**, 936-946 (2015). https://doi.org:10.1016/j.neuron.2015.03.065

86 Krashes, M. J. *et al.* Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J Clin Invest* **121**, 1424-1428 (2011). https://doi.org:10.1172/JCI46229

87 Sherathiya, V. N., Schaid, M. D., Seiler, J. L., Lopez, G. C. & Lerner, T. N. GuPPy, a Python toolbox for the analysis of fiber photometry data. *Sci Rep* **11**, 24212 (2021). https://doi.org:10.1038/s41598-021-03626-9

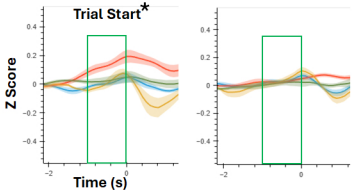

a
igluSnFR/ GRABDA
igluSnFR/ GRABDA
DLS- AP: 0.8
DMS- AP: -0.2
a Infuse biosensors and implant fibers
b Discrimination paradigm (10 days)
d Strategy switching test (1-3 days)
f Cocaine SA training (2-7 days)
h Continuous access, 2-hr (10 days)
h Intermittent access, 5-min (10 days)
Continuous access test (1 day)
j Punishment-resistance shock test (1 day)
DMS-iGluSnFr
DMS-GRABDA
DLS-iGluSnFr
DLS-GRABDA
Discrimination Paradigm: Visual Strategy (e.g. Light cue)
b
Trial Start
Correct
Incorrect
Pellet Retrieval
Incorrect
Correct
No Pellet
c
Percent Trials Correct
Slow switchers
Fast switchers
Pre-switch
Discrimination Day
Strategy Switching Test: Visual to Response (e.g. Left)
d
Incorrect
Correct
Correct
Incorrect
e
Errors (Incorrects)
Total Trials
All
Fast
Slow
Cocaine Self-Administration and Punishment Resistance Test
f
g
Days to SA Acquisition
P = 0.0990
Fast
Slow
h
125 mins
Continuous Access (Cont)
5-on
25-off
5-on
Intermittent Access (Int)
k
Foot Shock
i
Infusions/ day
P<0.0001
P = 0.0211
Infusions/ minute
P<0.0001
P = 0.0161
Int-fast
Int-slow
Cont-fast
Cont-slow
j
Cont-slow
Int-slow
Cont-fast
Int-fast
Cocaine Infusions
Day of Self-Administration
All Cont
l
Slow Switchers
Fast Switchers
Cocaine Infusions
No Shock
Low Shock
Med Shock
High Shock

Discrimination Paradigm: Trial Start

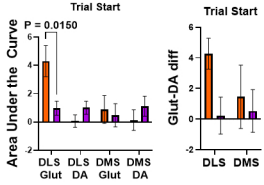


DLS Glut | DLS DA | DMS Glut | DMS DA

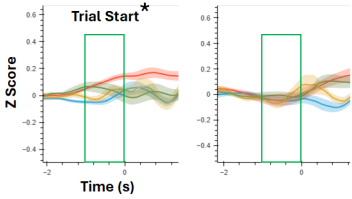
a Fast Switchers, N=11 | Slow Switchers, N=11



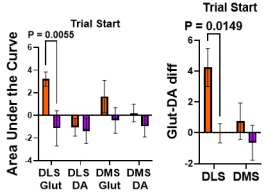
b Day 1



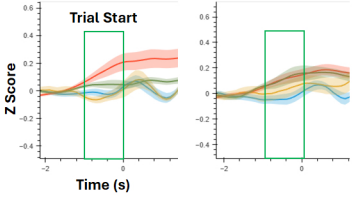
c Fast Switchers, N=11 | Slow Switchers, N=11



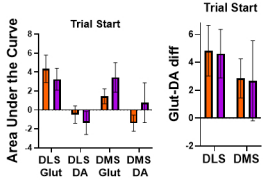
d Day 5



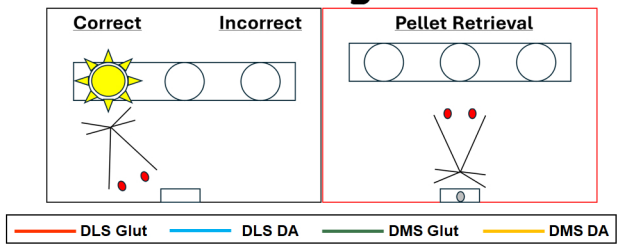
e Fast Switchers, N=11 | Slow Switchers, N=11



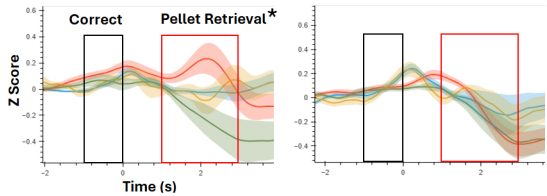
f Day 10



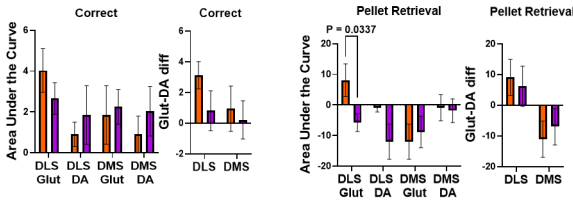
Discrimination Paradigm: Correct Choice



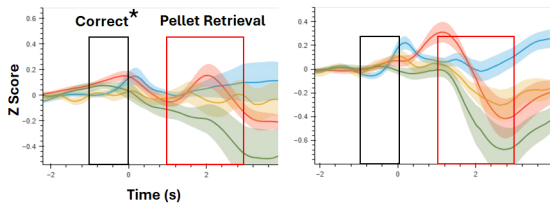
a Day 1 — Fast Switchers, N=11 — Slow Switchers, N=11



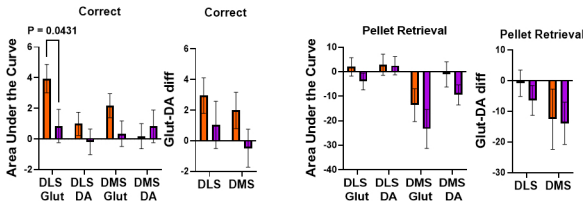
b



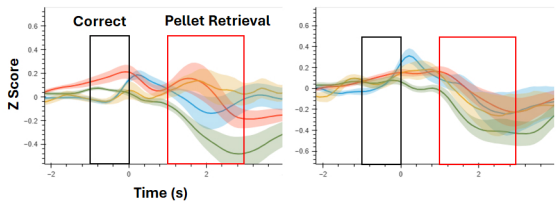
c Day 5 — Fast Switchers, N=11 — Slow Switchers, N=11



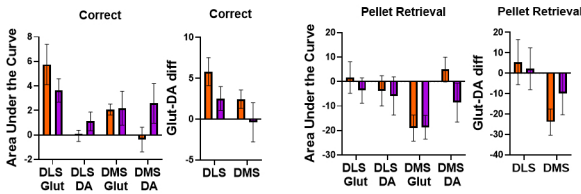
d



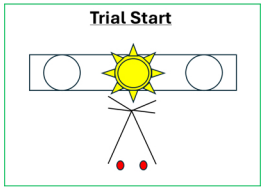
e Day 10 — Fast Switchers, N=11 — Slow Switchers, N=11



f



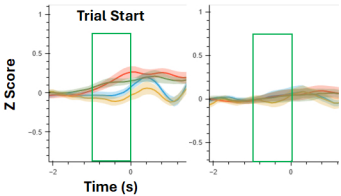
Strategy Switching Test: Trial Start



DLS Glut | DLS DA | DMS Glut | DMS DA

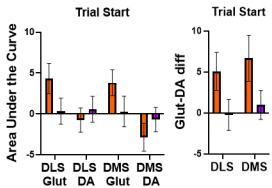
a

Fast Switchers, N=11 | Slow Switchers, N=11



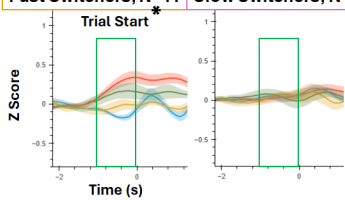
b

Early



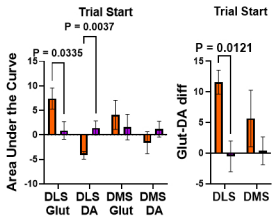
c

Fast Switchers, N=11 | Slow Switchers, N=11



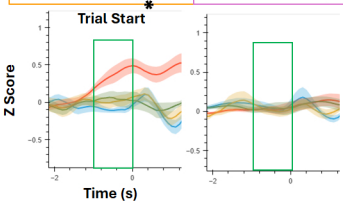
d

Middle



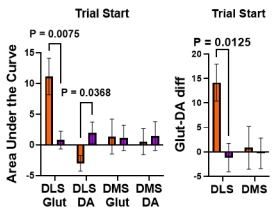
e

Fast Switchers, N=11 | Slow Switchers, N=11

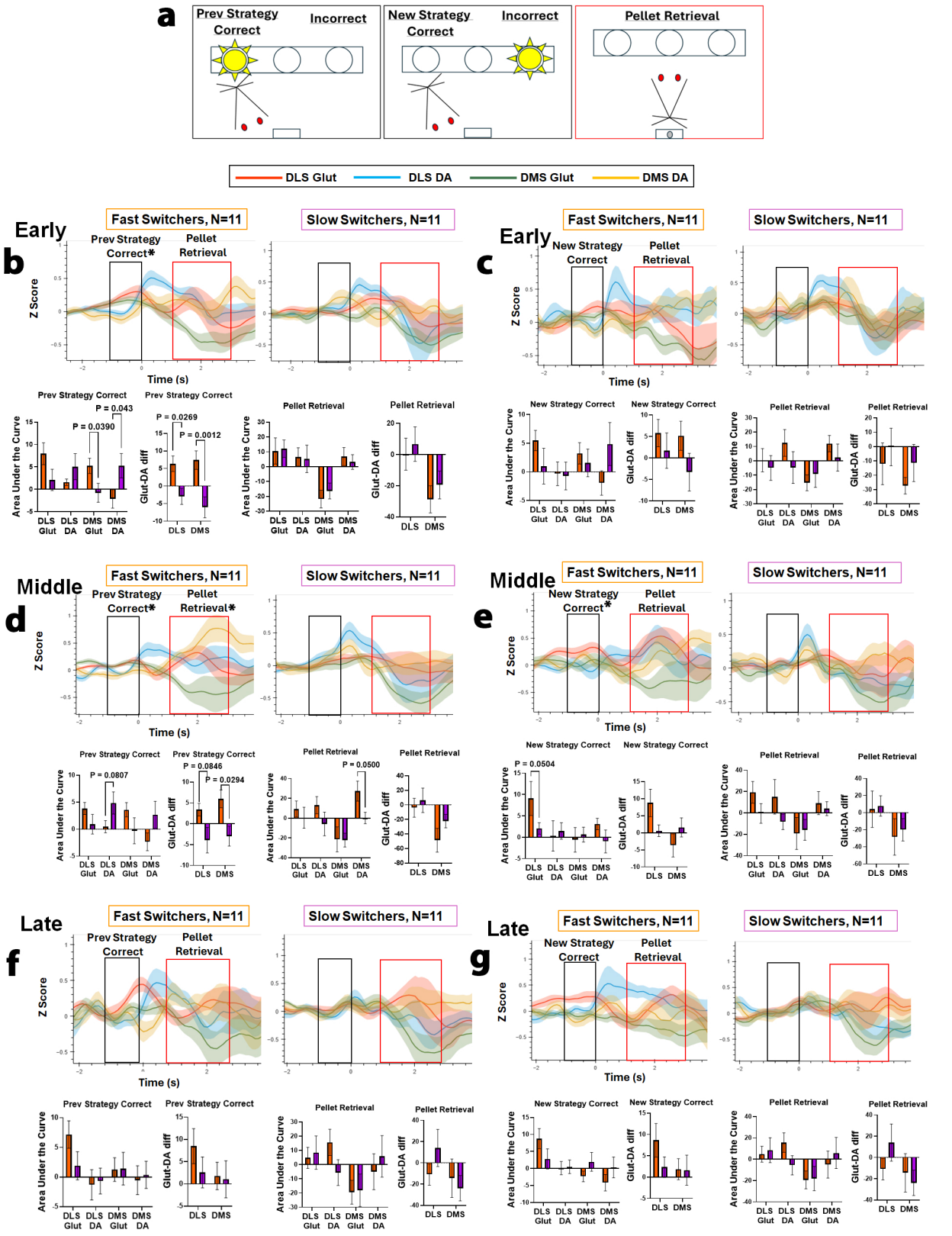


f

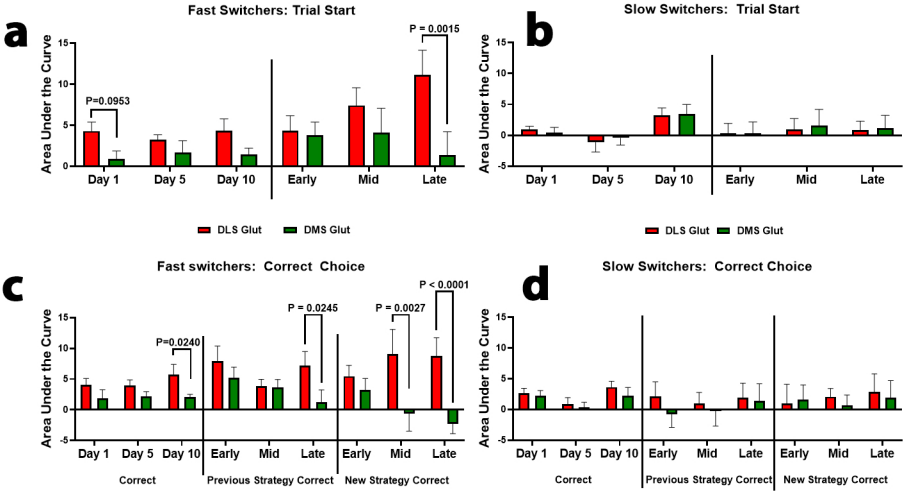
Late

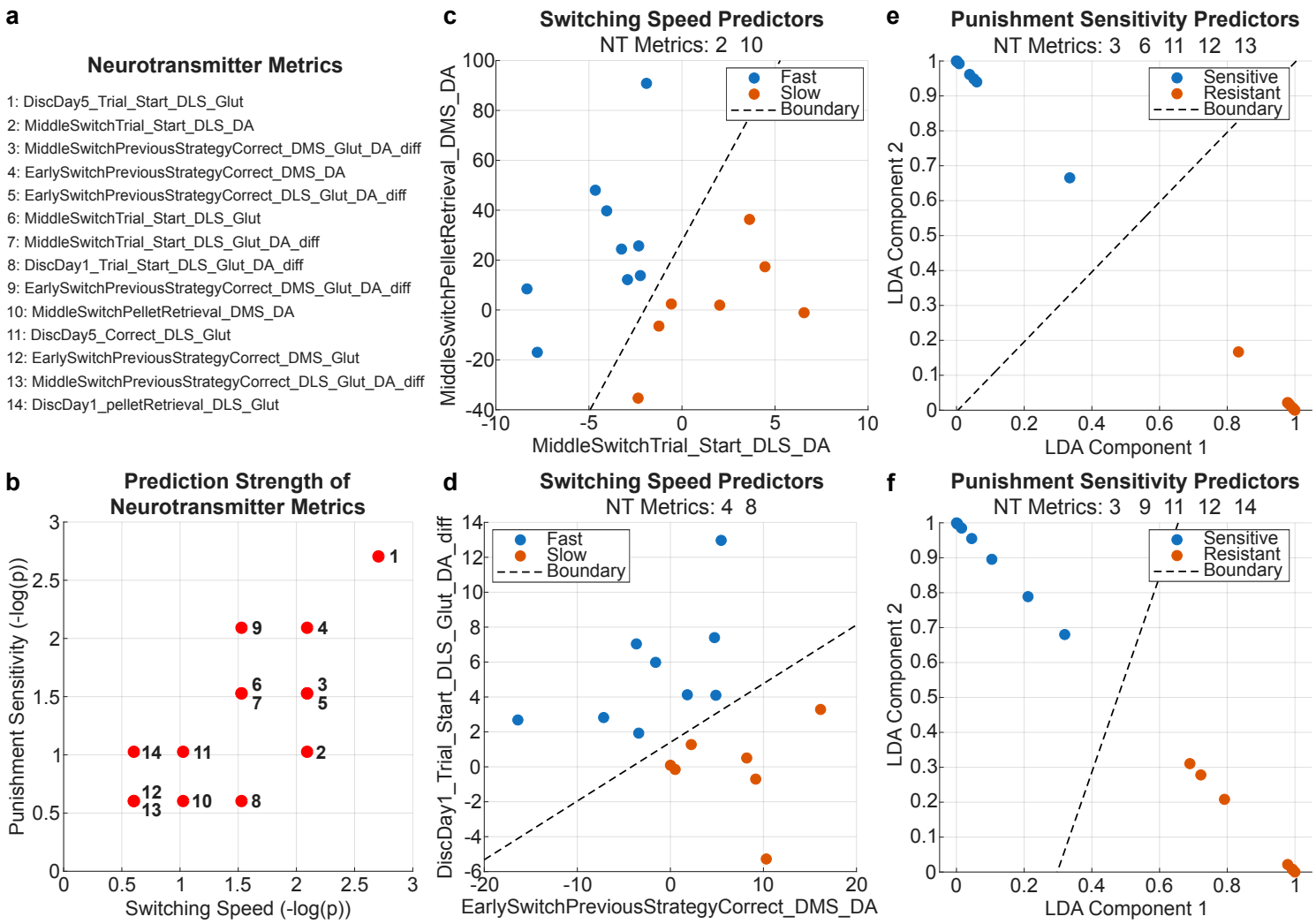


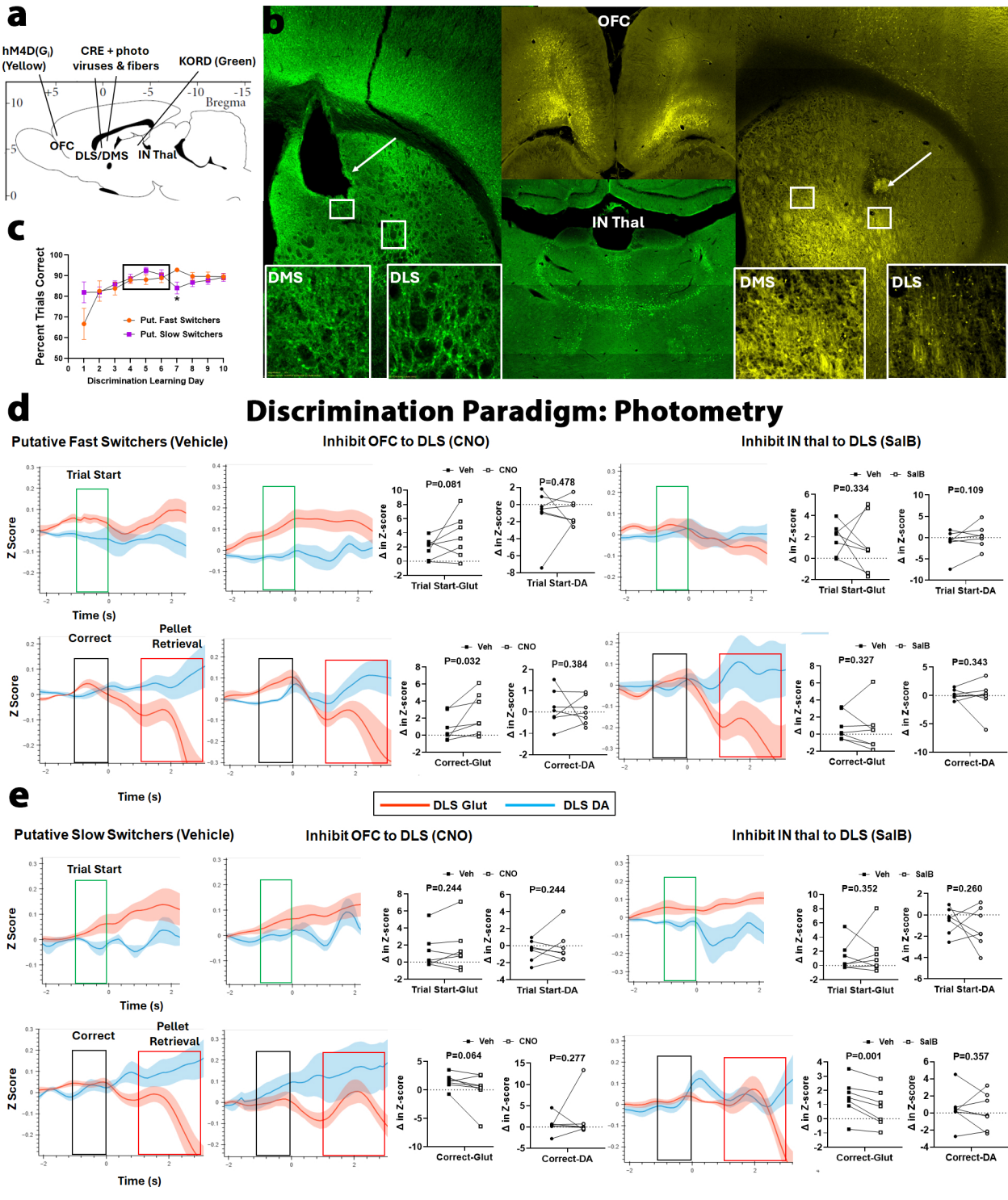
Strategy Switching Test: Correct Choice



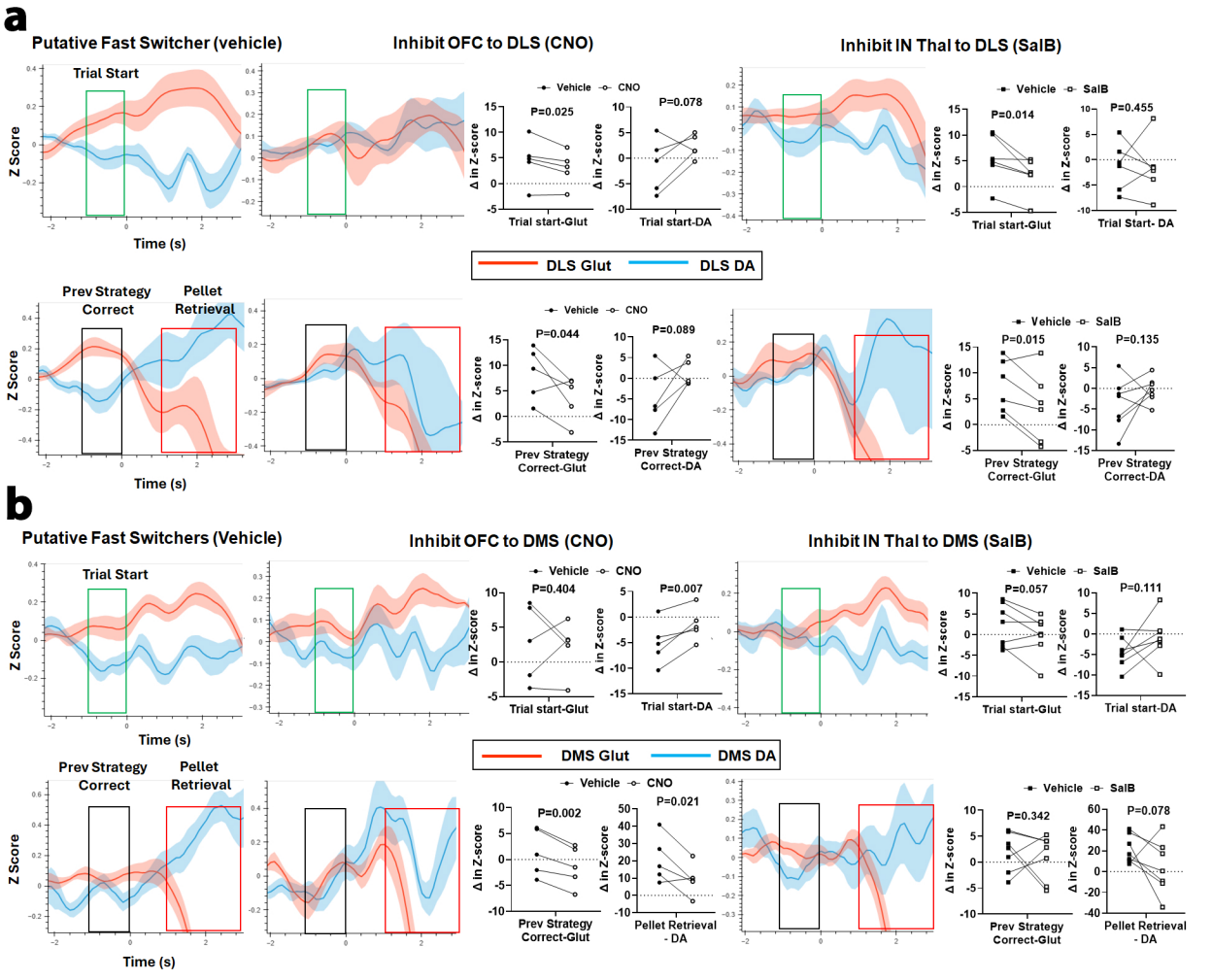
DLS Glut Versus DMS Glut



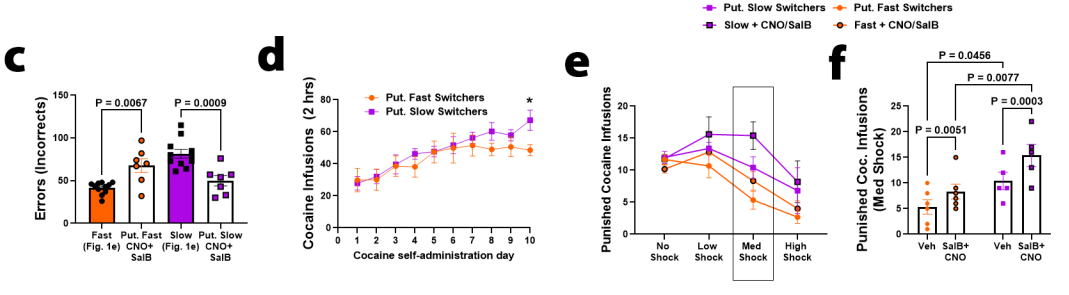

a
Neurotransmitter Metrics
1: DiscDay5_Trial_Start_DLS_Glut
2: MiddleSwitchTrial_Start_DLS_DA
3: MiddleSwitchPreviousStrategyCorrect_DMS_Glut_DA_diff
4: EarlySwitchPreviousStrategyCorrect_DMS_DA
5: EarlySwitchPreviousStrategyCorrect_DLS_Glut_DA_diff
6: MiddleSwitchTrial_Start_DLS_Glut
7: MiddleSwitchTrial_Start_DLS_Glut_DA_diff
8: DiscDay1_Trial_Start_DLS_Glut_DA_diff
9: EarlySwitchPreviousStrategyCorrect_DMS_Glut_DA_diff
10: MiddleSwitchPelletRetrieval_DMS_DA
11: DiscDay5_Correct_DLS_Glut
12: EarlySwitchPreviousStrategyCorrect_DMS_Glut
13: MiddleSwitchPreviousStrategyCorrect_DLS_Glut_DA_diff
14: DiscDay1_pelletRetrieval_DLS_Glut
b
Prediction Strength of Neurotransmitter Metrics
Punishment Sensitivity (-log(p))
Switching Speed (-log(p))
c
Switching Speed Predictors
NT Metrics: 2 10
Fast
Slow
Boundary
MiddleSwitchPelletRetrieval_DMS_DA
MiddleSwitchTrial_Start_DLS_DA
d
Switching Speed Predictors
NT Metrics: 4 8
Fast
Slow
Boundary
DiscDay1_Trial_Start_DLS_Glut_DA_diff
EarlySwitchPreviousStrategyCorrect_DMS_DA
e
Punishment Sensitivity Predictors
NT Metrics: 3 6 11 12 13
Sensitive
Resistant
Boundary
LDA Component 2
LDA Component 1
f
Punishment Sensitivity Predictors
NT Metrics: 3 9 11 12 14
Sensitive
Resistant
Boundary
LDA Component 2
LDA Component 1

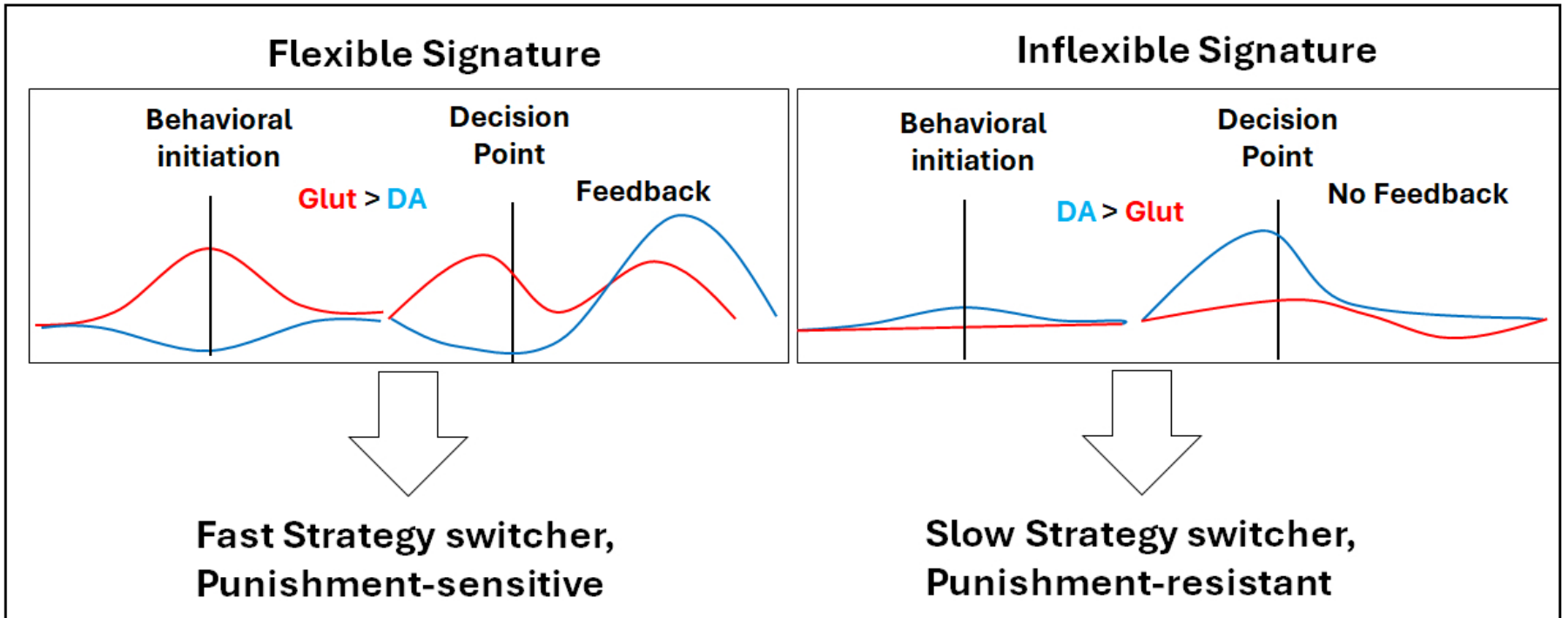

a
hM4D(Gi) (Yellow)
CRE + photo viruses & fibers
KORD (Green)
Bregma
OFC
DLS/DMS
IN Thal
b
OFC
IN Thal
DMS
DLS
DMS
DLS
c
Percent Trials Correct
Put. Fast Switchers
Put. Slow Switchers
Discrimination Learning Day
d
Discrimination Paradigm: Photometry
Putative Fast Switchers (Vehicle)
Inhibit OFC to DLS (CNO)
Inhibit IN thal to DLS (SalB)
Trial Start
Z Score
Time (s)
Veh
CNO
SalB
P=0.081
P=0.478
P=0.334
P=0.109
Δ in Z-score
Trial Start-Glut
Trial Start-DA
Correct
Pellet Retrieval
P=0.032
P=0.384
P=0.327
P=0.343
Correct-Glut
Correct-DA
e
DLS Glut
DLS DA
Putative Slow Switchers (Vehicle)
Inhibit OFC to DLS (CNO)
Inhibit IN thal to DLS (SalB)
P=0.244
P=0.244
P=0.352
P=0.260
P=0.064
P=0.277
P=0.001
P=0.357

Strategy Switching Test: Photometry



Strategy Switching, Cocaine Self-Administration: Behavior




Flexible Signature
Behavioral initiation
Decision Point
Glut > DA
Feedback
Fast Strategy switcher, Punishment-sensitive
Inflexible Signature
Behavioral initiation
Decision Point
DA > Glut
No Feedback
Slow Strategy switcher, Punishment-resistant

	Punished cocaine infusions		Strategy switching errors	
	Pearson r:	P value:	Pearson r:	P value:
Strategy switching errors	**0.51**	**0.041**		
Disc day 1 Trial Start DLS Glut	-0.34	0.194	-0.40	0.064
Disc day 1 Trial Start DLS Glut-DA diff	-0.41	0.118	**-0.43**	**0.047**
Disc day 1 pellet retrieval DLS Glut	**-0.67**	**0.004**	-0.24	0.286
Disc day 5 Trial Start DLS Glut	-0.47	0.069	-0.35	0.117
Disc day 5 Correct DLS Glut	-0.29	0.279	**-0.56**	**0.009**
Early switch previous strategy correct DMS Glut	-0.44	0.086	-0.34	0.127
Early switch previous strategy correct DMS DA	**0.54**	**0.032**	0.22	0.317
Early switch previous strategy correct DLS Glut-DA diff	**-0.55**	**0.029**	-0.28	0.199
Early switch previous strategy correct DMS Glut-DA diff	**-0.75**	**0.001**	-0.38	0.085
Middle switch Trial Start DLS Glut	-0.50	0.067	-0.32	0.164
Middle switch Trial Start DLS DA	**0.61**	**0.021**	**0.46**	**0.042**
Middle switch Trial Start DLS Glut-DA diff	**-0.68**	**0.008**	**-0.46**	**0.039**
Middle switch previous strategy correct DLS DA	0.27	0.359	0.35	0.126
Middle switch previous strategy correct DLS Glut-DA diff	**-0.57**	**0.034**	-0.36	0.116
Middle switch previous strategy correct DMS Glut-DA diff	**-0.58**	**0.029**	**-0.49**	**0.030**
Middle switch pellet retrieval DMS DA	-0.17	0.551	**-0.54**	**0.014**
Middle switch new strategy correct DLS Glut	-0.14	0.657	-0.36	0.136
Late switch Trial Start DLS Glut	-0.44	0.098	**-0.49**	**0.029**
Late switch Trial Start DLS DA	0.34	0.213	**0.46**	**0.041**
Late switch Trial Start DLS Glut-DA diff	-0.44	0.100	**-0.53**	**0.017**