Experience-Dependent Changes in Nucleus Accumbens Activity Predict Cued Approach Learning: Contribution of NMDA Receptors

Mercedes Vega Villar
The Graduate Center, City University of New York

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Behavioral Neurobiology Commons, Biological Psychology Commons, Experimental Analysis of Behavior Commons, and the Systems Neuroscience Commons

How does access to this work benefit you? Let us know!

Recommended Citation

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
EXPERIENCE-DEPENDENT CHANGES IN NUCLEUS ACCUMBENS ACTIVITY PREDICT CUED APPROACH LEARNING: CONTRIBUTION OF NMDA RECEPTORS

by

MERCEDES VEGA VILLAR

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2020
Experience-dependent changes in nucleus accumbens activity predict cued approach learning: contribution of NMDA receptors

by

Mercedes Vega Villar

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Jon C. Horvitz
Chair of Examining Committee

Date

Richard Bodnar
Executive Officer

Supervisory Committee:

Jeff Beeler
Andrew Delamater
Eleanor Simpson

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

Experience-dependent changes in nucleus accumbens activity predict cued approach learning: contribution of NMDA receptors

by

Mercedes Vega Villar

Advisor: Jon C. Horvitz

Animals learn associations between environmental cues and the natural rewards they predict (e.g., food, water, sex). As a result, reward-predictive cues come to trigger vigorous reward-seeking responses. Many neurons in the nucleus accumbens (NAc) become excited upon presentation of an already-learned reward-predictive cue. These NAc responses encode the motivational value of the cue and are necessary for the expression of the subsequent approach behavior. However, the precise temporal relationship between the emergence of cue-evoked excitations in the NAc and the acquisition of cued approach behavior remains unknown. In Experiment 1, NAc activity was recorded as rats learned to approach a reward receptacle upon presentation of a cue. The results from this experiment indicate that cue-evoked excitations begin to grow a few trials before cued approach behavior is detected and they continue to escalate as the learned response becomes more vigorous.

NAc neurons undergo N-methyl-D-Aspartate receptor (NMDAR)-mediated plasticity. However, whether NMDAR-dependent plasticity in this structure is necessary for learning to seek and secure rewards in the presence of reward-predictive stimuli remains unclear. In Experiments 2 and 3, intra-accumbens infusions of NMDAR antagonists at different points of training reveal the dynamic involvement of these receptors, untangling their specific contribution to the acquisition and expression of cued approach behavior. To understand the neural mechanisms by which NAc
NMDARs participate in appetitive learning, in Experiment 4, local infusions of an antagonist were combined with electrophysiological recordings from NAc neurons during training. The results from this experiment reveal that the potentiation of training-induced cue-evoked signals in the NAc depends on NMDAR-dependent plasticity within this structure. Taken together, the results from this work link NAc plasticity, changes in NAc activity and the emergence of conditioned behavior, revealing a neural mechanism that allows animals to predict, approach and procure what is beneficial to them.
ACKNOWLEDGMENTS

I want to express my gratitude to my advisor, Dr. Jon C. Horvitz, for mentoring and guiding me throughout this difficult journey with tremendous kindness and a great sense of humor. I also want to thank Dr. Saleem M. Nicola, for allowing me to be part of his lab, providing me with encouragement, generous support and for teaching me so much. I have had the rare privilege of having not one, but two incredible mentors, both of which have been perfect role models as scientists, educators and human beings. I will be forever grateful for it. I also want to thank my dissertation committee for selflessly offering their time and efforts. Dr. Rosa Caamaño Tubío and Dr. Kevin Caref for kindly sharing their expertise and being patient with me. I also want to acknowledge the support of “La Caixa” Foundation for believing in me and allowing me to even consider the possibility of pursuing a graduate degree in this country.

I want to thank my parents, Mercedes and Florentino, my brother Juan and his fiancée Guadalupe, as well as the rest of my family and my husband’s family for accepting my choices and for their unconditional love and support. To my friends, especially Silvia, Susana and Blanca, for always being there for me. Finally, I want to thank my husband, Andrew, for putting up with my graduate school anxieties and for getting me through it with food, love and adventures.

DEDICATION

I would like to dedicate this work to my late grandmother Pepita and my late grandfather Enrique, who lived long enough to see me embark on this project. They are a constant source of good memories and inspiration.
The work presented in this dissertation has recently been published in article format. The article’s figures, supplementary materials and parts of the text have been reused and adapted in this manuscript.

# TABLE OF CONTENTS

1. INTRODUCTION ................................................................................................................................. 1
   1.1. General overview .......................................................................................................................... 2
   1.2. The nucleus accumbens (NAc) ...................................................................................................... 3
      1.2.1. The nucleus accumbens in the context of basal ganglia circuitry .............................................. 3
      1.2.2. Identity of neurons in the striatum ......................................................................................... 8
      1.2.3. Dopamine in the nucleus accumbens ...................................................................................... 12
      1.2.4. The “limbic-motor interface” hypothesis of NAc function .................................................... 20
   1.3. Connecting synaptic plasticity with associative learning ............................................................. 23
      1.3.1. Experience-dependent synaptic plasticity: a historical perspective ........................................ 23
      1.3.2. Mechanisms: inducing, expressing and maintaining synaptic plasticity .................................. 28
      1.3.3. Are plasticity and learning causally linked? ........................................................................... 32
   1.4. Plasticity in the nucleus accumbens and appetitive learning ....................................................... 38
      1.4.1. Artificial induction of synaptic plasticity in the NAc ............................................................. 38
      1.4.2. Connecting plasticity in the NAc with learning ..................................................................... 39
      1.4.3. Unresolved questions: goals of this research project ................................................................. 55
   1.5. General materials and methods .................................................................................................... 57
      1.5.1. Subjects .................................................................................................................................. 57
      1.5.2. Operant chambers .................................................................................................................. 58
      1.5.3. Surgeries ............................................................................................................................... 58
      1.5.4. Drugs ..................................................................................................................................... 59
      1.5.5. Behavioral indices .................................................................................................................. 59
      1.5.6. Change point analysis ............................................................................................................ 63
      1.5.7. Statistical analyses .................................................................................................................. 67
      1.5.8. Histology ............................................................................................................................... 69
2. EXPERIMENTS ....................................................................................................................................... 70
   2.1. Experiment 1: Experience-dependent changes in NAc activity during appetitive learning ........... 71
      2.1.1. Introduction ............................................................................................................................ 71
      2.1.2. Materials and methods .......................................................................................................... 72
      2.1.3. Results .................................................................................................................................... 76
   2.2. Experiment 2. Transient contribution of NAc core NMDARs to the expression of cued approach behavior and NAc cue encoding ................................................................. 85
      2.2.1. Introduction ............................................................................................................................ 85
      2.2.2. Materials and methods .......................................................................................................... 86
      2.2.3. Results .................................................................................................................................... 90
   2.3. Experiment 3: Effects of intra-accumbens NMDAR blockade on acquisition of cued approach
behavior (above and beyond performance effects)........................................................................................................... 92
2.3.1. Introduction.......................................................................................................................................................... 92
2.3.2. Materials and methods......................................................................................................................................... 93
2.3.3. Results................................................................................................................................................................. 96
2.4. Experiment 4: NMDAR contribution to the emergence of training-induced changes in NAc activity. 98
2.4.1. Introduction........................................................................................................................................................ 98
2.4.2. Materials and methods......................................................................................................................................... 99
2.4.3. Results................................................................................................................................................................. 100
3. DISCUSSION.............................................................................................................................................................. 106
3.1. Summary of the results. ........................................................................................................................................ 107
3.2. General discussion.................................................................................................................................................. 111
   3.2.1. How experience shapes behavior in this paradigm......................................................................................... 111
   3.2.2. How experience shapes NAc activity: features, mechanisms and significance........................................... 112
   3.2.3. NAc NMDARs (transiently) contribute to the expression of cued approach behavior ................................ 121
   3.2.4. What these results say about the role of NAc plasticity on appetitive learning........................................ 125
   3.2.5. Final considerations........................................................................................................................................ 140
4. APPENDIX................................................................................................................................................................. 142
   4.1. Supplementary Figures .................................................................................................................................. 142
   4.2. Supplementary Table......................................................................................................................................... 160
       4.2.1. Statistical tests. Main figures. .......................................................................................................................... 160
       4.2.2. Statistical tests. Supplementary figures. ......................................................................................................... 167
5. REFERENCES............................................................................................................................................................... 174
LIST OF FIGURES

Figure 1. Cortico-striatal-thalamic-cortical circuits. .................................................................5
Figure 2. Nucleus accumbens afferents ...................................................................................7
Figure 3. Performance index calculation ..................................................................................61
Figure 4. Identifying the behavioral change point .................................................................65
Figure 5. Schematic depiction of the cannulated microelectrode array ..................................73
Figure 6. During training, an identifiable transition (“change point”, CP) between exploratory and cued approach behavior occurs .........................................................82
Figure 7. Cue-evoked excitations increase before CP and continue to escalate as cued approach becomes more vigorous .................................................................83
Figure 8. NAc firing patterns during rewarded entries switch from excitations (before CP) to inhibitions (after CP) .................................................................84
Figure 9. Setup for simultaneous electrophysiological recordings and microinjections ........88
Figure 10. NMDARs play a transient role in the expression of NAc cue-evoked excitations and the vigor of the subsequent approach response .................................................91
Figure 11. NMDARs in NAc are necessary for the acquisition of cued approach behavior ........97
Figure 12. Unilateral blockade of NMDARs in NAc core did not interfere with learning in many subjects .............................................................................................................103
Figure 13. In the AP5-treated hemisphere (but not in the vehicle hemisphere), large and prolonged NAc cue-evoked excitations failed to emerge during training ........................................104
Figure 14. The effects of AP5 on NAc firing outlast the presence of the drug in the brain ..........105
Figure 15. Model of the stages of learning cued approach behavior and the mechanisms implicated in the emergence and expression of learning-related S+-evoked excitations in NAc neurons. ......110
Supplementary Figure 1. Representative individual neurons at different points of training …..144

Supplementary Figure 2. Additional graphs showing NAc core firing activity during.................146

Supplementary Figure 3. Activity in the NAc around the time of receptacle entry. ................148

Supplementary Figure 4. NAc cue-evoked excitations emerged during training regardless of whether the electrodes were driven down in between sessions or not .............................................149

Supplementary Figure 5. Bilateral blockade of NMDARs during training disrupts the emergence of cue-evoked excitations in NAc as well as the acquisition of cued approach behavior ............150

Supplementary Figure 6. Additional graphs showing how behavior and NAc core activity were affected by bilateral AP5 microinjections after moderate or extended training.................................151

Supplementary Figure 7. Baseline firing rate in NAc core was not affected by unilateral AP5 infusions............................................................................................................................................153

Supplementary Figure 8. NAc activity around the time of S+, S- or ITI entry in animals treated with unilateral AP5 microinjections......................................................................................................................155

Supplementary Figure 9. Cue-evoked excitations did not emerge in the NAc core neurons of animals that failed to learn the task under daily unilateral AP5 injections (“non-learners”)........157

Supplementary Figure 10. Anatomical location of injection and recording sites....................159
LIST OF ABBREVIATIONS

AMPAR: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor

AP5: D(-)-2-Amino-5-phosphonopentanoic acid

BLA: basolateral amygdala

CaMKII: calcium-calmodulin dependent kinase II

cAMP: cyclic adenosine monophosphate

CNI: cholinergic interneuron

CP: change point

CS: conditioned stimulus

EPN: entopeduncular nucleus

EPSP: excitatory postsynaptic potential

ERK-MAPK: extracellular-regulated kinase-mitogen activated protein kinase

FR: fixed ratio

GABA: gamma-aminobutyric acid

GECI: genetically encoded calcium indicator

GPe: external globus pallidus

GPi: internal globus pallidus

IEGs: immediate-early genes

ITI: intertrial interval

LTD: long-term depression

LTP: long-term potentiation/ L-LTP: long-lasting long-term potentiation

MAGUKs: membrane-associated guanylate kinases

mPFC: medial prefrontal cortex

MSN: medium spiny neuron

NAc: nucleus accumbens
NMDAR: N-Methyl-D-Aspartate receptor
PBS: phosphate buffered saline
PFC: prefrontal cortex
PIT: Pavlovian-to-instrumental transfer
PKA: protein kinase A
PKC: protein kinase C
PSD: postsynaptic density
PV/FSI: parvalbumin-expressing fast-spiking interneurons
RPE: reward prediction error
SNC: substantia nigra pars compacta
SNR: substantia nigra pars reticulata
SOM: somatostin-expressing interneurons
SPM: synaptic plasticity and memory
STDP: spike-timing dependent plasticity
STN: subthalamic nucleus
TAN: tonically active neuron
US: unconditioned stimulus
VM: ventral mesencephalon
VP: ventral pallidum
VTA: ventral tegmental area
1. INTRODUCTION
1.1. General overview

Reward-predictive stimuli signal animals to interrupt ongoing behavior and move towards the predicted location of a reward. In this form of adaptive behavior, approach responses are invigorated by cue-reward associations. This function requires the integration of motivational and motor neural systems, which has long been thought to depend on the nucleus accumbens (NAc; Mogenson, Jones, & Yim, 1980). Accordingly, individual neurons in the NAc of trained animals exhibit cue-evoked excitations that encode the reward-predictive value of the stimulus (Schultz, Apicella, Scarnati, & Ljungberg, 1992; Hassani, Cromwell, & Schultz, 2001; Cromwell & Schultz, 2003; Nicola, Yun, Wakabayashi, & Fields, 2004a; Day, Wheeler, Roitman, & Carelli, 2006) while simultaneously determining the vigor of the ensuing approach response (Ambroggi, Ghazizadeh, Nicola, & Fields, 2011; Caref & Nicola, 2018; du Hoffmann & Nicola, 2014; McGinty, Lardeux, Taha, Kim, & Nicola, 2013; Morrison, McGinty, du Hoffmann, & Nicola, 2017). If the role of the NAc is to allow reward-paired cues to gain access to motor output systems, one would expect to find, during training, changes in cue-evoked NAc firing that precede or accompany the emergence of cued approach behavior. However, despite evidence that the NAc participates in the acquisition of appetitive conditioned responses (Saunders, Richard, Margolis, & Janak, 2018; Day & Carelli, 2007; Gerdjikov, Giles, Swain, & Beninger, 2007; Di Ciano, Cardinal, Cowell, Little, & Everitt, 2001; Hernandez, Andrzejewski, Sadeghian, Panksepp, & Kelley, 2005; Kelley, Smith-Roe, & Holahan, 1997), no study to date has examined the precise temporal relationship between experience-dependent changes in NAc activity and the emergence of cued approach behavior.

Plasticity mechanisms in the amygdala and the dorsal striatum are widely accepted as the

---

1 When the conditioned stimulus is in a specific spatial location, animals can also exhibit approach responses to the cue itself ("sign-tracking") as opposed to the predicted location of the reward ("goal-tracking"; Hearst & Jenkins, 1974).
likely substrates of fear conditioning (Luchkina & Bolshakov, 2019) and procedural skill learning (Perrin & Venance, 2019), respectively. Surprisingly, it is still unclear whether plasticity in the NAc plays the same crucial role in cued approach learning. While excitatory synapses onto NAc neurons are known to undergo N-methyl-D-Aspartate receptor (NMDAR)-dependent long-term potentiation (LTP; Floresco, Blaha, Yang, & Phillips, 2001; Goto & Grace, 2005; Kombian & Malenka, 1994; LeGates et al., 2018; Pennartz, Ameerun, Groenewegen, & Lopes da Silva, 1993; Popescu, Saghyan, & Paré, 2007; Yagishita et al., 2014), the relevance of this neuroplastic potential in the context of natural learning is poorly understood. If NMDAR-dependent NAc plasticity were necessary for cued approach learning, NMDAR activation might be expected to mediate the growth of cue-evoked excitations in the NAc during training. By blocking NMDARs within the NAc while monitoring the firing rate of NAc neurons, the present work revealed the dynamic contribution of NMDARs to both NAc activity and cued approach behavior at different stages of learning. The results identify and characterize a likely physiological substrate for the natural acquisition of cued approach behavior.

1.2. The nucleus accumbens (NAc)

1.2.1. The nucleus accumbens in the context of basal ganglia circuitry

The basal ganglia are a set of interconnected nuclei located in the base of the forebrain. These nuclei are embedded in a series of recurrent circuits that connect cortical, thalamic and mesencephalic structures (Alexander, DeLong, & Strick, 1986). Deficits as diverse and debilitating as Parkinson’s and Huntington’s diseases, Tourette’s syndrome or addiction are the result of basal ganglia pathology (Albin & Mink, 2006; Dauer & Przedborski, 2003; Robbins & Everitt, 2002), which explains why so much research has been devoted to understanding how these nuclei operate.
in the disease-free state.

The basal ganglia receive information from brain areas that process sensory, motor, cognitive and motivational information. This information is conveyed to the basal ganglia in the form of excitatory inputs from the widespread areas of the cortex, thalamus and other subcortical regions as well as dopaminergic inputs from mesolimbic structures. These inputs converge onto the striatum, which is the largest structure in the basal ganglia. Striatal neurons integrate these inputs and project to motor output areas. The striatum is subdivided into a dorsal and a ventral portion. The dorsal portion consists of the caudate and the putamen in primates, two nuclei separated by the internal capsula. In rodents, these structures are less differentiated, and they are usually known as “dorsomedial” and “dorsolateral” striatum respectively. The ventral striatum includes the main part of the olfactory tubercle, the subcommissural part of the main body of the striatum and the NAc (Heimer & Van Hoesen, 1979).

Dorsal striatal neurons project directly (“direct pathway”) or indirectly (“indirect pathway”) to two structures: the internal globus pallidus (GPi) and the substantia nigra pars reticulata (SNr). The GPi (in primates, the homologous region in rodents is the entopeduncular nucleus, EPN) and the SNr are the link that connects basal ganglia structures to motor nuclei in the thalamus, tectum (superior colliculus) and pedunculopontine nucleus, and therefore represent the output stage of this circuit. While striatal neurons of the “direct pathway” project to the GPi (EPN)/SNr monosynaptically, neurons in the “indirect pathway” do so polysynaptically, via the external globus pallidus (GPe) and the subthalamic nucleus (STN). The direct and indirect pathways exert opposite effects on basal ganglia output nuclei (Figure 1).

Neurons in the ventral striatum regulate thalamic motor output via the SNr and the ventral tegmental area (VTA), collectively referred to as the ventral mesencephalon (VM). The equivalent of the GPe in the ventral striatal downstream circuit is the ventral pallidum (VP; (Heimer, Zahm,
Churchill, Kalivas, & Wohltmann, 1991). Generally speaking, the direct/indirect dichotomy also applies to the ventral striatum, as some of its output neurons project directly to the VM whereas others do so via the VP and the STN. However, in NAc circuitry, these pathways are less segregated than the direct/indirect pathways of the dorsal striatum (Kupchik et al., 2015; Kupchik & Kalivas, 2017). For example, although the VP is relay between NAc neurons and the VM in the “indirect” NAc pathway, many VP neurons project directly to motor output nuclei in the thalamus (Kupchik et al., 2015; Tripathi, Prensa, & Mengual, 2013).

![Cortico-striatal-thalamic-cortical circuits](image)

**Figure 1. Cortico-striatal-thalamic-cortical circuits.** Inhibitory projections from the dorsal (left) and ventral (right) striatum exert bidirectional control over thalamocortical motor output structures depending on the basal ganglia areas they innervate. *GPe*: external globus pallidus; *GPi*: internal globus pallidus; *SNr*: substantia nigra pars reticulata; *VTA*: ventral tegmental area; *STN*: subthalamic nucleus; *VP*: ventral pallidum.

The largest component of the ventral striatum is the nucleus accumbens (NAc), and sometimes these terms are used interchangeably. The NAc has a central portion known as the “core” that is surrounded on its lateral, medial and ventral aspects by another portion known as the “shell”. These two regions differ considerably in terms of their pattern of afferents and efferents and their
Numerous behavioral studies confirm a functional specialization (e.g., Ambroggi, Ghazizadeh, Nicola, & Fields, 2011; Ito, Robbins, & Everitt, 2004; Kelley, 1999; Parkinson, Olmstead, Burns, Robbins, & Everitt, 1999). Both regions receive inputs from the medial prefrontal cortex (mPFC), basolateral amygdala (BLA), hippocampal formation, thalamus and mesencephalic dopaminergic areas, but different parts of these structures project to different parts of the NAc. For example, the prelimbic, anterior cingulate and dorsal agranular insular areas of the prefrontal cortex project predominantly to the core, whereas the infralimbic and ventral agranular cortices project mostly to the shell (Berendse et al., 1992). The segregation of afferents from the hippocampus, amygdala and mesencephalic dopaminergic neurons into different accumbens subterritories is less well defined, but projections from these areas roughly follow a dorsolateral-ventromedial organization (Figure 2; Friedman, Aggleton, & Saunders, 2002; Groenewegen, Wright, Beijer, & Voorn, 2006; Voorn et al., 1989).

In terms of efferents, neurons in the NAc core project mostly to the dorsolateral portion of the ventral pallidum, EPN, STN, SNr and SNC, whereas neurons in the shell project to the dorsomedial ventral pallidum, lateral hypothalamus, lateral preoptic area, substantia innominata, VTA and SNC (Heimer et al., 1991; Usuda, Tanaka, & Chiba, 1998). Based on their different connectivity profiles, the NAc core has traditionally been compared with dorsal striatal structures — which are thought to regulate action selection via control of somatic motor output systems— whereas the NAc shell has been associated with the extended amygdala, a forebrain circuit that is involved in autonomic behaviors via control of visceral and neuroendocrine output systems (Kelley, 1999; Zahm, 2006).
Figure 2. Nucleus accumbens afferents. Structures that project to the core (dark purple) and the shell (light purple) subregions of the NAc. *NAc*: nucleus accumbens; *BLA*: basolateral amygdala; *PFC*: prefrontal cortex; *IL*: infralimbic; *PL*: prelimbic cortex; *ACC*: anterior cingulate cortex; *AI*: dorsal agranular insular cortex; *AIv*: ventral agranular insular cortex; *Pv*: paraventricular thalamic nucleus; *Il*: intralaminar thalamic nucleus; *SN*: substantia nigra; *VTA*: ventral tegmental area; *Glu*: glutamate; *DA*: dopamine.
1.2.2. Identity of neurons in the striatum

1.2.2.1. Medium-sized, densely spiny neurons

Two kinds of neurons populate the striatum: medium spiny neurons (MSNs) and aspiny local interneurons. As their name suggests, MSNs have medium-sized somas (~ 12-14 μm in diameter) and their dendrites exhibit a high density of spines (Ramón y Cajal, 1909; Wilson, Groves, Kitai, & Linder, 1983). MSNs constitute the vast majority of striatal neurons (95% in rodents; the proportion is lower in higher vertebrates, about 77% in primates; Chang, Wilson, & Kitai, 1982; Chang & Kitai, 1985; Graveland & Difiglia, 1985). MSNs are the main recipients of inputs coming into the striatum and also its main output neurons (Gerfen, 1988). In the NAc core, MSNs receive excitatory glutamatergic projections from the mPFC, BLA, ventral hippocampus and medial thalamus (Finch, 1996; Friedman et al., 2002; Groenewegen et al., 2006). MSNs also receive modulatory afferents from the midbrain, mainly from dopaminergic projections from the VTA and SNc and cholinergic projections from local interneurons (for a review, Groenewegen et al., 2006). Finally, MSNs receive inhibitory inputs from GABAergic interneurons, axon collaterals of neighboring MSNs, the VTA and reciprocal GABAergic projections from the VP (Brog, Salyapongse, Deutch, & Zahm, 1993; Groenewegen, Berendse, & Haber, 1993; Tunstall, Oorschot, Kean, & Wickens, 2002; Van Bockstaele & Pickel, 1995; Wilson & Groves, 1980; Wu, Hrycyshyn, & Brudzynski, 1996).

Thanks to their dense dendritic arbor, MSNs have the ideal cytoarchitectonic design for the integration of all these inputs. Neuroanatomical tracing studies have found that axons coming from multiple afferent sources meet onto dendrites of single MSNs in the NAc. For example, single NAc MSNs receive convergent glutamatergic inputs from the BLA and the hippocampus (DeFrance, Marchand, Stanley, Sikes, & Chronister, 1980; French & Totterdell, 2003), as well as convergent inputs from the hippocampus and the prefrontal cortex (PFC; French & Totterdell, 2002). Given the
high degree of convergence of projections from these afferent areas, it is likely that BLA and PFC inputs also converge onto single NAc MSNs. In addition, dopaminergic terminals have also been shown to synapse onto the same NAc MSN dendrites that receive excitatory inputs from each one of these areas (Johnson, Aylward, Hussain, & Totterdell, 1994; Sesack & Pickel, 1990, 1992; Totterdell & Smith, 1989).

Electrophysiological studies suggest that these multiple NAc afferents coordinate to drive activity of single MSNs (Britt et al., 2012; Finch, 1996; Groenewegen et al., 1999; McGinty & Grace, 2009; Mulder, Hodenpijl, & da Silva, 1998; O’Donnell & Grace, 1995). The membrane potential of MSNs oscillates between two states. At rest, MSNs are highly hyperpolarized (~ -85 to -75 mV; Chang & Kitai, 1986; O’Donnell & Grace, 1995; Uchimura, Higashi, & Nishi, 1989), exhibiting a relative insensitivity to small excitatory inputs and low baseline firing rates (“Down state”). Occasionally, MSNs experience a switch to a sustained subthreshold plateau depolarization (~ -60 to -40 mV) during which excitatory inputs are more likely to trigger action potentials (“Up states”; Stern, Kincaid, & Wilson, 1997; Wilson & Kawaguchi, 1996; Wilson & Groves, 1981). “Up state” transitions require convergent excitatory synaptic input (Wilson, 1993).

These large subthreshold membrane fluctuations reveal the complex integrative dynamics of which MSNs are capable. For example, in vivo studies with anesthetized animals show that bursts of excitatory input from the ventral hippocampus are necessary to elicit “Up states” in NAc MSNs, providing a gating mechanism for other weaker excitatory inputs to push MSNs over the spike firing threshold (O’Donnell & Grace, 1995). This seminal study also found that electrical stimulation of amygdalar inputs elicit prolonged depolarizations in MSNs whereas stimulation of cortical afferents yielded fast and transient EPSPs. Optogenetic stimulation of each major NAc afferent confirms that hippocampal inputs evoke the greatest depolarization of NAc neurons (Britt et al., 2012). However, it has been argued that the failure of other projections to bring NAc MSNs closer to “Up states” has
to do with the inability to experimentally replicate natural firing patterns of each specific pathway. For example, PFC neurons tend to fire in bursts. When stimulated in bursts instead of single pulses, PFC neurons were also able to elicit “Up states” in NAc MSNs (Gruber & O’Donnell, 2009). *In vivo* studies with behaving animals suggest that, depending on the task in which the animal is engaged, MSNs in the NAc can rely on different afferent structures to reach the “Up state” (Gruber, Hussain, & O’Donnell, 2009). Pennartz, Groenewegen, & da Silva (1994) estimate that ~14-46 excitatory afferent terminals must release glutamate onto the terminals of an MSN near-simultaneously for the postsynaptic neuron to fire action potentials. Although this is a rough estimate and it overlooks the temporal, spatial and neuromodulatory factors that affect postsynaptic responding, it nonetheless emphasizes the notion that individual excitatory synaptic inputs are unlikely to drive firing of MSNs. Both the cytoarchitectonic and electrophysiological properties of MSNs make them naturally equipped for the integration of a vast array of inputs.

1.2.2.2. *Striatal interneurons*

The other ~5% of cells in the striatum are either cholinergic or GABAergic local interneurons (Kawaguchi, 1993; Tepper, Tecuapetla, Koos, & Ibañez-Sandoval, 2010). Cholinergic interneurons (CINs) have traditionally been referred to as “giant aspiny interneurons” due to their large smooth somas (~ 30-50 μm in diameter) or “tonically active neurons” (TANs) due to their regular spontaneous firing rate (Kawaguchi, 1993). Although they only constitute about 1%-2% of striatal neurons (Matamases, Götz, & Bertran-Gonzalez, 2016), they have a significant influence over striatal function (for a review, Lim, Kang, & McGehee, 2014). They receive sparse innervation, mostly excitatory inputs from the thalamus (Meredith & Wouterlood, 1990) and inhibitory GABAergic inputs from MSNs and other striatal interneurons. They send dense projections within the striatum, and their activity can affect the striatal excitability and plasticity.
either directly (i.e., by releasing acetylcholine onto MSNs) or indirectly (i.e., by regulating the activity of GABAergic interneurons or modulating presynaptic release of dopamine; for a review, Lim et al., 2014). CINs are heterogeneously distributed across the striatum, with the NAc core containing less CINs than any other striatal region, including the NAc shell (Matamales et al., 2016).

The remaining ~ 3%-4% of striatal neurons are a heterogeneous group of GABAergic interneurons (Tepper et al., 2010). The most common GABAergic interneurons are: parvalbumin-expressing fast-spiking interneurons (PV/FSI) and somatostin-expressing interneurons (SOMs). PV/FSIs are perhaps the most studied striatal interneurons. They exhibit low baseline firing rates, but they can emit fast trains of action potentials (~ 400 Hz) when stimulated (Kawaguchi, 1993). They exert strong inhibitory control over MSNs. PV/FSIs have dense collaterals sprouting from the axon, and more than 300 MSNs can be innervated by a single PV/FSI. These dense connections allow PV/FSIs to hyperpolarize entire populations of neighboring MSNs. Recent work has shown that selective inhibition of PV/FSIs leads MSNs to fire more and more asynchronously, confirming the critical role of PV/FSIs in coordinating MSN activity (Owen, Berke, & Kreitzer, 2018). SOMs are medium-sized interneurons and are relatively depolarized at rest, which results in a characteristic a low-threshold for spike output. The axons of these neurons can be long, usually establishing synaptic contact with distant MSNs and CINs (for a review, Castro & Bruchas, 2019). Recent studies have identified new GABAergic interneuron subtypes whose functional role in the intricate striatal circuitry is not well understood yet (for a review, Burke, Rotstein, & Alvarez, 2017).
1.2.3. Dopamine in the nucleus accumbens

1.2.3.1. Striatal expression of dopamine receptor subtypes. Direct and indirect pathways.

The role of the striatum in motivated behavior cannot be understood without examining the ways in which dopaminergic transmission regulates striatal activity. The main source of dopamine in the brain are neurons in two midbrain regions, the SNc and the VTA. Dopamine neurons in the SNc primarily project to the dorsal striatum (“nigrostriatal pathway”). The VTA, as stated above, sends projections to the ventral striatum (“mesolimbic pathway”) but also other areas, such as the amygdala, hippocampus, cortex and thalamus (Sesack & Grace, 2010).

There are five kinds of dopaminergic receptors (D₁-D₅) divided into two families depending on their pharmacological, physiological and molecular properties. D₁ and D₅ receptors are part of the D₁-like family whereas D₂, D₃ and D₄ receptors are part of the D₂-like family (Civelli, Bunzow, & Grandy, 1993). D₁-class receptors activate Gₛ/olf proteins, and D₂-class receptors activate Gᵢ/o proteins. The activation of these different G-protein-coupled receptors has opposite effects on intracellular levels of the second messenger cyclic adenosine monophosphate (cAMP)-dependent protein kinase A (PKA): D₁-class receptors upregulate cAMP and PKA whereas D₂-class receptors have the opposite effect (Missale, Nash, Robinson, Jaber, & Caron, 1998). PKA can act on numerous intracellular targets such as transcription factors and various ion channels that determine the cell’s excitability.

All five kinds of dopaminergic receptors are expressed in the striatum, but D₁ and D₂ are the most common ones. MSNs in the striatum tend to express either D₁ or D₂ receptors (Gerfen et al., 1990; Levey et al., 1993). D₁-expressing MSNs often also express dynorphin and substance P. D₂-expressing MSNs coexpress adenosine 2A receptors and enkephalin (Besson, Graybiel, & Quinn, 1990; Schiffmann & Vanderhaeghen, 1993). Traditionally, these two populations of striatal MSNs have been considered to be highly segregated in terms of their downstream projections: most D₁-
expressing MSNs project to basal ganglia output areas monosynaptically ("direct pathway"), whereas D2-expressing MSNs do so polysynaptically ("indirect pathway"; Albin, Young, & Penney, 1989; Gerfen & Surmeier, 2011). According to this classical view of striatopallidal pathways, dopamine acting on D1-expressing MSNs constitutes a “go” signal whereas activation of D2-expressing MSNs indicates a “no go” signal. In the dorsal striatum, support for this view has come from studies that found that optogenetic activation of D1 and D2-expressing MSNs had opposing effects on basal ganglia output activity, locomotion and reinforcement (Freeze, Kravitz, Hammack, Berke, & Kreitzer, 2013; Kravitz, Tye, & Kreitzer, 2012; Lee et al., 2016). However, other studies have found that both pathways cooperate with each other, rather than oppose each other, to control motor output (Cui et al., 2013; Tecuapetla, Jin, Lima, & Costa, 2016). Manipulation of direct and indirect pathway neurons in the NAc has also challenged the “go”-“no go” hypothesis. Intra-accumbens infusions of both D1 and D2 receptor antagonists reduce conditioned appetitive responding by reducing cue-evoked excitations in NAc core neurons (du Hoffmann & Nicola, 2014). Optogenetic manipulation of both D1 and D2-expressing MSNs suggests that both pathways can promote reward-oriented behavior (Francis et al., 2015; Natsubori et al., 2017; Soares-Cunha, Coimbra, David-Pereira, et al., 2016). Soares-Cunha et al. (2018) propose that this apparently paradoxical effect is due to an increase in NAc dopamine release caused by D2-MSNs disinhibiting VTA neurons via VP projections. Ca\(^{2+}\) imaging studies also contradict the notion of a functional segregation of MSNs in the NAc by showing that D1 and D2-MSNs are co-activated at trial start cue and at first lever press (Natsubori et al., 2017).

Anatomical studies have questioned the degree to which D1 vs. D2-expressing MSNs are segregated into direct vs. indirect pathways, particularly in the NAc. For example, most D1-MSNs in the NAc collateralize to both the VP and the ventral mesencephalon (Pardo-Garcia et al., 2019). Also, some D2-MSNs target VP neurons that project to the thalamus (bypassing the VM), thus
establishing a direct pathway (Kupchik et al., 2015). In addition, many NAc D₁ and D₂-MSNs in the direct and indirect pathways co-express D₃ receptors (Le Moine & Bloch, 1996; Sokoloff, Giros, Martres, Bouthenet, & Schwartz, 1990). Finally, some MSNs express both kinds of receptors, although the exact proportion of striatal MSNs that co-express D₁ and D₂-like receptors is still a topic of debate (Baiezonski, Trifilieff, Meszaros, Javitch, & Kellendonk, 2015; Kupchik et al., 2015; Surmeier, Song, & Yan, 1996). In summary, it is not clear that D₁- and D₂-expressing MSNs in the NAc form fully segregated pathways.

Striatal expression of dopamine receptors is not restricted to the somas of MSNs. D₁ and D₂ receptors can also be found in the terminals of glutamatergic corticoaccumbens projections (Dumartin, Doudnikoff, Gonon, & Bloch, 2007; Wang & Pickel, 2002). In addition, D₂ receptors can act as autoreceptors in the terminals of dopaminergic midbrain neurons innervating the NAc, providing an inhibitory feedback mechanism (i.e., release of dopamine onto striatal cells activates these autoreceptors, which in turn decreases the probability of release upon further stimulation; Limberger, Trout, Kruk, & Starke, 1991). Finally, dopaminergic receptors are also expressed in striatal interneurons. The ubiquity of dopaminergic receptors in these synapses makes it difficult to study the ways in which dopamine regulates striatal function.

1.2.3.2. Behavioral function of NAc dopamine

The role of NAc dopamine has been the subject of a longstanding debate in behavioral neuroscience (Beeler & Dreyer, 2019; Berke, 2018; Salamone & Correa, 2012). The controversy stems, in part, from the fact that motivated behavior is multifaceted, and different lines of evidence link dopamine to different aspects of it. This is illustrated by the diversity of the symptoms associated with abnormal dopaminergic function. Dopamine depletion in Parkinsonian patients leads to a disruption in locomotion (Dauer & Przedborski, 2003). Anhedonia, the inability to find
joy in pleasurable experiences (a common symptom of depression), is also associated with a
downregulation of dopamine (Belujon & Grace, 2017). Drugs of abuse, which increase
dopaminergic transmission in the NAc, are thought to confer excessive motivational value to drug-
paired cues by inducing aberrant learning (Robbins & Everitt, 2002). Attention deficit disorder is
associated with a hypodopaminergic state (LaHoste et al., 1996) whereas schizophrenia (Grace,
2016) and compulsive behaviors (Goodman, McDougle, & Price, 1992) are associated with a
hyperdopaminergic state. Although great progress has been made in the last few decades, the
contribution of dopamine to attention, motivated locomotion and reinforcement learning is still a
subject of contention in the neuroscientific community.

In the 1970s, a seminal study (Wise, Spindler, deWit, & Gerberg, 1978) demonstrated that,
under the effects of dopamine blockers, lever-pressing for food was gradually extinguished in well-
trained rats, just as if the reinforcer had been omitted. This, together with the finding that animals
would vigorously engage in self-stimulation of structures that activate the VTA-NAc pathway
(Wise, 1996), led to the conclusion that this pathway mediates the hedonic impact of reinforcers, the
feeling of pleasure associated with reward consumption. However, animals whose dopaminergic
neurons have been destroyed will still eat food placed in their mouths and their orofacial reactions
suggest that they are capable of enjoying it (Berridge, Venier, & Robinson, 1989). Dopamine is not
necessary for “liking” primary reinforcers (Berridge, 2007). A wealth of studies outlines a much
more complex picture of mesolimbic dopamine function. For example, dopamine release in the NAc
has been linked to the regulation of ongoing motivated behavior. Changes in dopamine
concentration within the NAc change in relation to the value and proximity of rewards (Howe,
Tierney, Sandberg, Phillips, & Graybiel, 2013) and the vigor of reward-seeking movement (Freed
& Yamamoto, 1985; Roitman, Stuber, Phillips, Wightman, & Carelli, 2004). NAc dopamine is
necessary to energize motivated behavior, particularly when it requires energy expenditure (Beeler,
Frazier, & Zhuang, 2012; Salamone, Correa, Farrar, & Mingote, 2007) and/or interrupting ongoing behavior to approach a reward-associated target (Nicola, 2010).

Electrophysiological recordings from midbrain dopamine neurons in behaving animals also challenged the “hedonic theory” of mesolimbic dopamine. Schultz (1998) found that dopamine neurons undergo phasic excitations to rewards and to reward-predictive stimuli, but only when these were presented unexpectedly. Furthermore, these neurons will exhibit phasic inhibitions upon omission of an expected reward. Thus, dopamine signals did not seem to be signaling reward but rather a “reward prediction error” (RPE). Before this finding, learning and artificial intelligence theorists had predicted that reinforcement learning would have to rely on RPE signals (Rescorla & Wagner, 1972; Sutton & Barto, 1981). RPEs are teaching signals, they help organisms update the reward-predictive value of recent events in order to optimize their future behavior. This finding pointed to a likely involvement of dopamine in reinforcement learning, a notion that has been further supported by the fact that striatal dopamine is necessary for inducing plasticity in corticostriatal synapses (Calabresi et al., 2000; Horvitz, 2009; Kerr & Wickens, 2001; Wickens, Begg, & Arbuthnott, 1996; Wickens, Reynolds, & Hyland, 2003; Yagishita et al., 2014). In addition, behavioral pharmacology studies established that disruption of dopaminergic transmission interferes with the acquisition phase of appetitive conditioned behavior (Di Ciano et al., 2001; Eyny & Horvitz, 2003; Hernandez et al., 2005; Horvitz, 2001; Horvitz & Ettenberg, 1988). Finally, optogenetic studies have shown that manipulations of dopaminergic cell firing lead to behavioral changes that are consistent with dopamine acting as an RPE signal (Chang et al., 2016; Hamid et al., 2016; Saunders, Richard, Margolis, & Janak, 2018; Sharpe et al., 2017; Steinberg et al., 2013).

How can dopamine encode ongoing changes in motivational value while keeping tally of discrepancies between the present and the recent past via RPE-like signaling? Some have argued that the firing patterns of dopaminergic cells encode two different signals at different timescales,
with slow changes in tonic firing encoding changes in motivation and RPEs being encoded by rapid phasic changes in firing (Niv, Daw, Joel, & Dayan, 2007; Schultz, 2007). Although levels of dopamine concentration in the NAc ramp up as animals get closer to reward (Howe et al., 2013) and as animals move faster (Freed & Yamamoto, 1985), there is little evidence to suggest that changes in tonic firing rate of dopaminergic neurons ramp up accordingly. Recent work by Joshua Berke’s group has brought some clarity to this issue by comparing the firing patterns of dopaminergic neurons in the VTA with dopaminergic transmission in the NAc of animals engaged in a probabilistic operant task (Hamid et al., 2016; Mohebi et al., 2019). The authors confirmed that RPEs are encoded both by spiking activity of VTA dopaminergic neurons and by subsequent surges in NAc dopamine transmission. However, changes in expectation of reward availability (across or within trials) only covaried with release of dopamine in the NAc, but not with changes in VTA tonic firing rate, which highlights the contribution of local modulation of dopamine release at the terminals to motivational value encoding. The authors suggest that mesolimbic dopamine conveys a single signal that dynamically encodes temporally discounted estimates of future reward. In this context, RPEs can be interpreted as rapid changes in expected value (i.e., cues that predict reward is closer or larger than expected quickly boost this signal). They try to reconcile previous ideas about the role of mesolimbic dopamine function by proposing that this single dopamine signal is a moment-to-moment estimate of “the value of work” (Hamid et al., 2016), therefore guiding decisions about how and when to allocate resources (Berke, 2018).

Dopaminergic transmission in the striatum exerts acute effects on MSN excitability (Gerfen & Surmeier, 2011; Nicola, Surmeier, & Malenka, 2000), which is likely the mechanism via which it affects ongoing motivated behavior (du Hoffmann & Nicola, 2014). Consistent with a role in learning, it also regulates corticostriatal plasticity (Gerfen & Surmeier, 2011; Horvitz, 2009; Nicola et al., 2000; Wickens et al., 2003; Yagishita et al., 2014). Behaviorally, mesolimbic dopamine is
both necessary (Chang, Gardner, Di Tillio, & Schoenbaum, 2017; Di Ciano et al., 2001; Eyny & Horvitz, 2003; Hernandez et al., 2005) and sufficient (Saunders et al., 2018; Sharpe et al., 2017; Steinberg et al., 2013) for learning to predict positive outcomes from previously neutral cues or responses. However, it remains unclear how target neurons in the striatum would be able to decode a motivational and a learning message from one single signal. One candidate mechanism for determining how MSNs interpret dopamine signaling is via local regulation exerted by striatal interneurons. These neurons may provide a gating mechanism over corticostriatal plasticity (Crossley, Horvitz, Balsam, & Ashby, 2016; Morris, Arkadir, Nevet, Vaadia, & Bergman, 2004; Owen et al., 2018).

Finally, it is worth noting that some aspects of dopamine signaling do not fit with encoding of either a motivational or an RPE signal. Activity of VTA neurons has been shown to regulate arousal and sleep-wake cycles (Eban-Rothschild, Rothschild, Giardino, Jones, & de Lecea, 2016). Also, dopaminergic neurons respond phasically to surprising, salient, novel and even aversive stimuli (Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Horvitz, 2000; Pezze & Feldon, 2004; Redgrave & Gurney, 2006). Phasic changes in activity of identified dopaminergic neurons can also encode spontaneous movement in naïve rats (Coddington & Dudman, 2018; Dodson et al., 2016) or initiation/pause of a motor sequence (Jin & Costa, 2010). Furthermore, recent evidence suggests that RPE correlates of dopamine neurons are only observed in well-trained animals learning new associations but not early in training in naïve animals (Coddington & Dudman, 2018). These findings remind us of the complex nature of dopamine’s role in behavior. They highlight the importance of differentiating the activity patterns of specific subpopulations of dopamine neurons, their effect on specific target regions and the need to employ behavioral paradigms that can capture the rich landscape of dopamine-dependent functions.
Dopaminergic modulation of excitability and plasticity in corticostriatal synapses

Dopamine does not drive rapid signaling the way classic excitatory or inhibitory neurotransmitters do (Hernández-López, Bargas, Surmeier, Reyes, & Galarraga, 1997). Instead, it augments or attenuates the response of striatal neurons to excitatory inputs and it modulates their plasticity (Cepeda & Levine, 1998; Reynolds & Wickens, 2002). Thus, dopamine regulates basal ganglia output by providing a gating mechanism for the corticolimbic signals arriving at the striatum (Horvitz, 2002). In vitro studies suggest that dopamine affects the probability that an MSN will fire action potentials in response to glutamatergic input in a non-linear manner, at least via D₁ receptors (Nicola et al., 2000). When MSNs are in the “down” state, D₁ receptor activation promotes resistance to glutamatergic inputs by increasing the activity of inwardly rectifying potassium channels (Calabresi, Mercuri, Stanzione, Stefani, & Bernardi, 1987; Flores-Hernandez et al., 2000). However, when synchronized depolarizing inputs bring MSNs closer to the “up” state, D₁ receptor activation enhances excitability by increasing L-type Ca\(^{2+}\) currents (Hernández-López et al., 1997) and NMDA-evoked excitation (Cepeda, Colwell, Itri, Chandler, & Levine, 1998). Hence, dopamine release amplifies the signal-to-noise ratio in striatal neurons, making sure that only strong convergent inputs get further processed in the circuit (Gruber et al., 2003; Horvitz, 2002; Nicola et al., 2000).

Corticostriatal plasticity requires dopamine in vitro (Calabresi, Maj, Mercuri, & Bernardi, 1992; Reynolds & Wickens, 2002; Schotanus & Chergui, 2008; Yagishita et al., 2014) and in vivo (Fisher et al., 2017). D₁ and D₂ receptors are differentially involved in LTP and long-term depression (LTD), although studies investigating their specific contributions to plasticity have obtained contradicting results. D₁ receptor activation (Reynolds, Hyland, & Wickens, 2001) is necessary for LTP whereas dopamine acting on D₂ receptors (in concert with endocannabinoids) promotes LTD (Kreitzer & Malenka, 2007). Striatal cells tend to express one or the other receptor,
which would mean that most cells are only capable of unidirectional plasticity. This would make it very difficult for striatal cells to reset their synaptic weights. Shen, Flajolet, Greengard, & Surmeier (2008) demonstrated that, in fact, both D₁- and D₂-MSNs are capable of bidirectional Hebbian plasticity and that dopamine is a critical determinant of the direction of that plasticity. Using an STDP paradigm, they found that presynaptic stimulation of cortical afferents followed by postsynaptic firing in D₁-MSNs (“pre-post” protocol) induces robust NMDAR-dependent LTP. In the presence of a D₁ receptor antagonist, the same protocol induces LTD in D₁-MSNs. D₂-MSNs also undergo LTP in a pre-post induction protocol. This is prevented by blockade of D₂ or NMDARs. Post-pre protocols induce LTD in these cells, and that was also disrupted by D₂ receptor blockade. Although these results conflict with an earlier study that found bidirectional plasticity in striatal neurons in the opposite direction (i.e., a post-pre protocol induced LTP and a pre-post protocol induced LTD; Fino, Glowinski, & Venance, 2005), it is consistent with other in vitro (Pawlak & Kerr, 2008) and in vivo (Fisher et al., 2017) studies. The parameters of the plasticity-inducing protocols and the placement of the stimulation electrodes were different in Fino et al., (2005) than in the other studies, which may account for the differences. Importantly, regardless of the contradictions regarding the direction of the effects, all these studies highlight the role of dopamine as a critical modulator of plasticity in the striatum.

1.2.4. The “limbic-motor interface” hypothesis of NAc function

In the 1976 Society for Neuroscience meeting in Toronto, Ann Graybiel observed that, due to its pattern of connections, the NAc was well situated to connect limbic structures with motor output regions. Mogenson, Jones, & Yim (1980) formalized this hypothesis and postulated the NAc as a “limbic-motor interface”. This kind of interface would allow motivational information from the
external environment and internal milieu to gain access to the brain’s motor resources, thus promoting context-appropriate biologically significant actions.

The firing patterns of NAc neurons are consistent with their hypothetical role as limbic-motor translators (Morrison et al., 2017). In well-trained rats, nearly half of NAc neurons exhibit brief excitations and inhibitions upon presentation of reward-predictive cues (Cromwell & Schultz, 2003; Day, Wheeler, Roitman, & Carelli, 2006; Hassani, Cromwell, & Schultz, 2001; Nicola, Yun, Wakabayashi, & Fields, 2004; Schultz, Apicella, Scarnati, & Ljungberg, 1992; Setlow, Schoenbaum, & Gallagher, 2003). These signals are time-locked to the cue and reach their peak within ~100-300 ms of cue onset. Critically, they precede the onset of reward-seeking locomotion (McGinty et al., 2013). Cue-evoked responses in the NAc jointly encode the incentive value of the cues well as spatial and locomotor properties of the subsequent response —likelihood, latency and speed— (Ambroggi et al., 2011; McGinty et al., 2013; Morrison et al., 2017; Taha, Nicola, & Fields, 2007). Disruption of cue-evoked excitations decreases the likelihood and vigor of cued approach responses (Caref & Nicola, 2018; du Hoffmann & Nicola, 2014), which suggests that these neural signals are necessary for vigorous appetitive behavior.

Cue-evoked excitations in the NAc and the subsequent expression of cued approach behavior are dependent on input from the VTA (du Hoffmann & Nicola, 2014; I. A. Yun, Wakabayashi, Fields, & Nicola, 2004), the PFC (Ishikawa, Ambroggi, Nicola, & Fields, 2008) and the BLA (Ambroggi, Ishikawa, Fields, & Nicola, 2008; Jones, Day, Wheeler, & Carelli, 2010). These upstream areas also undergo phasic excitations upon presentation of reward-predictive cues (Jodo, Suzuki, & Kayama, 2000; Paton, Belova, Morrison, & Salzman, 2006; Schoenbaum, Chiba, & Gallagher, 1999; Schultz, 1998). Hence, it is plausible that neurons in the NAc integrate information about the predictive value of cues from upstream cortical and limbic areas and, in turn, set the likelihood and vigor of the approach behavior via downstream projections.
In trained animals, many NAc neurons that are excited upon presentation of the reward-predictive cue remain excited throughout the sequence of behaviors necessary to obtain reward (Gmaz, Carmichael, & van der Meer, 2018). However, once animals enter into the reward receptacle, many NAc neurons exhibit strong inhibitions that span the interval during which the reward is consumed (Day et al., 2006; Janak, Chen, & Caulder, 2004; Nicola, Yun, Wakabayashi, & Fields, 2004b). These inhibitory responses around the time of reward collection have been directly linked to movements involved in reward consumption (Roitman, Wheeler, & Carelli, 2005). Consistent with this, inactivation of the NAc with GABA agonists (Stratford & Kelley, 1997) or AMPA antagonists (Maldonado-Irizarry, Swanson, & Kelley, 1995) promotes consumption, even in subjects that are satiated. Also, electric (Krause, German, Taha, & Fields, 2010) or optogenetic (O’Connor et al., 2015; Prado et al., 2016) activation of NAc neurons arrests consummatory behavior. Thus, NAc neurons not only regulate reward-seeking behavior by promoting approach but also by initiating and maintaining reward consumption. Ca^{2+} imaging of NAc afferents in freely behaving animals has revealed that consumption-related pauses in NAc firing are elicited by a coordinated reduction in the excitatory drive from afferent structures (Reed et al., 2018).

When animals are engaged in tasks that do not impose long time intervals in between trials, lesions of the NAc and manipulations that disrupt dopaminergic transmission in the NAc have minimal effects (Bowman & Brown, 1998; Brown & Bowman, 1995; Robbins, Giardini, Jones, Reading, & Sahakian, 1990). However, the same manipulations impair cued responding in tasks with long intertrial intervals (ITIs; Di Ciano, Cardinal, Cowell, Little, & Everitt, 2001; du Hoffmann & Nicola, 2014; Parkinson et al., 2002; Wakabayashi, Fields, & Nicola, 2004; Yun, Nicola, & Fields, 2004). When intertrial intervals are long, it is harder for animals to predict when cues will be presented. Also, during long intertrial intervals, animals usually disengage from the reward-receptacle and the instrumental operandia and engage in other behaviors (e.g., exploring other parts...
of the chamber, grooming, etc.). When the cue comes on, they need to interrupt those other behaviors and initiate an approach response from the location where they find themselves at that point, which usually varies from trial to trial. These observations led to the proposal of the “flexible approach hypothesis” (Nicola, 2007, 2010). According to this hypothesis, the role of the NAc — and more specifically, NAc dopamine— is to interrupt ongoing behavior and promote approach to a reward-associated target upon presentation of a reward-predictive cue, but specifically when cues are presented at temporally unpredictable times and animals must determine a novel path to reach that target. In brief, intact NAc function is necessary for animals to initiate flexible approach behavior upon presentation of reward-predictive stimuli, which is consistent with the “limbic-motor” hypothesis of NAc function (Mogenson et al., 1980).

1.3. Connecting synaptic plasticity with associative learning

1.3.1. Experience-dependent synaptic plasticity: a historical perspective.

In 1894, Santiago Ramón y Cajal gave a landmark lecture in the Royal Society of London in which he outlined some concepts that would become the cornerstone of modern neuroscience (Ramón y Cajal, 1894). In his lecture, Ramón y Cajal defended the “neuron doctrine”, the idea that neurons do not form a web-like network and instead, are discrete cells separated by a small distance. He also described the delicate architecture of different neuronal types, diagramed pathways and reflected on the dynamic polarization of neurons. However, one of the ideas that attracted most interest was the suggestive speculation that experience had the power to modify the structure of neurons and strengthen their connections, and that this experience-dependent sculpting of the brain might underpin our ability to learn. This idea would not be formalized into a concrete model until about 50 years later. The Canadian neuropsychologist Donald Hebb hypothesized that experience would sometimes result in the repeated and simultaneous activation of groups of neurons. He
thought that this pattern of co-activation would strengthen synapses. In turn, this strengthening would give rise to memories and learning. Hebb’s (1949) famous postulate states the following:

> When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.

Since then, decades of experimental research have confirmed that synapses undergo long-lasting changes in response to experience. This concept, usually referred to as “experience-dependent synaptic plasticity”, is almost universally accepted as the neurobiological basis for learning and memory, although divergent opinions also exist (e.g., Gallistel & Matzel, 2013).

In the late 1960s, Eric Kandel and colleagues embarked in one of the first attempts to experimentally examine synaptic plasticity and its relationship to learning (Kandel & Tauc, 1964; Pinsker, Kupfermann, Castellucci, & Kandel, 1970). They were able to investigate this question by focusing on the *Aplysia californica*, a large sea slug that has a relatively simple and well-understood nervous system. This sea slug exhibits a defensive behavior (the gill withdrawal reflex) when its skin is stimulated, and the strength of the behavior can change with experience. In their experiments, they applied patterns of stimulation that induced enduring changes of the gill withdrawal reflex of the *Aplysia* (e.g., habituation, sensitization) while electrophysiologically recording activity from its sensory and motor neurons and testing the effect of pharmacological manipulations (Kandel, 2001). Kandel and colleagues were able to causally connect changes in the strength of the sensory-motor synapses of this sea slug with its acquired behavior. Even though the changes in the gill withdrawal reflex are far less complex than adaptive behavior in mammals, the contribution of Kandel and colleagues was enormous. To this day, it is arguably the most conclusive evidence of the connection between experience-dependent synaptic plasticity and the acquisition of an adaptive response.
Around the same time period as Kandel was exploring synaptic plasticity in invertebrates, other researchers had started looking for it in the mammalian brain. In 1966, Terje Lømo provided experimental evidence supporting Hebb’s learning rule. He demonstrated that granule cells in the dentate gyrus of the hippocampus of anesthetized rabbits underwent a long-lasting increase in excitability upon presynaptic input from the perforant path as a result of repeated high-frequency stimulation of those afferents (Lømo, 1966). In 1968, in collaboration with Timothy Bliss, he conducted a series of experiments further characterizing this activity-dependent strengthening in synaptic transmission (Bliss & Lømo, 1973). “Long-term potentiation” (LTP), which was the term later coined to refer to that phenomenon (Douglas & Goddard, 1975), became one of the most active fields of research in neuroscience (Nicoll, 2017).

LTP’s usefulness as a cellular mechanism for encoding information would be limited if the opposite process (i.e., experience-dependent weakening of synapses) did not exist. In the 1970’s, LTP’s counterpart, long-term depression (LTD), was discovered in the hippocampus. Lynch, Dunwiddie, & Gribkoff (1977) demonstrated that LTP of an excitatory afferent pathway can depress the postsynaptic cell’s response to a different excitatory afferent pathway (“heterosynaptic LTD”). Soon after, Dunwiddie & Lynch (1978) discovered that “homosynaptic LTD” could also be experimentally obtained by applying prolonged low-frequency (1 Hz) stimulation. Homosynaptic LTD was later identified in other brain regions such as the visual (Kirkwood & Bear, 1994) or somatosensory cortex (Feldman, Nicoll, Malenka, & Isaac, 1998) using the same low-frequency stimulation protocol. Because of its replicability, this LTD-inducing protocol quickly became paradigmatic in the field. However, under natural circumstances, neurons are unlikely to exhibit extended periods of firing at perfectly regular long intervals, which raised some skepticism about the biological relevance of this experimental protocol (e.g., Perrett, Dudek, Eagleman, Montague, & Friedlander, 2001).
Early studies focused on the firing rate of presynaptic inputs and the degree of postsynaptic depolarization as the decisive factors determining the plastic fate of a synapse. High-frequency stimulation paired with strong postsynaptic depolarization would lead to LTP, and low-frequency stimulation together with sustained weak postsynaptic depolarization would result in LTD. Aside from one exception (Levy & Steward, 1983), early plasticity studies did not take into account the precise timing intervals between pre- and postsynaptic spiking activity. In the 1990’s, Markram and colleagues (1997) demonstrated that the sign and magnitude of LTP/LTD depended on the precise order and temporal relationship between pre- and postsynaptic firing activity. The term “spike-timing dependent plasticity” (STDP) was later coined to refer to the strong influence of temporal parameters on plasticity (Song, Miller, & Abbott, 2000). In classic STDP protocols, potentiation of a synapse occurs when presynaptic input (and associated EPSPs) lead to postsynaptic action potentials by up to 20 ms. LTD occurs when the postsynaptic firing precedes presynaptic spikes and associated EPSPs by 20-100 ms. The transition between LTP and LTD is strikingly sharp (~ 5 ms; Bi & Poo, 1998). This kind of canonical STDP is also referred to as “Hebbian STDP”, because it experimentally proves Hebb’s postulate (i.e., inputs that causally lead to postsynaptic firing are strengthened whereas inputs that fail to predict postsynaptic firing are depressed). Hebbian STDP has been found in many excitatory and inhibitory synapses throughout the neocortex as well as in subcortical structures such as the hippocampus or striatum, although other kinds of STDP that do not follow Hebb’s postulate also exist (for a review, Feldman, 2012).

In awake and behaving animals, neurons are constantly receiving inputs from numerous afferent fibers. Also, once fired, postsynaptic action potentials propagate back to the dendritic tree, adding to the excitatory noise in the postsynaptic membrane. Just based on temporal parameters, how do neurons know what specific afferent signals were responsible for specific output activity? The dynamic fate of many synapses has been found to be dictated by a “third factor” (Gerstner,

26
In this kind of plasticity, the first and second factors refer to pre- and postsynaptic activity respectively. The “third factor” is thought to be provided by neuromodulators (e.g., dopamine, acetylcholine, noradrenaline), which are known to signal behaviorally-relevant events such as reward, punishment, surprise or novelty (Horvitz, 2000; Kafkas & Montaldi, 2018; Mirenowicz & Schultz, 1996; Schultz, 1998). In vitro, neuromodulators have been shown to influence plasticity by regulating the number of input-output pairings required to induce plasticity or by adjusting the temporal window during which LTP/LTD can be obtained (for a review, Pawlak, 2010). Nonetheless, in the synapses that are presumably shaped by the release of neuromodulators, input-output excitatory activity is extremely fast (in the order of milliseconds) whereas the time course of neuromodulator release and action is slow (in the order of seconds). Because these factors take place in such different time scales, three-factor learning rules of synaptic plasticity have largely focused on solving this temporal conundrum. Very recent studies have postulated that precisely-timed pre- and postsynaptic spiking sets a synaptic flag called an “eligibility trace”, and that synaptic plasticity only occurs if a neuromodulator is released onto the synapse before that trace fades (e.g. Bittner, Milstein, Grienberger, Romani, & Magee, 2017; Brzosko, Schultz, & Paulsen, 2015; Fisher et al., 2017; He et al., 2015; Shindou, Shindou, Watanabe, & Wickens, 2019; Yagishita et al., 2014). This synaptic flag has not been definitively identified, but it has been suggested that it could be a calcium-based mechanism (Shindou et al., 2019; Yagishita et al., 2014).

In their seminal paper, Bliss & Lømo (1973) raised two questions that have framed the field of neuroplasticity research ever since and that will be addressed in the next two subsections: a) what are the cellular and molecular mechanisms that underlie activity-dependent synaptic strengthening? and b) is synaptic plasticity causally linked to learning and memory in mammals?
1.3.2. Mechanisms: inducing, expressing and maintaining synaptic plasticity.

Although many kinds of activity-dependent plasticity have been identified throughout synapses of the mammalian central nervous system, NMDAR-dependent LTP in hippocampal excitatory synapses is the best understood form of plasticity. Because excitatory corticostriatal synapses are also known to undergo this form of plasticity and that is the focus of this work, this section will concentrate on the mechanisms implicated on this kind of LTP.

1.3.2.1. Induction of LTP: NMDA receptors as coincidence detectors

The synaptic strengthening caused by an LTP-inducing stimulus is orchestrated by the interplay between two kinds of ionotropic glutamatergic receptors: AMPARs and NMDARs. AMPARs have a channel that allows the flow of monovalent cations (Na\(^+\) and K\(^+\)) in and out of the cell. When the neuron is at resting membrane potential, glutamate binding to AMPARs mediates a fast EPSP by allowing an inward current of Na\(^+\), making AMPARs the workhorse of excitatory transmission. NMDARs, in contrast, have a moderate contribution to basal excitatory transmission but are crucial for plasticity, a property that puzzled early researchers of LTP induction mechanisms (Collingridge, Kehl, & McLennan, 1983). NMDARs have two unique properties that explain their role in LTP induction: a) The ion channel of the NMDAR is not only ligand-gated but also voltage-gated, and b) once activated, NMDARs allow the influx of Ca\(^{2+}\) into the cell, and this event initiates a series of intracellular processes that ultimately change the composition of the synapses, altering its strength (Malenka, 1991).

The voltage-dependent behavior of NMDARs stems from their channel being blocked by extracellular magnesium when the cell is at rest. When glutamate acting on non-NMDA receptors depolarizes the neuron, the magnesium block is released from the NMDAR channel, allowing
cations to enter into the postsynaptic cell (Mayer, Westbrook, & Guthrie, 1984). Because activation of NMDARs requires coincident presynaptic release of glutamate and postsynaptic depolarization, they are often referred to as “coincidence detectors”. This unique property of NMDARs can explain the basic properties of LTP induction: associativity and cooperativity. Associativity refers to the finding that an input that is too weak to induce LTP by itself can nonetheless exhibit robust LTP when it is tetanized together with a stronger afferent. Cooperativity refers to the fact that coincident activation of several weak synapses can induce LTP. Because sufficient postsynaptic depolarization is required to activate NMDARs, it makes sense that this would be facilitated by multiple synapses being active at a given time. In conclusion, the unique physiology of NMDARs can account for the Hebbian mechanisms underlying the induction of LTP.

1.3.2.2. Expression of LTP: the role of intracellular calcium and AMPAR trafficking

The major mechanism for LTP expression involves an increase in the availability and responsivity of AMPA receptors in the postsynaptic density (PSD; for a review, Herring & Nicoll, 2016). After induction of LTP, there is a rapid accumulation of GluR1-expressing AMPARs in the PSD that increases the likelihood and magnitude of the postsynaptic response. This process is referred to as “AMPAR trafficking”. These receptors, due to their enhanced conductance and higher permeability to Ca$^{2+}$, make a larger contribution to EPSPs than the AMPARs that were in the PSD prior to the induction of LTP (usually GluR2/GluR3-expressing AMPARs that only allow the influx of monovalent cations). These GluR1-expressing AMPARs come from the cytoplasm and are inserted into the membrane via activity-dependent exocytosis. Interestingly, AMPARs are not directly trafficked into the PSD but, instead, exocytosed at a region of the membrane adjacent to the PSD and later captured by the PSD via lateral diffusion (Lledo, Zhang, Südhof, Malenka, & Nicoll, 1998; Opazo, Sainlos, & Choquet, 2012; Shi et al., 1999). This process is orchestrated by Ca$^{2+}$-
calmodulin-dependent kinase II (CaMKII). When Ca\(^{2+}\) enters the dendrite upon activation of NMDARs, it binds to a Ca\(^{2+}\)-binding protein (“calmodulin”) which in turn binds to and activates CaMKII. CaMKII promotes the anchoring of GluR1-expressing AMPARs into the PSD and it also increases the channel conductance of these receptors once they are inserted into the PSD (for a review, see Lisman, Schulman, & Cline, 2002). AMPARs are thought to be “trapped” into slots in the PSD because they interact with membrane-associated guanylate kinases (MAGUKs), a family of scaffolding proteins that exist in the PSD (e.g., Chen et al., 2015).

Other kinases have been found to play an important role in the early expression of LTP. For example, cAMP-dependent PKA is known to enhance the activity of CaMKII by inhibiting the activity of a competing phosphatase. Also, PKA and protein kinase C (PKC) cooperate with CaMKII to enhance the contribution of AMPARs to synaptic depolarization after induction of LTP (for a review, Rudy, 2013). Although great strides have been made to connect CaMKII to LTP, the precise intracellular mechanisms underlying this connection are still in the process of being characterized (Herring & Nicoll, 2016).

1.3.2.3. Maintenance of L-LTP: protein synthesis and changes in spine morphology

If a strong stimulus is used to elicit LTP, the enhanced synaptic strengthening can last for months in a living animal (e.g., Abraham, Logan, Greenwood, & Dragunow, 2002). This long-lasting form of LTP (L-LTP) does not only involve the postsynaptic mechanisms described above, it also requires the synthesis of new proteins. L-LTP requires the postsynaptic neuron to initiate the synthesis of plasticity-related proteins via transcriptional mechanisms in two different loci: in the dendritic zone at the base of spines and in the nucleus. Local protein synthesis takes place within minutes, and it is made possible by the presence of translational machinery and mRNA in dendrites (Sutton & Schuman, 2005). Transcription in the nucleus is slower (~ 30-40 minutes) and it is
initiated via two kinds of intracellular pathways: synapse-to-nucleus (Lim, Lim, & Ch’ng, 2017) or soma-to-nucleus (Dudek & Fields, 2002) signaling cascades. The endpoint of both kinds of cascades is the phosphorylation of transcription factors such as the CREB protein (Yin & Tully, 1996). When that happens, mRNA is transcribed and later translated into proteins that are crucial for inducing long-lasting changes in synaptic strength. Synapse-to-nucleus signaling cascades involve many different pathways. Most of these pathways engage an array of protein kinases, including PKA (Nguyen & Woo, 2003), which is triggered by cAMP, and extracellular-regulated kinase-mitogen activated protein kinase (ERK-MAPK; Thomas & Huganir, 2004), which is triggered by neurotrophic factors. Soma-to-nucleus signaling that results in protein synthesis occurs when repetitive action potentials open voltage-dependent Ca\(^{2+}\) channels on the cell body. The entry of Ca\(^{2+}\) into the soma triggers a series of intracellular events that also result in the transcription of plasticity-related mRNAs.

Another mechanism via which synaptic strengthening is maintained for long periods of time is the morphological alteration of dendritic spines. A strong LTP-inducing stimulus will change the shape and increase the size of the spine by acting on actin, a cytoskeleton protein filament (Fifková, 1985). Actin is regulated by some of the same mechanisms that promote functional strengthening of the synapse (e.g., NMDAR-mediated influx of Ca\(^{2+}\) into the postsynaptic neuron). Larger spines are more stable than smaller spines, and they also contain more AMPARs, which makes them more sensitive to presynaptic glutamate release (for a review, Yuste & Bonhoeffer, 2001). Additionally, the induction of L-LTP is also associated with the strengthening of the connection between the modified spine and its presynaptic partner. Neural cadherins are cell adhesion molecules that form strands of proteins held together by Ca\(^{2+}\) cations. These strands build strong bridges that hold the pre- and postsynaptic membranes in apposition, promoting the maintenance of a stable synaptic relationship between them (Bozdagi, Shan, Tanaka, Benson, & Huntley, 2000).
1.3.3. Are plasticity and learning causally linked?

It is commonly assumed that changes in synaptic strength constitute the mechanism via which experiences are encoded and stored in the brain. This is known as the “synaptic plasticity and memory (SPM)” hypothesis. Based on their properties alone, LTP and LTD are likely physiological substrates for learning and memory. However, establishing a causal link between neuroplasticity and learning in freely moving vertebrates has proven to be a challenging task. Martin, Grimwood, & Morris (2000) established an epistemological framework for rigorously testing the SPM hypothesis. According to these authors, empirical support for the SPM hypothesis has to satisfy four logical criteria: detectability, anterograde alteration, retrograde alteration, and mimicry. Each one of the following subsections includes a description of each criterion and of methodological approaches commonly used to test each one of them.

1.3.3.1. Detectability

If the formation of a memory can be explained by LTP/LTD, the experience that gave rise to that memory should also induce detectable changes in synaptic efficacy somewhere in the brain. Even though Martin et al. (2000) do not explicitly state this, detectability is naturally subjected to a temporal requisite. Logic dictates that a cause must temporally precede an effect, and thus only synaptic changes that are detected before (or in association with) the induction of memory can possibly be causally linked to it.

A wide range of ex vivo and in vivo techniques can be used to detect changes in synaptic strength associated with the emergence of a specific memory. In ex vivo protocols, animals are trained to learn something and, after their brains have been removed and sliced, the strength of the synapse that is suspected to underlie the memory is compared to the strength of the same synapse in control animals. Synaptic strength can be measured using electrophysiological recordings of the
postsynaptic response to an experimenter-controlled presynaptic input. Input activity is elicited, for example, by administering an electric pulse to afferents or by using light to optogenetically activate presynaptic terminals expressing light-sensitive ion channels or to uncage glutamate in the synapse. In this kind of protocol, the strength of the postsynaptic response is quantified based on indices such as the relative amplitude of synaptic currents mediated by AMPAR vs. NMDA receptors, the EPSP amplitude/slope or the threshold to spike. In vivo studies in awake animals provide real-time estimates of synaptic changes associated with the acquisition of a new memory or behavior. Probing the strength of a synapse in vivo often involves the use of electrophysiological recordings to monitor changes in the postsynaptic response (i.e., spiking activity or local field potentials) to a controlled input —either a sensory stimulus or the presynaptic pathway whose activity presumably represents that stimulus—. Another detectable symptom of experience-dependent synaptic strengthening is the escalation of input-output activity coherence after variable amounts of behavioral training.

Because induction of LTP implicates intracellular signaling cascades that promote protein synthesis, another way of detecting synaptic plasticity is monitoring the expression of plasticity-evoked genes. Some of these genes, such as c-fos, egr-1 or Arc (“immediate-early genes”, IEGs), are upregulated rapidly after strong neuronal activation. Therefore, some studies have monitored the expression of these genes during behavioral training as a proxy for changes in neuronal responsivity that would be expected to be found in a synapse after induction of plasticity (Minatohara, Akiyoshi, & Okuno, 2016).

An increasingly popular alternative to electrophysiology for monitoring neuronal activity involves the use of genetically encoded Ca$^{2+}$ indicators (GECIs). GECIs are fluorescence-based sensors can be transfected to specific cell types, populations or subcellular locations for long-term measurement of intracellular Ca$^{2+}$ dynamics (Tian, Hires, & Looger, 2012). When intracellular levels of Ca$^{2+}$ change, GECIs alter their levels of fluorescence. Because action potentials elicit a
localized sharp increase in intracellular Ca\textsuperscript{2+} levels, fluctuations in fluorescence elicited by Ca\textsuperscript{2+} transients are a good proxy for changes in neuronal activity (Girven & Sparta, 2017). Although electrophysiology offers a higher temporal resolution as well as a more comprehensive and direct characterization of neuronal activity dynamics, Ca\textsuperscript{2+} imaging offers some advantages. Similarly to other techniques that use targeted viral-assisted transgene expression, Ca\textsuperscript{2+} imaging can be used to obtain cell- and pathway-specific recordings. In addition, the same neurons can be tracked over time, something that becomes useful when studying learning-related changes in neuronal activity.

Finally, because synaptic potentiation not only involves changes in function but also in structure, experience-dependent changes in spine volume and morphology have also been used to detect learning-related synaptic plasticity (Gipson & Olive, 2017).

1.3.3.2. Anterograde alteration

This criterion posits that manipulations that interfere with mechanisms that induce synaptic changes during a learning experience should disrupt the memory of that experience. Anterograde alteration has often been implemented experimentally by using behavioral pharmacology techniques. In these experiments, animals are trained while under the effects of drugs that disrupt plasticity-inducing mechanisms (i.e., NMDAR antagonists or agents that interfere with the intracellular consequences of NMDAR activation, such as Ca\textsuperscript{2+} chelators, CaMKII inhibitors, protein synthesis inhibitors, etc.; e.g., Kelley, Smith-Roe, & Holahan, 1997; Morris, Anderson, Lynch, & Baudry, 1986; Schafe & LeDoux, 2000). Other approaches include using NMDAR-knockout animals (e.g., Koralek, Jin, Long Li, Costa, & Carmena, 2012; Tsien, Huerta, & Tonegawa, 1996), saturating LTP to occlude subsequent learning (e.g., Moser, Krobert, Moser, & Morris, 1998) or manipulating neuronal activity at synapses where plasticity is expected to be induced by experience using optogenetics or chemogenetics (e.g., Kakegawa et al., 2018; Tipps,
Marron Fernandez de Velasco, Schaeffer, & Wickman, 2018). A challenge that these studies face is that all these manipulations may alter behavior by interfering with aspects of synaptic transmission that are unrelated to plasticity. Thus, experiments that use this kind of approach must ensure that the effects of the treatment on learning cannot be explained by a deficit in performance-related factors (e.g., perceptive or motor skills, motivation for rewards, arousal, expression of exploratory or learned behavior, etc.).

1.3.3.3. Retrograde alteration

If memory is the result of synaptic changes induced by a prior experience, interventions that alter the mechanisms involved in the expression of those changes should alter the memory of that episode. The techniques used to test the effect of retrograde alteration on memory retrieval overlap with those used to induce anterograde alteration (the difference, of course, is that these treatments are applied after learning has taken place). Before an experience takes place, it is hard to pinpoint what cells will be allocated to it, in what specific ways that experience will modify those cells and when will those changes happen. Therefore, achieving anatomical, temporal and mechanistic specificity is more challenging in studies that use anterograde alteration in naïve animals than in studies that use retrograde alteration in already-trained animals. Nabavi et al. (2014) provide an example of the possibilities of retrograde alteration approaches by achieving bidirectional mechanistic control over a fear memory once it had already been formed. These authors trained animals to associate a foot shock with optical stimulation of auditory inputs targeting the lateral amygdala. After animals had learned to freeze upon presentation of the optical conditioned stimulus, the authors were able to erase the memory by optogenetically inducing LTD in the potentiated synapse and later reinstate the memory by inducing LTP in the same synapse.

The “detectability” criterion requires finding a correlation between plasticity and memory,
which is a prerequisite for possibly establishing a causal relationship between the two, but causality is not implied by it. Detection of a plasticity-memory correlation informs us of where to look for learning-related neural hotspots. To declare that neuroplasticity is causally linked to the formation of a specific memory, the former has to be both necessary and sufficient for the latter to occur. Anterograde and retrograde alteration help us establish necessity. Establishing sufficiency is more challenging methodologically because it requires inducing artificial memories by artificially replicating the changes that occur in the brain during natural learning. Although this possibility seemed far-fetched just a few years ago, the development of techniques that allow activity-dependent and pathway- and cell-specific manipulation of neuronal activity is making this increasingly feasible.

1.3.3.4. Mimicry

In the same way that engineering and technological creations validate our hypotheses about how the physical world works, being able to artificially generate a memory by inducing LTP/LTD in the right synapses constitutes a critical test of the SPM hypothesis. The difficulty in this kind of experiment lies in making sure that the artificial stimuli adequately simulate patterns of activity elicited by natural experiences. An ingenious approach is that of optogenetics experiments that rely on activity-dependent expression of light-gated ion channels (e.g., by infusion of cre-dependent opsins into a cfos-cre line of animals). For example, Ramirez et al. (2013) were able to express the excitatory opsin ChR2 in the dentate gyrus but only in neurons that were activated during exploration of context A. Later, in an alternative context B, animals were exposed to foot shock while the neurons that encoded context A were being photostimulated. When animals were subsequently placed in context A —but not in another neutral context C—, they displayed a freezing response, even though they had never experienced a foot shock in context A. Similarly,
animals learned to fear a neutral context after concurrent optogenetic stimulation of c-fos expressing neurons in the BLA (associated with a previous experience of foot-shock) with c-fos expressing hippocampal neurons encoding the neutral context (Ohkawa et al., 2015).

Because that kind of study requires previous experience with the to-be-associated stimuli in order to identify the neurons that encode those stimuli, it could be argued that the implanted memory is not completely artificial. In a very recent study, Vetere et al. (2019) were able to engineer a fully artificial memory in mice, bypassing sensory experience altogether, by optogenetically stimulating pathways that encode sensory and motivational information. They were able to do so thanks to the segregated topography of the olfactory system. Individual olfactory sensory neurons have receptors for specific odors, and neurons expressing the same olfactory receptor converge downstream onto a particular olfactory glomerulus. The authors expressed ChR2 in the glomerulus involved in encoding a particular odorant (acetophenone) but not a different odorant that they used as a control (carvone). They used photostimulation of the acetophenone-encoding glomerulus as the conditioned stimulus (CS). The unconditioned stimulus (US) was also replaced by photostimulation. In this case, activation of inputs from the lateral habenula to the VTA were used to mimic an aversive US and activation of inputs from the laterodorsal tegmental nucleus to the VTA were used to mimic an appetitive US. When the real odors were presented, animals avoided or approached the smells depending on whether the aversive or the appetitive pathway had been stimulated in association with the artificial CS during training, even though they had never smelled those odors before.

Experience-dependent plasticity in some brain structures has been linked to different forms of learning by evidence that satisfies each one of the four criteria. In particular, great strides have been made to connect NMDAR-dependent LTP in the hippocampus with spatial learning, in the amygdala with fear conditioning and in the dorsal striatum with motor skill acquisition (Bliss,
In contrast, the connection between neuroplasticity in the nucleus accumbens and acquisition of cued reward-seeking behavior has been poorly tested—despite this being commonly invoked mantra in addiction research (Kauer & Malenka, 2007)—. In section “1.4.2.” the SPM hypothesis linking appetitive conditioned behavior and plasticity in the nucleus accumbens will be assessed using the four criteria laid out in this section.

1.4. Plasticity in the nucleus accumbens and appetitive learning

1.4.1. Artificial induction of synaptic plasticity in the NAc

Synapses between glutamatergic afferents and MSNs in the NAc can undergo LTP and LTD. LTP can be induced in vitro and in vivo in synapses formed by afferents from the BLA (Floresco, Blaha, Yang, & Phillips, 2001; Uno & Ozawa, 1991), the hippocampus (LeGates et al., 2018; Mulder, Arts, & da Silva, 1997) and the cortex (Goto & Grace, 2005; Pennartz, Ameerun, Groenewegen, & Lopes da Silva, 1993). LTP in these synapses has been induced employing a variety of protocols: high frequency stimulation (Pascoli, Turiault, & Lüscher, 2012), low frequency stimulation paired with moderate postsynaptic depolarization (Kombian & Malenka, 1994) or STDP protocols (Ji & Martin, 2012; Popescu et al., 2007; Yagishita et al., 2014). Both D1-MSNs and D2-MSNs can undergo LTP (Pascoli et al., 2012).

LTP in the NAc depends on NMDARs and on the intracellular events triggered by their activation, such as increases in Ca^{2+} concentration, CaMKII activation, ERK signaling and protein synthesis (Floresco et al., 2001; Ji & Martin, 2012; LeGates et al., 2018; Mazzucchelli et al., 2002; Pascoli et al., 2012; Yagishita et al., 2014). Several studies also suggest that LTP in the NAc depends on D1 receptors and PKA but not on D2 receptors (Floresco et al., 2001; Kombian &
LTD can also be induced in the NAc using different protocols. High frequency stimulation paradigms can induce non-NMDAR-dependent LTD in a small percentage of synapses (Pennartz et al., 1993). In addition, although tetanic stimulation enhances the AMPAR-mediated component of the synaptic response in the NAc, it also induces a decrement of the NMDAR-mediated component (Kombian & Malenka, 1994). Also, while tetanic stimulation at cortical afferents induces LTP, it simultaneously induces heterosynaptic LTD at hippocampal synapses onto the NAc and vice versa (Goto & Grace, 2005). Low frequency stimulation of cortical afferents for 5-10 minutes can also depotentiate a synapse that had previously undergone LTP (Pascoli et al., 2012) or elicit endocannabinoid-dependent LTD (Robbe, Kopf, Remaury, Bockaert, & Manzoni, 2002). In summary, synapses between excitatory inputs and MSNs in the NAc are highly plastic, with NMDARs and endocannabinoid receptors mediating LTP and LTD respectively.

1.4.2. Connecting plasticity in the NAc with learning

Demonstrating that excitatory synapses in the NAc are capable of long-lasting changes in strength does not establish the functional relevance of this dynamic property. Establishing a link between plasticity in the NAc and acquisition of appetitive conditioned responses requires: a) detecting changes in synaptic efficacy in vivo during natural learning, and b) demonstrating that those changes are necessary and sufficient for learning. Rigorously answering this question requires gathering empirical evidence that meets the detectability, anterograde alteration, retrograde alteration, and mimicry criteria described above (section “1.3.3.”). This involves tremendous scientific efforts, and whether it is worthwhile to engage in such an endeavor has been questioned (Malenka & Bear, 2004):
We would also argue that it is no longer particularly productive to debate the generic question of whether LTP and LTD are cellular/synaptic mechanisms for memory. LTP and LTD are experimental phenomena, which can be used to demonstrate the repertoire of long-lasting modifications of which individual synapses are capable. It is a daunting task to demonstrate that identical modifications due to the same mechanisms underlying some form of LTP or LTD occur in vivo in response to experience. It is even more difficult to prove that these LTP or LTD-like modifications subserve essential functional roles (p. 5).

However, optimism is not unwarranted. Great progress has already been made to causally connect experience-dependent plasticity in other brain regions with other kinds of learning — namely, amygdala/fear learning (Luchkina & Bolshakov, 2019), dorsal striatum/procedural learning (Perrin & Venance, 2019) and hippocampus/spatial and episodic learning (Bliss et al., 2018; Takeuchi, Duszkiewicz, & Morris, 2014) —. In contrast, however, efforts to connect experience-dependent changes in NAc synapses with acquisition of conditioned reward-seeking behaviors have been limited. In the case of instrumental learning, it has been argued that this connection does not exist (Yin, Ostlund, & Balleine, 2008).

Despite the lack of conclusive experimental evidence, the idea that NAc plasticity plays a causal role in natural learning of cued reward-seeking behavior is a pillar of addiction research (Hyman, 2005; Hyman, Malenka, & Nestler, 2006). Environmental stimuli associated with the consumption of drugs induce cravings and promote relapse, even after years of abstinence (Self & Nestler, 1998; Stewart, 2008). Also, drugs of abuse induce synaptic plasticity in synapses at the VTA and NAc (Kauer & Malenka, 2007; Pierce, Bell, Duffy, & Kalivas, 1996; Russo et al., 2010). Modern theories of addiction consider that compulsive drug-seeking behaviors are a form of aberrant associative learning caused by drug-induced plasticity mechanisms in the NAc (Di Chiara
et al., 2006; Hyman, 2005; Kalivas, Volkow, & Seamans, 2005; Robbins & Everitt, 2002; Robinson & Berridge, 1993; Torregrossa, Corlett, & Taylor, 2011). Hence, these theories implicitly assume that natural appetitive conditioning takes place because primary rewards induce neuroadaptations in the NAc. Consequently, investigating the natural functions of plasticity in the NAc is important not only to determine how our brains promote adaptive responses, but also to interpret addiction research and understand how drugs of abuse induce addictive behaviors.

1.4.2.1. Detectability: does appetitive conditioning induce detectable changes in NAc neuronal responses to conditioned cues?

A few studies have identified changes in NAc firing patterns during natural learning. Roitman et al. (2005) recorded activity of NAc neurons while training naïve animals in a purely Pavlovian design. In this task, two different audiovisual cues were paired with intraoral delivery of either a rewarding (sucrose) or aversive (quinine) solution. The authors found that, although very few NAc neurons exhibited innate responses upon presentation of neutral audiovisual stimuli at the beginning of training, those responses emerged quickly (within the first 10 trials) as animals learned to anticipate the upcoming outcomes associated with those stimuli.

Setlow et al. (2003) conducted a study along the same lines, although in this case, both cues were olfactory and the delivery of either sucrose or quinine was contingent upon entry into a fluid well. These authors also found that neurons in the NAc come to encode the motivational value of cues after repeatedly experiencing the cue-outcome contingency. Interestingly, based on their learning-related firing patterns, two subpopulations of NAc neurons were identified in this study. One subpopulation of neurons (“rapidly selective”) came to encode the learned motivational value of the cues very early in training, during sessions in which animals still exhibited poor behavioral performance. Another subpopulation of neurons (“slowly selective”) developed selective responses
to the cues only during later phases of training, when the animal’s behavior was highly accurate and vigorous. During a reversal phase, about half of the rapidly selective neurons reversed selectivity whereas ~80% of slowly selective neurons lost cue-evoked responses altogether. Setlow et al. (2003) conclude that rapidly selective neurons only encode the motivational value of the cue whereas the slowly selective neurons encode particular combinations of motivational value and motor responses (“go/ no go”). On the grounds of the baseline firing rate of both subpopulations of neurons, the authors speculate that perhaps rapidly selective neurons are interneurons whereas slowly selective neurons are MSNs. The possibility of different subtypes of neurons in the NAc encoding different information during natural learning was systematically explored by Atallah, McCool, Howe, & Graybiel, (2014). In this study, authors recorded from cells identified as interneurons and MSNs in the ventromedial striatum as animals were trained in a T-maze task. Atallah et al. (2014) found that MSNs quickly develop phasic increases in activity to a cue that indicates the beginning of a trial. This finding challenges the hypothesis that neurons that slowly acquire responsivity to learned cues are MSNs (Setlow et al., 2003). These results suggest that NAc neurons, probably regardless of neuronal subtype, acquire phasic cue-evoked responses at some point during training.

If an excitatory synapse is strengthened due to experience, we would also expect to observe an increase in the coherence of pre-postsynaptic activity during learning in this synapse. This is what Popescu, Popa, & Paré (2009) found in the BLA-to-NAc synapse of cats as they learned the association between an auditory cue and the availability of reward. This coupling was reduced by infusions of a GABA agonist. Additionally, they did not find correlated activity between other structures (i.e., auditory cortex or intralaminar thalamus) and the NAc. In accordance with this, the AMPA/NMDA receptor ratio increases after the acquisition of a cued approach behavior in the BLA-to-NAc synapse (Namburi et al., 2015). These results situate the BLA-to-NAc synapse as a
likely locus of plasticity underlying the acquisition of cued appetitive responses.

Because the induction of plasticity is thought to elicit gene expression and protein synthesis, some studies have measured the presence of the immediate early gene c-fos to identify synapses where plasticity may have emerged as a result of experience. For example, after training in a Pavlovian cued approach task, animals that failed to learn (due to NMDAR blockade in the VTA) showed much lower levels of c-fos in the NAc than vehicle animals that did learn the task (Ranaldi et al., 2011). Human brain studies have also identified learning-related changes in NAc activity during associative learning. Investigations that measure blood oxygenation levels using fMRI have identified changes in activity in the ventral striatum that are consistent with with learned anticipation of outcome upon presentation of reward-predictive cues (Diekhof, Kaps, Falkai, & Gruber, 2012; Knutson, Adams, Fong, & Hommer, 2001; O’Doherty et al., 2004; O’Doherty, Dayan, Friston, Critchley, & Dolan, 2003).

Changes in NAc dopamine release during learning are not a direct measure of putative plasticity in the NAc. However, due to the critical role of dopamine in corticostriatal plasticity (see sections “1.2.3.3.” and “1.4.1.”), detecting changes in NAc dopamine release during training can shed light on potential mechanisms underlying learning-related changes in NAc activity. Day, Roitman, Wightman, & Carelli (2007) found that, early in conditioning, increases in dopamine release in the NAc were elicited by delivery of reward, but not by the presentation of conditioned stimuli. After several cue-reward presentations, increases in NAc dopamine release shifted to the conditioned stimulus and they were no longer elicited by the reward. Similarly, during acquisition of an instrumental action sequence, peak levels of NAc dopamine concentration first took place after reward delivery and, with training, they gradually occurred earlier during action sequence (Collins et al., 2016). Taken together, these data suggest that appetitive conditioning is associated with an increase in the encoding of reward-predictive cues in the NAc.
1.4.2.2. Anterograde and retrograde alteration

Anterograde alteration: does the interference with the induction of NAc plasticity disrupt aspects of appetitive learning?

The earliest attempts to investigate the mnemonic effects of NAc disruption during training consisted of lesion studies. Electrolytic or excitotoxic lesions of the NAc core disrupt learning in a T-maze or a Morris water maze in which rats have to learn to locate a hidden platform relying on external cues (Annett, McGregor, & Robbins, 1989; Sutherland & Rodriguez, 1989). NAc lesions also impair the acquisition of an autoshaped appetitive approach to a visual stimulus (Parkinson, Willoughby, Robbins, & Everitt, 2000) or a lever (Chang, Wheeler, & Holland, 2012). Pavlovian-instrumental transfer (PIT), a phenomenon that reveals the ability of Pavlovian stimuli to enhance instrumental responding, is affected by lesion or pharmacological inactivation of the NAc core (Corbit & Balleine, 2011). The timing of reward prediction errors signaled by VTA neurons is also disrupted by NAc lesions (Takahashi, Langdon, Niv, & Schoenbaum, 2016), indicating at least one possible indirect mechanism via which the NAc is implicated in reinforcement learning.

Regarding the effect of NAc lesions on instrumental learning, there have been some seemingly contradictory findings. NAc lesioned monkeys were able to learn an fixed ratio (FR)-15 button-pressing task as quickly and accurately as control animals (Stern & Passingham, 1996), although their rate of responding declined faster during an extinction session. Cardinal & Cheung (2005) found that excitotoxic lesions of the NAc core retarded the acquisition of an FR-1 lever-pressing response, but only when there was a temporal delay between the operant response and the delivery of the reward. In a two-armed bandit task, monkeys with NAc lesions exhibited impaired learning of a stimulus-reward contingency, but only in when there is some uncertainty in the predictive value of the cues (i.e., in non-deterministic environments; Costa, Dal Monte, Lucas,
Murray, & Averbeck, 2016). This may help explain the lack of effect in the study by Stern & Passingham (1996), in which the operant response was rewarded in a deterministic and instant manner. In any case, because NAc neurons can exert opposite influences on motor output regions based on their pattern of downstream projections (section “1.2.1”), the effects of lesion and inactivation studies are hard to interpret.

Behavioral pharmacology studies have perhaps been the most productive avenue of research in the investigation of NAc-specific mechanisms implicated in reward-related learning. Because induction of plasticity in the NAc is known to depend on NMDA receptors (section “1.4.1.”), most behavioral pharmacology studies have focused on the contribution of these receptors to the acquisition phase of appetitive learning. Maldonado-Irizarry & Kelley (1995) trained animals in a food-search task in which they had to remember the location of four pellets across four training days (5 trials per day). The chamber had a 16-hole grid on the floor and the pellets were always placed in the same holes for each animal. Compared to vehicle infusions, infusions of the NMDAR antagonist AP5 into the NAc core or shell prior to each training session markedly impaired learning. AP5-treated rats made more errors (entries into the empty holes) and took significantly longer to find all the pellets on each trial. This was not due to a reduction in locomotion, animals just engaged in a more erratic and inefficient search pattern. Blockade of AMPA and kainate receptors in the NAc core (but not the shell) also impaired learning in this task but the effect was less pronounced. Similarly, AP5 infusions into the NAc core impaired learning in a task in which rats had to remember, across sessions, which four arms of an 8-arm radial maze were baited (Smith-Roe, Sadeghian, & Kelley, 1999). Acquisition of an appetitive Pavlovian response in an autoshaping paradigm is also disrupted by NMDAR blockade during training (Di Ciano et al., 2001; Kelley et al., 1997). Instrumental learning is impaired by pretraining infusions of an NMDAR antagonist into the NAc core too. For example, animals under the effects of intra-accumbens AP5 failed to learn
what arm to choose in a Y-maze upon perceiving discriminative olfactory stimuli that were acting as instrumental cues (Atallah, Lopez-Paniagua, Rudy, & O’Reilly, 2007). In more standard instrumental lever-pressing paradigm, pre-session infusions of AP5 into the NAc core during training also disrupted the acquisition of the operant response (Hernandez et al., 2005; Kelley et al., 1997), consistent with the anterograde alteration criterion.

Although these studies suggest that NAc core NMDARs are necessary for reward-related learning, their results should be interpreted with caution. AP5 treatment in the NAc core does not affect motivation for food or inhibit motor output (Kelley et al., 1997) –if anything, it increases locomotion (Burns, Everitt, Kelley, & Robbins, 1994) –. Nevertheless, critically, it affects exploratory behaviors that might be crucial for bringing the animal in contact with important environmental stimuli. Maldonado-Irizarry & Kelley (1994) found that AP5 infused into the NAc core renders naïve rats less likely to explore the periphery of an open field or interact with novel objects. In Hernandez et al. (2005), AP5-treated animals were not only less likely to learn to lever-press for reward but they were also less likely to check the reward compartment, even when free rewards were audibly delivered every ~15-30 s at the beginning of training. If an AP5-induced performance deficit precludes animals from experiencing contingencies that are necessary to learn the associative structure of their environment, a learning deficit might just be a side effect of a more general behavioral deficit. Some of the studies mentioned in the previous paragraph addressed this possibility by administering infusions of AP5 into the NAc core after animals had already acquired to Pavlovian or operant response (Di Ciano et al., 2001; Hernandez et al., 2005; Kelley et al., 1997). Because these infusions failed to suppress the learned conditioned response, the observed AP5-induced learning impairment has been interpreted as a purely associative deficit. However, these studies gave these test AP5 infusions after animals had received extended training. Intra-accumbens core infusions of AP5 in animals that have only been moderately trained do indeed disrupt
expression of a previously learned behavior by either increasing the number of errors (Smith-Roe et al., 1999), the latency to respond (Atallah et al., 2007) or both (Maldonado-Irizarry & Kelley, 1995).

It is challenging to provide an exact definition of what constitutes moderate vs. extended training, particularly given the diversity of paradigms and the variability of learning curves across animals (Gallistel, Fairhurst, & Balsam, 2004). In this text, the definition will be based on the amount of training animals receive between the end of the training period during which AP5 disrupts acquisition of a conditioned response and the administration of a drug test to probe the impact of AP5 on the expression of that same response. For example, in Maldonado-Irizarry & Kelley (1995), AP5 was found to disrupt learning in a food-search paradigm when administered on days 1-4 of training. When a different group of animals received intra-accumbens AP5 infusions on the fifth day (“moderate training”), expression of the previously learned food-search pattern was found to be disrupted by NMDAR blockade. In contrast, AP5 disrupted the emergence of an autoshaped response on days 1-6 of training, but it failed to alter its expression when administered after ~17 days of training (“extended training”; Di Ciano et al., 2001). It is also worth noting that under some conditions, blockade of NAc NMDARs in trained animals can affect expression of learning even after extended training. For example, after eight weeks of training, intra-accumbens infusions of AP5 prevent the expected reward magnitude associated with a cue from determining the vigor of a conditioned response (Hauber, Bohn, & Giertler, 2000). Thus, NAc core NMDARs are necessary for normal expression of a learned response after moderate (and sometimes after extended) training.

The fact that NMDARs in the NAc core are involved in the expression of conditioned reward-seeking behaviors does not necessarily contradict their contribution to learning. It just complicates the interpretation of an AP5-induced learning impairment. However, some studies have
managed to disrupt learning with manipulations that have minimal or no effects on performance. For example, because artificial induction of plasticity into the NAc is thought to require the synergistic activation of D₁ and NMDA receptors, Smith-Roe & Kelley (2000) hypothesized that combined D₁ and NMDA receptor antagonists should disrupt learning even at very low doses. They administered a mixture of an NMDA and a D₁ receptor antagonist (AP5/SCH23390) at doses that were 10-fold lower than those necessary to impair learning when administered independently. This combination of low dose drugs did not affect exploratory or feeding behavior but it did strongly disrupt acquisition of an instrumental lever-pressing response. Both drugs failed to impair learning when administered separately, suggesting that coactivation of NMDA and D₁ receptors during training are necessary for natural reinforcement learning. A caveat of this study is that the administration of the low dose AP5/SCH23390 cocktail had a mild disruptive effect on lever pressing, even after extended training. This, again, raises the possibility that the effect of low dose AP5/SCH23390 on learning is at least partially explained by a performance deficit (i.e., it is possible that, earlier in training, infusions of even a low dose AP5/SCH23390 solution affect behavior and deprive animals from experiencing the task-related contingencies they are supposed to learn).

An alternative approach to avoid the performance confound is to administer the drug after the training session(s). A learning impairment caused by post-training infusions is not likely to be explained by a disruption in exploratory or motor performance, because animals can engage in the task drug-free. While a lack of effect would indicate that the targeted mechanism is not necessary for the consolidation of a particular kind of learning, it could still be necessary for encoding the memory at the precise moment the experience takes place and for initiating plasticity mechanisms. Bilateral infusions of NMDAR antagonists into the NAc core immediately after training (but not 2 h later) disrupt learning in tasks with a strong spatial component (Roullet, Sargolini, Oliverio, &
Posttraining infusions of AP5 into the NAc core also prevented acquisition of an autoshaped approach response to a location where an LCD screen signaled upcoming delivery of reward (Dalley et al., 2005). On the other hand, post-training infusions of AP5 did not disrupt acquisition of an operant lever-pressing response (Hernandez et al., 2005). Learning to press a lever requires repeated pairings of specific motor patterns followed by reward within very tight temporal parameters. However, this is probably not the case in a setting in which what animals are learning might be the reward-predictive value of specific locations in an arena. This may explain why in instrumental protocols (and not in protocols with a strong spatial component) NMDARs mediate plasticity while animals are engaged in the task but not afterwards.

NMDARs are known to contribute to the induction of plasticity by initiating a series of intracellular events that result in protein synthesis and strengthen the efficacy of that synapse (see section “1.3.2.”). Some studies have tested the effects of inhibiting some of these second messengers or protein synthesis during learning. For example, Baldwin, Sadeghian, Holahan, & Kelley (2002) gave intra-accumbens infusions of a broad-based serine/threonine kinase inhibitor (which inhibits protein kinases A, C and G) or a selective inhibitor of PKA during training. Both manipulations impaired learning of an instrumental response. These results are unlikely to be explained by a disruption in performance because the infusions were given immediately after the training session. Similarly, post-training infusions of the protein synthesis inhibitor anysomicin into the NAc core, but not the shell, disrupted learning in the same training paradigm (Hernandez, Sadeghian, & Kelley, 2002). The same infusions, when given 2-4 h after the session had no effect. These studies strongly suggest that NMDAR-associated intracellular mechanisms are necessary for the consolidation of associative appetitive memories. However, due to the lack of studies that appropriately rule out a potential performance confound, the precise contribution of NAc NMDARs
to the encoding of reward-reinforced experiences as they occur is still an open question.

**Retrograde alteration: does interference with the expression of NAc synaptic plasticity disrupt aspects of a previously learned appetitive response?**

If reward-reinforced behaviors emerge as a result of synaptic strengthening in the NAc, we would expect their expression to be impaired when synaptic weights are reset. This is what was found by Pascoli et al. (2012) in an experiment that examined cocaine-induced behavioral changes. Cocaine exposure followed by a withdrawal period results in the potentiation of AMPAR-mediated glutamatergic transmission in the NAc (Kourrich, Rothwell, Klug, & Thomas, 2007). This has been associated with increases in locomotion upon subsequent exposure to the drug, a phenomenon known as behavioral sensitization (Pierce et al., 1996). By applying 5-10 min of low-frequency optical stimulation of NAc neurons in animals that had been withdrawn from cocaine, Pascoli et al. (2012) were able to depotentiate corticostriatal synapses in the NAc and abolish the acquired behavioral sensitization.

Induction of LTP depends on increased AMPAR-mediated neurotransmission (section “1.3.2.”). Therefore, if LTP in the NAc underlies the acquisition of vigorous cued reward-seeking behaviors, AMPAR blockade in the NAc core after training should disrupt this kind of responding. However, that approach has yielded conflicting results. On one hand, AMPAR antagonists infused into the NAc core decrease the ability of drug-associated stimuli to drive previously acquired cocaine-seeking behaviors (Di Ciano & Everitt, 2001). A cocaine self-administering response that has already been acquired and extinguished can be reinstated after a single intraperitoneal injection of cocaine, a phenomenon that is interpreted as an animal model of relapse (Bossert, Marchant, Calu, & Shaham, 2013). AMPAR antagonists infused into the NAc prior to the cocaine injection prevent reinstatement of the drug-seeking response in this kind of paradigm (Cornish & Kalivas, 2000). In contrast, AMPAR agonists potentiate the reinstatement of the cocaine-seeking
response without increasing food-seeking behavior or responding to an inactive lever (Cornish, Duffy, & Kalivas, 1999). These results suggest that the strength of excitatory transmission in the NAc is a critical regulator of the strength of previously learned cued cocaine-seeking responses.

On the other hand, the results are less clear in studies that investigate the contribution of AMPARs to the expression of previously learned food-seeking behaviors. AMPAR antagonists infused into the NAc core of well-trained rats did not abolish previously learned autoshaped responses (Di Ciano et al., 2001) or instrumental lever-pressing (Hernández et al., 2005). Instead, these studies found that AMPAR blockade led to an increase in task-irrelevant responses (e.g., autoshaped approach to a non-reward predictive cue, repetitive nose-poking in the absence of reward, etc.). In contrast, in a different study, AMPAR antagonists impaired expression of a previously-learned food-search pattern after a few training sessions (Maldonado-Irizarry & Kelley, 1995). Strikingly, although the expression of food-reinforced conditioned behaviors is not suppressed by AMPAR blockade in some studies (Di Ciano et al., 2001; Hernandez et al., 2005), simultaneously blocking AMPA and NMDARs yields different results. For example, NAc core –but not shell– inactivation with a combination of AMPA and NMDAR antagonists diminishes cued responding in an instrumental task (Ambroggi et al., 2011). Because activation of AMPARs is a prerequisite for the activation of NMDARs, it is unlikely that the differences across these studies are solely explained by the kind of drug infused (i.e., AMPAR antagonists alone vs. AMPAR/NMDAR antagonists combined). Similarly, inactivation (using GABA agonists) of the NAc core specifically reduces the expression of a cued Pavlovian behavior, but lesions to the shell disinhibit task-irrelevant behavior (Blaiss & Janak, 2009). Other studies call attention to different details of the experimental design as likely determinants of the outcome of this manipulation. For example, when AMPAR/NMDAR antagonists are applied to the NAc as a whole and not just to the NAc core (Yun et al., 2004), the probability of responding to a discriminative stimulus is spared (only the latency to
respond is affected), and the probability of task-irrelevant responses increases. The same combination of AMPAR/NMDAR antagonists led to a reduction in cued approach behavior after three—but not after nine—days of training (Dobrovitsky, West, & Horvitz, 2019). Therefore, contradictory results could be explained by the anatomical reach of the infusions (i.e., NAc core alone vs. NAc core and shell) or the amount of training (i.e., moderate vs. extended).

In addition, experience-dependent learning of reward-seeking behavior is unlikely to be a monolithic phenomenon. It is possible that plasticity in different synapses contributes to the acquisition of different aspects of the motivated response. Perhaps an increase in synaptic efficacy at the level of the NAc is not necessary for establishing an association between the cue and the identity or magnitude of the reward it predicts. However, it could still be necessary for learning to respond to that cue in a vigorous and efficient manner. For example, using a task specifically designed to evaluate the vigor of operant responses, Giertler, Bohn, & Hauber (2003) found that activation of AMPARs in the NAc core is necessary for rats to respond faster when cues predict larger rewards. The lack of effect on CS+ responding observed in Di Ciano et al. (2001) could be due to the extent of training before the behavioral effects of AMPAR blockade were tested. Alternatively, it could also be possible that the animals’ latency to respond was affected by the drug but that, because the length of the cue presentation was long (10 s), the probability of responding to the cue (the only reported metric) was unaffected by this. Inactivation studies have raised the possibility that the NAc is necessary for acquiring other facets of adaptive behavior, like for example behavioral flexibility (i.e., learning to ignore previously important cues and focus on previously irrelevant cues; e.g., Floresco, Ghods-Sharifi, Vexelman, & Magyar, 2006). In conclusion, when evaluating the effects of anterograde or retrograde alteration studies, it is critical to consider the complex and multilayered nature of reward-reinforced behaviors.

Finally, optogenetically inhibiting NAc afferent projections after training can also modulate
the expression of previously learned appetitive behaviors. For example, briefly inhibiting the BLA-to-NAc pathway during cue presentation diminishes cue-evoked sucrose-seeking responses (Stuber et al., 2011). In a different study, glutamatergic release from the mPFC-to-NAc pathway was found to be downregulated by sleep deprivation, which was associated with a selective enhancement in reward-seeking behaviors without affecting normal levels of food consumption. Optical stimulation of mPFC-to-NAc afferents expressing a kind of opsin that increased presynaptic release of glutamate in this synapse restored normal levels of instrumental responding for sucrose (Liu et al., 2016). These studies, just like inactivation or lesion studies, do not just selectively disrupt the mechanisms that we know are involved in the expression of neuroplasticity (e.g., AMPAR transmission), making it hard to interpret their results from the perspective of the SPM hypothesis. Nonetheless, they reinforce the idea that strong excitatory drive in the NAc is necessary for the expression of previously learned reward-seeking behaviors.

1.4.2.3. Mimicry: is the induction of plasticity within the NAc sufficient to elicit appetitive conditioning (in the absence of experience)?

Very few studies have attempted to manufacture cued reward-seeking behaviors by artificially inducing plasticity in the NAc. However, optogenetic manipulation of some of the circuits and mechanisms implicated in NAc plasticity has proven to be sufficient to modulate appetitive learning (Stuber, Britt, & Bonci, 2012). For instance, optogenetic activation of dopamine neurons in the VTA promotes the acquisition of positively reinforced behaviors in a way that is consistent with an RPE teaching signal. Optogenetic stimulation of dopaminergic VTA neurons can induce conditioned place preference (Tsai et al., 2009) and facilitate positive reinforcement during a food-seeking operant task (Adamantidis et al., 2011), but only at stimulation frequencies that elicit a
detectable surge of dopamine release in the NAc (Adamantidis et al., 2011; Tsai et al., 2009). Cues that, under control conditions, would fail to acquire motivational value because of their lack of reward predictive power, come to elicit Pavlovian approach responses after being paired with optical stimulation of VTA neurons (Sharpe et al., 2017; Steinberg et al., 2013). Alternatively, optogenetically inhibiting VTA neurons at the same time as an expected reward is delivered has the behavioral effect of a negative RPE, leading to a decrease in subsequent cued responding (Chang et al., 2016). In addition, animals will learn to lever-press or enter into a receptacle when that response results in optogenetic stimulation of dopaminergic VTA neurons (Witten et al., 2011) that project to the NAc (Saunders et al., 2018; Steinberg et al., 2014). Once acquired, this operant response is attenuated by NAc injections of D₁ or D₂ receptor antagonists (Steinberg et al., 2014). In sum, these studies suggest that activation of the VTA (a main afferent of the NAc) is sufficient to promote appetitive conditioning and reinforce operant responses in the absence of a physical reward.

Instrumental responses are also reinforced when excitatory afferents to the NAc are optogenetically stimulated. Animals engage in intracranial self-photostimulation of BLA-to-NAc projections (Britt et al., 2012; Namburi et al., 2015; Stuber et al., 2011). Activation of PFC-to-NAc projections does not support self-stimulation when administered unilaterally (Stuber et al., 2011) but it does when performed bilaterally and at a higher frequency (Britt et al., 2012). Photostimulation of hippocampal inputs to the NAc in a specific location will induce conditioned place preference for that location, whereas their inhibition will prevent animals from learning to prefer a site where they experienced social reinforcement (LeGates et al., 2018). Similarly, optogenetic activation of hippocampus-to-NAc inputs elicits self-stimulation and place preference (Britt et al., 2012). Optical inhibition of these inputs reduced (and their activation increased) cocaine-induced behavioral sensitization (Britt et al., 2012). In sum, strong activation of all glutamatergic afferents to the NAc can promote the acquisition of reward-seeking behaviors. In fact, robust self-photostimulation is
also observed when mice are given the opportunity to activate their NAc MSNs upon entry into a receptacle (Britt et al., 2012). These results suggest that strong excitatory transmission within the NAc —regardless of the source— may be sufficient to reinforce behavior. However, it remains to be seen whether this finding holds true in physiological conditions when glutamate is not released at the unnaturally high frequencies elicited by optogenetic stimulation.

1.4.3. Unresolved questions: goals of this research project

The body of literature reviewed in the previous section (“1.4.2.”) generally supports the hypothesis that plasticity in the NAc is a likely substrate of reward-related learning. However, some basic questions remain unanswered.

- Are experience-dependent changes in NAc firing patterns an eligible substrate of cue-reward learning? In trained animals, cue-evoked responses in the NAc core have been causally linked to the initiation of vigorous cued approach behavior (Caref & Nicola, 2018; du Hoffmann & Nicola, 2014; Ishikawa et al., 2008; McGinty et al., 2013; Yun et al., 2004). *In vivo* electrophysiological recordings in behaving naïve rats have identified populations of NAc neurons that come to encode the motivational value of the cue and the strength of the subsequent approach behavior after appetitive conditioning (Atallah et al., 2014; Roitman et al., 2005; Setlow et al., 2003). However, if NAc core neurons exhibit cue-evoked excitations only after animals begin to show cued reward-seeking behavior, these signals might merely be correlates of behavioral performance rather than reflections of reward-related NAc plasticity. **Experiment 1** investigates the precise temporal relationship between the changes in behavior and NAc neuronal activity during
training.

- **Does intra-NAc NMDAR blockade disrupt the expression of motivated behavior?** In the field of anterograde alteration studies, the strongest evidence to date to support that plasticity in the NAc is necessary for appetitive learning comes from a series of studies that impair learning by training rats under the effects of the NMDAR antagonist AP5 (section “1.4.2.2.”). The interpretation of these results would be put into question if, early in training, intra-accumbens blockade of NMDARs disrupted normal behavior, thereby preventing animals from experiencing critical task-related contingencies (section “1.4.2.2.”). **Experiment 2** tackles this question by exploring the neural and behavioral effects of blocking NMDARs in the NAc core of animals that had been exposed to moderate vs. extended training. By simultaneously performing electrophysiological recordings from the same structure that received these infusions, this experiment reveals a neural mechanism underlying the behavioral effects of the drug.

- **Do NMDARs in the NAc core contribute to learning above and beyond their involvement in the expression of exploratory and learned behavior?** **Experiment 2** established that bilateral infusions of NMDAR antagonists into the NAc core can disrupt cued approach performance early in training. By implementing a design that circumvents the expression/acquisition confound, **Experiment 3** examines the specific contribution of NAc core NMDARs to learning during early training.

- **Is NAc core plasticity necessary for the learning-related changes observed in NAc core neurons?** The NAc core receives strong excitatory input from other structures that also
exhibit dynamic cue encoding during training (Paton et al., 2006; Schoenbaum, Chiba, & Gallagher, 1998; Tye, Stuber, de Ridder, Bonci, & Janak, 2008; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010; Wirth et al., 2003). Therefore, it is possible that cue-evoked excitations emerge as a result of synaptic strengthening upstream of the NAc and that NAc neurons merely transmit these signals to downstream structures. Because NMDARs have a well-established role in NAc LTP (section “1.4.1.”) and in the acquisition of cued approach behavior (section “1.3.3.”), the emergence of cue-evoked excitations in the NAc is likely dependent on NMDAR activation during training. This hypothesis was tested in Experiment 4.

In sum, this research project describes changes in conditioned activity in the NAc core neurons during training, it connects the acquired firing patterns to the acquisition of cued approach behavior, and it establishes NMDAR-dependent plasticity within the NAc (not elsewhere in the circuit) as a necessary mechanism for this kind of learning. Altogether, these results contribute to causally implicate NAc plasticity in cued approach learning by meeting three of the four criteria proposed by Martin et al. (2000): detectability (Experiments 1 and 4), retrograde alteration (Experiment 2) and anterograde alteration (Experiments 3 and 4).

1.5. General materials and methods

This section describes the methodological details that are common to all experiments. The materials and methods specific to each experiment will be introduced in the “Materials and methods” section that correspond to each experiment.

1.5.1. Subjects

Male Long-Evans rats (N=58; 350-375 g; Charles River, NY) were used in this study. Upon
arrival, they were placed on a 12 h light/12 h dark schedule and housed in groups of two or three. They were handled regularly for at least one week before surgery (or training if that came first). Following surgery, animals were singly housed, received postsurgical care as necessary and were allowed to recover for at least one week. After recovery, they were placed on a restricted diet of 13 g of standard chow. Behavioral training started after animals had been food deprived for a week and the restricted diet was maintained for the remainder of the experiment. Animal care was in accordance with the U.S. National Institute of Health Guide for the Care and Use of Laboratory Animals and approved by the Institutional Animal Care and Use Committee at Albert Einstein College of Medicine.

1.5.2. Operant chambers

Custom-made Plexiglass operant chambers (40 cm square base and 60 cm tall) were used in this study. Chambers were controlled using Med-PC software. On one of the walls of each chamber there was a reward receptacle equipped with an infrared beam for detection of head entries and exits. Chambers were also equipped with 28 V house lights, a white noise generator and speakers. A syringe pump adjacent to the chamber delivered the reward into a well inside the receptacle via steel-reinforced PVC tubing. All chambers were kept inside soundproof cabinets and 65 dB white noise was played throughout each session to minimize acoustic interference. Timestamps associated with task-related events were recorded with a resolution of 1 ms.

1.5.3. Surgeries

Rats were anesthetized with isoflurane, set in the stereotaxic apparatus and their scalp was retracted. Bone screws were inserted in the surface of the skull for enhanced support of the implant and holes were drilled above the target structure. Implants (cannulae in “Experiment 3”, electrode
arrays in all other experiments) were fixed to the skull and screws using a combination of permanent glass ionomer cement and acrylic dental cement. In all experiments, implants were targeted at the NAc core. Rats were treated with intraperitoneal injections of an analgesic solution (Ketofen), subcutaneous injections of antibiotic (Baytril) and topical antibiotic powder (Neopredef).

1.5.4. Drugs

The competitive NMDA receptor antagonist D-(-)-2-Amino-5-phosphonopentanoic acid (AP5, Tocris; Ellisville, MO) was dissolved either in sterile 0.9% saline (“Experiment 3”) or in phosphate buffered saline (PBS) 0.1 M (all other experiments), divided in aliquots and stored in a -20º C freezer for up to a month before being used. The chosen concentration (2 µg/1 µl) has consistently been found to disrupt appetitive conditioning when bilaterally infused into the NAc (Di Ciano et al., 2001; Hernandez et al., 2005; Kelley et al., 1997). The solution in which the drug was mixed (either sterile 0.9% saline or PBS) was used as a control solution. Aliquots of both solutions were retrieved from the freezer ~ 30 minutes before behavioral testing.

1.5.5. Behavioral indices

Several variables were used to quantify the strength of cued approach behavior:

a) **S+ entry probability**: number of S+ trials in which the animal made a receptacle entry while the cue was on divided by the total number of S+ trials during the session

b) **S- entry probability**: number of S- trials in which the animal made a receptacle entry while the cue was on divided by the total number of S- trials during the session

c) **ITI entry probability**: number of S+ trials in which the animal made a receptacle entry during the 10 s window that preceded the S+ divided by the total number of S+ trials during the session
d) **S+ latency**: latency to enter the reward compartment upon S+ onset;

e) **S- latency**: latency to enter the reward compartment upon S- presentation (for both S+ and S- latency measures, a 10 sec maximum latency was assigned if no entry was made during the 10 s duration of the cue)

f) **ITI pseudolatency**: latency from the point 10 s prior to cue onset to the first receptacle entry prior to cue onset. If no entry was made during this 10 s ITI window for a particular trial, a value of 10 s was assigned to the ITI pseudolatency for that trial (Figure 3). In 9-15% of trials (depending on the experiment), the rat’s head was already inside the receptacle at the onset of this ITI window. Those trials were assigned the average ITI pseudolatency of the set of ±5 trials surrounding that trial.

g) **Performance index**: latency to respond to the cue compared to the latency that would be expected from the animal’s overall response frequency (Figure 3). It is calculated by subtracting the cued latency from the ITI pseudolatency on each trial (S+ trials and S- trials were treated separately to calculate S+ and S- performance index, respectively). It ranges from -10 to 10, with positive values indicating that the latency to respond to the cue was shorter than that predicted by the rate of receptacle entry during the preceding ITI. Negative numbers indicate the opposite. Values around 0 indicate a lack of influence of the cue on approach behavior.
Figure 3. Performance index calculation.

(a) Representative behavioral raster plots of one animal on the first (left, Day 1) and the last (right, Day 6) day of training. Within each panel, performance is divided into S+ and S- trials. Each trial is shown in a different row, and trials are sorted earliest to latest from bottom to top. Black horizontal lines within each trial represent periods when the rat’s head was inside the reward receptacle. Data is aligned to the time of cue onset (vertical red line). Arrows mark the 10 s interval before and after cue onset. The raster plots show that, early in training, an overall high frequency of entry into the reward receptacle may preclude the interpretation
of entry during the S+ as specifically cue-driven behavior. Note that fluctuations in S+ responding are accompanied by fluctuations in responding during the intertrial interval. This emphasizes the need to consider the rate of indiscriminate responding (i.e., during the ITI) when quantifying cued responding.

(b) Calculating the performance index. The left panel represents hypothetical performance on ten trials aligned to the time of S+ onset (vertical red line). Pink rectangles span the duration of the S+. Black rectangles depict entries into the receptacle. A dashed red line indicates the beginning of a window beginning 10 s prior to cue onset. For each trial, two latency values were calculated: the interval from the point 10 s prior to the cue to the first receptacle entry occurring prior to the cue (ITI pseudolatency), and the period during which the cue was on (cued latency, corresponding to the interval between cue onset and receptacle entry). If no entry was made during one of these periods, a value of 10 was assigned. To calculate the performance index of the animal on a given trial, its cued latency on that trial was subtracted from its ITI pseudolatency on the same trial. The performance index ranges from -10 to 10, with negative values indicating that the animal entered into the receptacle faster in the absence of the cue than in its presence, and positive values indicating the opposite. Values around zero suggest that the cue has no influence on receptacle entry behavior. The table on the right shows ITI pseudolatency, cued latency and performance index corresponding to the trials shown in the left panel.
1.5.6. Change point analysis

It has long been recognized that group averages of performance during training fail to capture critical dynamics of individual learning curves (Brown & Heathcote, 2003; Estes, 2002; Krechevsky, 1932). In many basic learning paradigms, group averages suggest that conditioned responses emerge gradually and follow a negatively accelerated curve. However, some authors suggest that the transition from a phase of no progress to a phase of mastery is abrupt rather than gradual (Gallistel et al., 2004). An alternative for quantifying the learning curve involves identifying the trial after which there is a consistent expression of cued behavior. This trial is called the “change point” (CP). In order to identify the CP, we used a variation of the method used in Gallistel et al. (2004).

Gallistel et al. (2004) suggest that the first appearance of conditioned responding can be identified by inspecting the cumulative record of each animal’s responses as a function of the trials experienced up to that point. These charts are a powerful tool in the identification of behavioral trends (Skinner, 1976) because random changes in behavior from one trial to the next are minimized while steady changes in performance are emphasized by changes in the slope of the line. These records sometimes undergo small changes in slope, and Gallistel et al. propose a method for quantifying the significance of these putative change points. The method is a recursive algorithm that is successively run over each data point in the cumulative record of an animal. It has four steps (Figure 4). First, it measures the degree of deviation of a particular data point from the cumulative record by drawing a straight line from that data point to either the very first trial or the previous change point (whichever is closest to the trial to which the algorithm is currently being applied). Second, it finds the point that maximally deviates vertically from the straight line, and this trial becomes the putative or “test” change point. The algorithm then calculates the strength of the evidence that the distribution of trials after the test change point is different from the distribution of
trials before the test change point (i.e. the log of the odds against the hypothesis that the test change point is not a true change point). In the third step, if the strength of the evidence is larger than a user-set logit value—the most sensitive value proposed by Gallistel et al. was used, logit = 1.3, which corresponds to $p < 0.05$ — the algorithm truncates the data and treats that true change point as the new origin (it becomes a “candidate” change point). Finally, the algorithm repeats the process using the new origin. Using this algorithm, each rat’s individual record was broken down in a series of candidate change points that identified changes in slope that are maintained across trials.

Gallistel et al. used the first upward change in the slope of cumulative behavioral responses as a function of trials to identify the trial in which the conditioned behavior debuts. In this study, the change point calculation was conducted based on the cumulative records of the performance index (see Figure 3 and section “1.5.5. Behavioral indices”). Unlike the cumulative conditioned response used by Gallistel et al., the performance index can exhibit both positive and negative values. If one of the candidate change points identified by the algorithm was followed by a negative slope, that would suggest that, for a series of trials, the animal was consistently responding more vigorously in the absence of the cue than in its presence. Even though it is not uncommon to observe fluctuations in conditioned behavior after it is acquired, it is unlikely that an animal that has truly learned the meaning of the cue would suddenly cease responding to the cue for a series of trials while maintaining a high overall rate of responding during the ITI. For that reason, an additional criterion was added to the algorithm: identification of a “definitive” change point (the first trial after which there is evidence of robust conditioned responding) required that the slopes of all the subsequent inter-candidate-change-point segments were positive (Figure 4). This is the change point used in our analyses and is referred to as “the change point” or “CP” throughout this work.
Figure 4. Identifying the behavioral change point.

(a) Hypothetical application of the change point algorithm to the cumulative record as of trial 30 (adapted from Gallistel, Fairhurst and Balsam, 2004). First, a straight line is drawn from trial 30 to the origin. Second, the trial that maximally deviates from that line is identified as a potential candidate change point (test change point). Third, performance values before and after the test point are compared. If the null hypothesis of no change can be rejected at a user-specified significance value, that test change point is considered a candidate change point. The algorithm then truncates the data at that point and treats that candidate change point as the new origin. Finally, the algorithm starts the process all over again, running successively over each trial in the cumulative record.

(b) The result of this iterative algorithm is typically a list of candidate change point trials. Gallistel et al. take the first candidate change point in the cumulative record as the definitive change point – the first trial after which cued behavior can be consistently detected. However, they applied the algorithm on behavioral variables that can only adopt null or positive values, which yield cumulative records in which the change of the slope can only detect an improvement in behavior or lack of thereof (i.e., the slope can only be positive or 0). In contrast, our performance index can also capture instances in which the animal’s likelihood or speed of cued responding is less than what would be expected from its baseline behavior. As a result, at the beginning...
of training, it is not unusual to find brief increases in the slope of the line followed by decreases. For that reason, for a candidate change point to be identified as definitive in our paradigm, the subsequent segments between candidate change points in the cumulative record had to have a positive slope, or the candidate change point was rejected. The slope of these segments could fluctuate – as is common for conditioned behavior even after it is acquired – but it could not be negative. Therefore, we determine the definitive change point as the first candidate change point for which all subsequent slopes are positive, and we report this trial as the change point (CP) in the main text. This trial corresponds to that on which consistent cued behavior first appears.

(c) Sample performance of one subject (“B”) throughout training in three graphs. Gray lines indicate the transition between sessions. Top: average S+ performance index in five-trial bins (blue). Middle: cumulative S+ performance index record. Blue dots mark all of the candidate change points identified by the algorithm. The vertical red line marks the change point. Bottom. Trial by trial S+ performance (black) and average performance before and after the change point (red).
1.5.7. Statistical analyses

Analyses of imported NeuroExplorer and Med-PC files were performed using custom routines in the R software environment. A summary of all the statistical tests shown in the main and supplementary figures can be seen in “4.2. Supplementary Table”. Whenever multiple comparisons were conducted on the same set of observations, p-values were adjusted using the Holm-Sidak correction. Throughout the study, p-values smaller than 0.05 resulted in the rejection of the null hypothesis. Labels indicate whether the data shown in the figure come from all neurons or only those that were classified as significantly excited or inhibited by the event of interest.

Unless otherwise indicated, firing rate data was converted to Z scores using the 2 s before cue onset as the baseline window. Comparisons of cue-evoked firing rate across conditions within the same neuronal population were conducted using Wilcoxon signed rank tests. Wilcoxon rank sum tests were used for comparisons across different neuronal populations. Proportions of significantly cue-excited or cue-inhibited (and entry-excited or entry-inhibited) neurons were compared using Fisher’s exact test for count data. The relationship between the magnitude of S+-evoked phasic signals and cued approach behavior was also evaluated. To do so, the post S+ firing rates of all the units recorded on a given session were averaged and plotted as a function of the average performance of that animal on that session according to three different behavioral indices: S+ latency, S+ response ratio and S+ performance index. Simple linear regression models were used to quantify the relationship between these variables. Observations whose Cook’s distance deviated more than three times from the average Cook’s distance were classified as outliers and excluded from the analysis (none of the results were substantially different when outliers were included, both versions of the analysis are always shown in Supplementary Table 1). A similar approach was used to quantify the relationship between firing rate after cue onset and after receptacle entry.

When animals learned the task successfully, behavioral and neural data from each rat were
aligned to the trial in which the change point took place and binned with respect to that trial. One disadvantage of aligning neural data to the trial in which the behavioral change point took place before averaging it is that two consecutive trial bins may contain data from the same neuron — which would not be the case if the data were simply averaged by session in order, because different populations of neurons were recorded every day. This makes it difficult to compare firing rate between two consecutive bins, because some of the data points in those two bins would have been recorded from independent samples (i.e., two different neurons on two different days) while others may have been recorded from the same sample (i.e., same neuron on the same day, two different sets of trials). To prevent violating assumptions of dependence/independence of observations required by the statistical tests that were employed, only non-consecutive bins were compared (e.g., it would be impossible for a neuron to contribute data to two non-consecutive 40-trial bins when each neuron was only recorded for 40 trials).

To test whether injections affected baseline firing, pre- and post-injection baseline firing rates were plotted against each other and the 99% confidence interval was constructed around the slope of the resulting regression line. Baseline firing rate was considered to be unaffected by the injections if the slope of the regression line did not significantly differ from the unity line (its confidence interval included the value 1). Firing rate during the 5 s window before cue onset was used to define baseline activity in this analysis.

Paired t-tests were used to compare performance before vs. after the behavioral change point. In Experiment 2, mixed two-factor ANOVAs with “injection” (AP5 vs. Vehicle) as a between-subject factor and “bin” (each 30 min bin of the session) as a within-subject factor were used to test the effects of AP5 on performance index throughout the test session in the “moderate training”. In the “extended training” group, “injection” (AP5 vs. Vehicle) was treated as a within-subject factor. Per bin post-hoc comparisons were conducted with two-sample one-tailed t-tests. Unpaired and
paired t-tests were used for between- or within-subject comparisons, respectively. Welch’s t-tests were used when the assumption of unequal variances was violated. A two-factor ANOVA was also used to compare performance during the extinction test sessions in Experiments 3 and 4. The Greenhouse-Geisser correction was applied whenever the assumption of sphericity was violated.

1.5.8. Histology

At the end of behavioral training and testing, animals were deeply anesthetized with Euthasol (39 mg/kg pentobarbital) and intracardially perfused with saline and a 4% paraformaldehyde solution. Their brains were removed after decapitation and then stored in 4% paraformaldehyde solution until further processing. After rinsing them with PBS 0.1 M, all brains were sliced into 50 μm sections using a vibratome. Sections were later mounted on slides and stained with cresyl violet to facilitate verification of injection and recording locations (Paxinos & Watson, 2007; Supplementary Figure 10).
2. EXPERIMENTS
2.1. Experiment 1: Experience-dependent changes in NAc activity during appetitive learning.

2.1.1. Introduction

In trained animals, NAc neurons exhibit cue-evoked firing responses that are necessary for initiating vigorous cued approach behavior (Ambroggi et al., 2008; Caref & Nicola, 2018; du Hoffmann & Nicola, 2014; McGinty et al., 2013; Morrison et al., 2017; Nicola et al., 2004a; section “1.2.4”). Many studies also suggest that dopaminergic and glutamatergic transmission in the NAc are necessary and sufficient for acquiring cue-reward associations and reward-reinforced responses (Britt et al., 2012; Chang et al., 2017; Di Ciano et al., 2001; Hernandez et al., 2005; Kelley et al., 1997; LeGates et al., 2018; Namburi et al., 2015; Saunders et al., 2018; Steinberg et al., 2013; sections “1.2.3.3.” and “1.4.2.2.”), suggesting that the NAc plays a critical role in this form of learning. However, the NAc neural mechanisms that underlie the acquisition of appetitive responses are not clear. If the NAc is causally implicated in appetitive learning, we would expect NAc neuronal activity to dynamically encode cue-reward and/or cue-action associations during learning (section “1.3.3.1.”). Although some studies have found larger NAc cue-evoked responses in trained rats than in naïve rats (Atallah et al., 2014; Roitman et al., 2005; Setlow et al., 2003; section “1.4.2.1”), it is not clear whether the emergence of these signals antecedes, accompanies or follows the emergence of the learned behavior. Also, if NAc core excitations encode the vigor of the reward-seeking behavior, we would expect these signals to become progressively larger as cued approach responses become more and more vigorous. The results of this experiment outline the precise temporal relationship between the evolution of learning-related signals in the NAc and cued approach behavior.

Acquisition of cued approach depends on unconstrained exploratory behavior to produce chance encounters with the reward in close temporal proximity to the cue. Consequently, the rate of learning is highly variable across individuals, which complicates the extraction of meaningful group
data. In addition, although high rates of exploratory behavior facilitate learning, they obscure the extent to which responses are motivated by the reward-predictive cue. An individualized analysis based on prior theoretical work (Gallistel, Fairhurst, & Balsam, 2004) allowed for the detection of the first trial after which each animal showed consistent cued approach behavior. This method provided the sensitivity required to establish the precise relationship between experience-induced increases in NAc encoding and the acquisition of cued approach in freely moving animals. The results of this experiment show that, with repeated cue-reward pairings, cue-evoked excitations in the NAc emerge and grow in the trials prior to the detectable expression of cued approach behavior. The magnitude and prevalence of these signals continues to escalate as the learned behavior becomes more vigorous.

2.1.2. Materials and methods

2.1.2.1. Cannulated microelectrode arrays: construction and intracranial implantation

Cannulated microelectrode arrays were assembled following the description in du Hoffmann, Kim, & Nicola (2011). Each one of these arrays consists of 8 tungsten microwires (A-M systems; Sequim, WA) surrounding a 27-gauge microinjection guide cannula (Figure 5). Before assembly, each microwire’s impedance was tested to ensure it was in the 90-110 MΩ range. Both the microwires and the cannula were assembled into a custom-designed plastic drive with two parts connected by a screw that allowed dorsoventral displacement of the bundle of microelectrodes and cannula with no rotation of the probes (each full turn of the screw displaced the array ~300 μm). Finally, the microelectrodes and a silver ground wire were soldered onto connectors (Omnetics; Minneapolis, MN) that were cemented behind the cannulated microarrays. The silver ground wire was wrapped around one of the most posterior screws and then inserted inside the brain about 0.7
cm deep. The coordinates for the tips of the arrays were, from bregma: AP: +1.44 mm, ML: ±1.5-1.6 mm, DV: -6.5 to -7 mm (the tips of the injectors extended 0.5 mm beyond the tips of the cannulated array once inserted).

![Diagram](image)

**Figure 5. Schematic depiction of the cannulated microelectrode array.** (a) Frontal view of a cannulated microelectrode array. A custom-made plastic drive provides the structure for a bundle of tungsten wires surrounding a 27-gauge microinjection guide cannula. The top and the bottom parts of the drive are connected by a drivable screw that allows dorsoventral displacement of the bundle without rotation. (b) Cross-section of the array. Eight tungsten wires surround a central guide cannula. In red, distance between each pair of wires and between each wire and the center of the cannula.

### 2.1.2.2. Training protocols

**Pretraining phase.** After a week of food restriction and regular handling, animals were habituated to 10% sucrose in their home cage (15 ml). After that, they received a mock infusion in which two injectors were bilaterally inserted and remained in place for ~ 30 s. Before training, rats were also habituated to the experience of being tethered to the recording cable. This was carried out in a plastic container of similar dimensions to the operant chamber. One end of a recording cable
was connected to a commutator above the container, and the other end was attached to the rat’s headstage for at least 30 min. Before training, all rats also received a receptacle training session. During the receptacle training session, animals received 40 rewards upon receptacle entry on an FR1 schedule. Delivery of the first 20 rewards was followed by a 10-15 s time out and delivery of the last 20 rewards was followed by a 30-45 s time out. If animals failed to collect all the rewards within ~45 min, they were given a second receptacle training session the next day. On this second session, the receptacle was baited with sucrose before putting the animal inside the chamber.

Training phase. Each trial consisted of an inter-trial interval (ITI) and one of two cues: an S+ or an S-. A head entry into the reward receptacle during presentation of the S+ resulted in delivery of a droplet (~150 μl) of 10% sucrose into the same compartment (“correct trial”) and offset of the cue. If the animal failed to enter the compartment during presentation of the cue (“missed trial”), the tone would terminate after 10 s and sucrose would not be delivered. The S- always ended after 10 s. Entries into the compartment during the S- or the ITI had no programmed consequences. Both cues consisted of non-localized stimuli of different sensory modalities to facilitate discrimination. One cue was an auditory stimulus (a “siren” alternating between 4-8 kHz) and the other one was a visual stimulus (four houselights turning on in an otherwise dark chamber). Assignment of sensory modality to cue type (S+ or S-) was counterbalanced across subjects. Each daily session consisted of 80 trials: 40 S+ and 40 S- trials. Both kinds of trials were randomly interleaved. The ITI was 15-45 s on day 1 and 20-100 s subsequently.

2.1.2.3. Acquisition of neural data

On recording days, a headstage containing unity-gain operational amplifiers was plugged into connectors that were cemented to the animal’s skull. A recording cable extended from the headstage
to a multichannel commutator above the chamber, allowing the animal free movement within the operant chamber. Signals were amplified (2,000-20,000X) and band-pass filtered at 250 Hz and 8.8 kHz before being sent to 40 kHz multi-unit acquisition processors. Once animals were tethered to the recording cable, each of the 16 channels was monitored using SortClient (Plexon Inc, Dallas TX) and the threshold and gain were adjusted to optimize the signal.

Electrophysiological recordings taken during behavioral sessions were subsequently processed using Offline Sorter (Plexon). Putative neurons were manually defined by identifying clusters of spikes in a 3D feature space. In that space, different combinations of each spike’s features —i.e., waveform projection onto its principal components, difference between the maximum and minimum amplitude of the waveform and the waveform height at a particular point in time— were represented on the X, Y and Z axes. In order to be identified as such, an individual neuron had to show a clear and consistent waveform (> 75 µV) and less than 0.1% of the interspike intervals could be less than 3000 µs. Cross-correlograms were also used to make sure that there was no overlap among different units identified within one single channel.

Given the nature of our inquiry, monitoring activity of the same neuron across days would be ideal. However, the exact location of the tips of the electrode arrays may slightly shift across days, and it would be virtually impossible to track specific neurons from one session to the next and distinguish them from new neurons that may appear on the record for the first time. The resulting data set would be a statistically unmanageable combination of repeated and non-repeated measures. Consequently, the standard practice of ventrally advancing electrode arrays in between sessions was adopted, thus ensuring that a new population of neurons would be recorded the next day. Arrays were lowered by turning the microdrive screw about half a turn (~ 150 µm). One drawback of this approach is that potential changes in neural activity observed during training could be accounted for by the changing location of the probes along the dorsoventral axis rather than by the increasing
degree of exposure to the task throughout training. In order to address that confound, electrode arrays were kept in place during training in a group of eight rats as a control (Supplementary Figure 4). Occasionally, no neurons were recorded on a particular day (either because the electrode arrays failed to capture any signals above the specified voltage threshold or, more rarely, because of technical difficulties with the recording system during the session). If that happened, rats were run anyway to keep the training schedule consistent across subjects.

Neurons were classified as significantly cue-excited if their firing rate exceeded the upper limit of the 99.9% confidence interval of a Poisson distribution comprised of the 5 s pre-cue baseline for at least three consecutive 50 ms bins (up to 500 ms after cue onset). Neurons were classified as significantly cue-inhibited if their firing rate fell below the lower limit of the same 99.9% confidence interval for at least two consecutive 50 ms bins within the post-cue 500 ms window. For units to be classified as significantly excited (or inhibited) upon S+, S- or ITI receptacle entry, their firing rate in the 1500 ms window after receptacle entry had to exceed the upper limit (or fall below the lower limit, in the case of inhibitions) of the 99.9% confidence interval of a Poisson distribution calculated based on the 5 s pre-cue baseline for at least six consecutive 50 ms bins. Offline Sorter files were saved as NeuroExplorer (Plexon) files and later imported as spike timestamps into R.

2.1.3. Results

To understand how experience modifies NAc neuronal activity around the time of task events, we recorded the activity of individual NAc core neurons in 6 naïve rats as they learned a cued approach task (Figure 6a). To earn a droplet of sucrose reward, animals had to enter a reward receptacle within 10 s of the onset of a discrete, non-localized auditory or visual cue (S+); responses at other times, including during presentation of the unrewarded cue (S-), had no programmed
consequences. Each session consisted of 40 S+ and 40 S- trials, randomly interleaved. Before examining changes in NAc activity during learning, we analyzed individual subjects’ behavior during task acquisition.

During the first 2-3 days of training, animals showed a high rate of receptacle entries indiscriminately during the S+, S- and the 10 s window at the end of intertrial intervals (ITIs; Figure 6b, left). With additional training, the probability of responding to the S+ gradually increased while the latency to respond to the S+ with a receptacle entry decreased (Figure 6b, right). Although these observations indicate that the animals learned to respond to the S+, the animals continued to respond with high probability during the S- and ITI. Neither S+ response probability nor S+ response latency is a sufficient metric for cued approach performance because neither reflects the difference in rate of responding to the S+ vs. responding in the absence of cues. To calculate this difference, it was determined, for each trial, the ITI pseudolatency, the interval between a point 10 s prior to cue onset and the next receptacle entry. Next, for each trial, the performance index was calculated. This is the difference between an animal’s ITI pseudolatency and its cue-evoked latency to enter the receptacle on the same trial (latencies were assigned values of 10 s if the animal made no response within 10 s. Performance index values range from -10 to 10, with positive values indicating a faster response to the cue than predicted by the rate of responding during the ITI, and negative values indicating a slower-than-predicted response (see Figure 3 and section “1.5.5.: Behavioral indices”).

To assess the time course of learning for individual subjects, the cumulative records of the performance index were examined (Figure 6c,d). These oscillated around 0 at the beginning of training, indicating that the cue did not yet influence the subject’s approach behavior. Later, at the trial on which faster-than-predicted responses to the S+ were first reliably evoked, a discrete change in the slope of the line can be observed (“change point”, CP; Gallistel et al., 2004; see Figure 4 and
section “1.5.6.: Change point analysis”). An upward CP occurs in S+ (Figure 6c, left), but not S-
trials (Figure 6c, right). Interestingly, some animals (e.g., Animal “F” in Figure 6c) may have
learned that the S- was a conditioned inhibitor before they learned about the S+-reward contingency.
Importantly, subjects varied in the amount of training necessary to display a CP (Figure 6d).
Therefore, the apparent gradual increase in conditioned behavior seen in group data (Figure 6b) is at
least partly an artifact of averaging sudden improvements in cued responding that occur at different
points in different subjects.

To characterize the behavioral changes associated with the CP, the mean probability and
latency to respond during the S+, S- and ITI was examined in trials aligned to each animal’s CP
(Figure 6e,f). Before the CP, a steady increase in the probability of entering the reward receptacle
was observed during the S+, the S- and the ITI periods (Figure 6e, left); animals also showed a
decreased latency to respond during each of the periods (Figure 6e, right). These observations
indicate that the rate of responding gradually increased prior to the CP irrespective of the presence
of cues. However, at the CP and afterward, there was a clear dissociation between the cue
conditions: the probability of entering the receptacle in the presence of the S+ continued to increase
towards asymptotic levels, and the latency to respond to the S+ decreased. At the same time, the
probability and latency of responding in the absence of the S+ stabilized at the CP (Figure 6e).
These results were confirmed by mean pre- and post-CP measures (Figure 6f). Thus, the CP
represents the trial on which responses to the S+ undergo a sudden and stable increase in probability
and vigor, while responses to the S- and to contextual cues (present during the ITI) remain
unchanged. Notably, S+ response probability and latency continued to gradually increase and
decrease, respectively, after the CP (Figure 6e), indicating that the gradual performance increase in
data averaged across animals (Figure 6b) was due not only to the staggered occurrence of CPs
across subjects but also to gradual performance increases occurring in trials after the CP within
individual subjects.

To examine how NAc activity evolved as animals (N=6) learned the task, 186 NAc neurons were recorded throughout training. Before combining neuronal activity data from different subjects, the trials from each recorded neuron were aligned to the CP trial in the corresponding subject (data from representative neurons can be found in Supplementary Figure 1). We focused first on cue-evoked excitations because they are prominent in well-trained subjects (Ambroggi et al., 2011; Cromwell & Schultz, 2003; Day et al., 2006; Hassani et al., 2001; Nicola et al., 2004a; Schultz et al., 1992) and play a causal role in cued approach behavior (Caref & Nicola, 2018; du Hoffmann & Nicola, 2014). In S+ but not S- trials, large cue-evoked excitations, which peaked at around 100-400 ms after S+ onset, emerged just prior to the CP (Figure 7a, b). This pattern of emergence of cue-evoked excitations did not depend on the sensory modality of the cue (Supplementary Figure 2c), or on whether the electrode arrays were advanced between sessions (Supplementary Figure 4b,c).

Cue-evoked excitations were significantly higher after S+ than after S- onset as early as the first 10-20 trials after the beginning of training (Supplementary Figure 2b), indicating that NAc neurons discriminated between the relative reward-predictive value of cues well before the difference in value of the cues was reflected behaviorally. This is consistent with previous findings (Setlow et al., 2003). During this early stage of learning, cue-evoked firing responses remained small. At a later stage, during the trials leading up to the CP, cue-evoked excitations progressively increased in magnitude (Figure 7a-c, e) and prevalence (Figure 7d). To address whether the growth in cue-evoked excitations precedes, follows or is concomitant with the debut of vigorous cued approach behavior, cue-evoked NAc activations occurring during the bin of 40 trials that preceded the CP was compared to an earlier and later trial bin (see section “1.5.7. Statistical analyses” for an explanation of how the three comparison trial bins were chosen). Both the proportion of neurons that became significantly excited upon cue presentation (Figure 7d) and the magnitude of their cue-
evoked excitations (Figure 7e) began to increase just prior to the CP. After the CP, as reward-seeking behavior became more vigorous, the prevalence (Figure 7d) and magnitude (Figure 7e) of cue-evoked excitations in NAc core neurons continued to increase. The magnitude of cue-evoked excitation was strongly correlated with the latency and probability of the cued approach (Figure 7f), consistent with observations in well-trained animals (Caref & Nicola, 2018; du Hoffmann & Nicola, 2014; McGinty et al., 2013; Taha et al., 2007) and supporting the hypothesis that S+-evoked firing promotes approach. (The proportion of cue-inhibited neurons significantly decreased after the CP; however, because these represent a small proportion of units – ~18% – our analyses focus on cue-evoked excitations.)

In reinforcement learning, reward is typically experienced some time after the reward-predictive cue and reward-eliciting actions occur, thus posing a theoretical conundrum known as the “credit assignment problem” (Minsky, 1961): how do neural representations that deserve credit for predicting the reward become selectively strengthened? One potential mechanism is that cue-excited neurons remain excited during subsequent reward delivery, providing an eligibility trace that facilitates further strengthening of cue-evoked excitation. To determine whether NAc neural activity follows a prolonged time course consistent with such a mechanism, we first established that the firing of cue-excited neurons just prior to and after the CP remained elevated during the 750-2000 ms post-S+ window, long after their initial firing peak (Supplementary Figure 2d). Next, we asked whether the increased firing persisted until receptacle entry in trials in which the entry occurred at least 5 s after cue onset. Firing during the 2 s window prior to such entries was indeed higher during S+ presentation than S- presentation (Supplementary Figure 2e).

Finally, we examined whether an additional excitation occurred during reward delivery by aligning firing to reward receptacle entry. Before the CP, some NAc neurons exhibited excitations within 1 s after rewarded entries occurring during S+ presentation (Figure 8a). Moreover, there was
a significant positive correlation between S+-evoked and entry-evoked activity at this point in training, suggesting that neurons that became cue-excited early in training also tended to exhibit elevated activity upon reward delivery (Figure 8b, top). In contrast, activity triggered by the S+ was not correlated with activity after unrewarded receptacle entries during the S- or ITI (Supplementary Figure 3b). Together, these results suggest that NAc neuronal firing forms a temporal bridge between cue presentation and reward, potentially serving as a neural substrate for plasticity.

In the CP session and afterward, receptacle entries during the S+ were followed by pronounced inhibitions (Figure 8a). The emergence of entry-related inhibitions was absent in non-rewarded entries (Supplementary Figure 3a). Also, cue- and reward-elicited activity became negatively correlated (units that were more excited upon S+ onset became more inhibited at the time of reward consumption; Figure 8b, bottom), consistent with previous observations in trained animals (Nicola et al., 2004a, 2004b). No such correlation was observed between post-S+ activity and that seen following non-reward receptacle entries (Supplementary Figure 3c). Thus, not only S+-evoked firing but also reward-associated firing changes during learning.
Figure 6. During training, an identifiable transition ("change point", CP) between exploratory and cued approach behavior occurs. (a) Task diagram. Only receptacle entries during the S+ were rewarded. S+ and S- trials were randomly interleaved. ITI: intertrial interval. (b) Average group entry probability (left) and latency (right) during three task periods: S+ (light blue), S- (dark blue) and the ITI (gray) throughout training in 5-trial bins. Data are expressed as mean ± SEM. (c) Cumulative performance index records on S+ (left) and S- (right) trials (see “1.5.5. Behavioral indices” and Figure 3 for an explanation of this index). Each animal’s behavior is represented by a black line and labeled with a gray letter. (d) Identification of the trial in which the consistent change in behavior took place (see “1.5.6. Change point analysis” and Figure 4). The black line on each panel is the cumulative record of the performance index of each individual animal. Small blue dots mark the abrupt changes in the slope of the line detected by the algorithm. The vertical red line marks the change point (CP). Vertical gray lines indicate the transition between sessions. (e) Same as "(b)" but with trials aligned to the CP trial. (f) Comparison of average S+ performance index (**p < 0.01, t-test), probability (**p < 0.01, t-test), latency (**p < 0.01, t-test) and ITI pseudolatency (p = 0.86, t-test) before vs. after change point. Details of all statistical tests are in “4.2. Supplementary Table”.
Figure 7. Cue-evoked excitations increase before CP and continue to escalate as cued approach becomes more vigorous. (a) Firing rate in Z scores (Z sc.) around S+ (light blue) or S- (dark blue) onset, trials aligned to CP. Mean (line) ± SEM (shaded area). (b) Activity (mean±SEM) 100-400 ms after S+ (light blue) or S- (dark blue) onset around CP (five-trial bins). (c) Same as “(b)” but median and interquartile range are shown in 40-trial bins. Numbers indicate sample size. Firing after S+ was higher than after S- on most bins. S+-evoked excitations increased after CP (*p < 0.05; **p < 0.01; ***p < 0.001, Wilcoxon). (d) The proportion of cue-excited neurons (solid blue) began to increase before CP. It continued to increase while the proportion of cue-inhibited neurons (white) declined (*p < 0.05, Fisher). (e) Same as "(c)" but for cue-excited neurons only. S+-evoked excitations began to increase before CP. (f) Activity in the 100-400 ms post-S+ window strongly correlates with S+ latency (top), S+ entry probability (middle) and S+ performance index (bottom) (**p < 0.001, Pearson). Sessions before/after CP are shown in light/dark purple, respectively.
Figure 8. NAc firing patterns during rewarded entries switch from excitations (before CP) to inhibitions (after CP). (a) Colors indicate average activity around S+ (left) or S+ entry (right) in 50 ms bins, in blocks before, during or after CP session (see legend). Neurons are sorted within each block based on the magnitude of their post-S+ response. (b) Firing elicited by S+ onset and S+ entry is strongly correlated. The relationship changes from positive before CP (top) to negative after CP (bottom). Z sc.: Z scores.
2.2. Experiment 2. Transient contribution of NAc core NMDARs to the expression of cued approach behavior and NAc cue encoding

2.2.1. Introduction

If the parallel emergence of strong cue-evoked excitations and vigorous cued approach (Experiment 1) means that the changes in cue-evoked firing are causal to learning, then preventing the acquisition of cue-evoked excitation over the course of successive trials should also prevent acquisition of cued approach. Given that NMDARs are necessary for both LTP of excitatory synapses in the NAc (Floresco, Blaha, et al., 2001a; Ji & Martin, 2012; Kombian & Malenka, 1994; Pascoli et al., 2012; C. M. Pennartz et al., 1993; Popescu et al., 2007; Yagishita et al., 2014); section “1.4.1.” and acquisition of reward-conditioned behaviors (Di Ciano & Everitt, 2001; Hernandez et al., 2005; Kelley et al., 1997; Smith-Roe et al., 1999; section “1.3.3.2.”), bilateral injections of an NMDAR antagonist into the NAc core during training should prevent both the emergence of cue-evoked excitations and cued approach learning. A preliminary study confirmed those predictions (Supplementary Figure 5), but the interpretation of those results is problematic. NMDAR blockade may have suppressed exploratory and motivated behavior earlier in training (Maldonado-Irizarry & Kelley, 1995; Maldonado-Irizarry & Kelley, 1994), precisely at the time when NMDAR blockade disrupted learning (section “1.4.2.2.”). If that were the case, the question of whether NMDAR blockade in the NAc impairs learning specifically, aside from potential performance effects, would have to be reexamined.

To test the hypothesis that the contribution of NMDARs to the expression of motivated behavior is transient, the behavioral and neuronal effects of NAc NMDAR blockade after moderate/extended training were examined. Intra-accumbens AP5 injections after moderate-but not extensive-training reduced both the magnitude of cue-evoked excitations in NAc neurons and the strength of the cued approach behavior.
2.2.2. Materials and methods

2.2.2.1. Cannulated microelectrode arrays: construction and intracranial implantation

The construction of cannulated microelectrode arrays and the surgery conducted to implant them into the NAc followed the same procedures described in Experiment 1.

2.2.2.2. Combined microinfusions and electrophysiological recordings

In this experiment (as well as in Experiment 4), NAc core electrophysiological recordings were combined with local microinfusions of either a vehicle or AP5 solution. Figure 9 shows how all the elements (cannulated arrays, injectors, fluid lines, recording cable and connectors) were arranged on the animal’s head. Infusions took place inside the operant chamber, with the animals already tethered to the recording system. The rate of infusion was slow (0.55 µl infused over 12 min) to facilitate simultaneous injections and recordings without loss of signal (du Hoffmann et al., 2011). The setup for these infusions consisted of two 33-gauge microinjectors that were affixed to two polyethylene tubes filled with mineral oil. The tubing was connected to a two-channel fluid swivel in the center of the roof of the chamber that allowed the animal to move freely without the tubes getting tangled. The top part of the swivel was connected to two other pieces of polyethylene tubing that ended at the tips of two 5 µl Hamilton syringes placed in a microinjection pump that stood atop the chamber. On days on which animals received injections while electrophysiological signals were being recorded, microinjectors were pre-loaded with the appropriate injectable solution before the session. The point where the water-based solution met the mineral oil was marked to enable post-hoc verification of the injection. Access to the reward receptacle was blocked with a metallic sheet before the animal was put in the chamber. This was done to prevent extinction of receptacle entry responses, which could delay learning, particularly early in training. The animal
was gently handled and restrained, injectors were inserted into its cannulae, the recording cable was
plugged into the connectors at the back of the rat’s headpiece and the polyethylene tubing was taped
onto the recording cable in a way that applied downward pressure on the injectors and locked them
in place inside the cannulae. After the microinjection pump stopped, the mark in the fluid lines was
checked to ensure that the injection had been properly delivered, the reward receptacle was
reopened, the door of the operant chamber was closed and the session started. During the test
session, infusions were delivered 30 min into the session. In this case, the mark on the fluid lines
was examined after the session ended.

2.2.2.2. Training protocol

The pretraining phase in this experiment was similar to the pretraining phase in Experiment 1. The task used during training was similar to the one used in Experiment 1 with some minor
modifications. The task used in Experiment 1 facilitates quick learning, but the sessions are long
(~1h 20min). In Experiment 4, we tested the effects of AP5 on learning while we recorded from the
NAc during training, which required adjusting the task to ensure that the effects of the drug lasted
throughout the whole session. Although this was not a concern the present experiment, training
animals with the exact same protocol as in Experiment 4 was important for reasons that will become
clear later. To shorten the session, every day, animals were given 35 S+ and 35 S- trials (instead of
40; randomly interleaved) and the intertrial interval was always chosen from an exponential
distribution truncated at 15-45 s (instead of 20-100 s). The downside of these modifications is that
they may slow down acquisition. To facilitate learning under these circumstances, two other
variations were introduced: a) the first trial was always an S+, it was presented 5 s after session
onset, lasting for a maximum of 60 s; b) when animals entered the compartment before S+ offset,
the cue remained on for two additional seconds rather than turning off immediately (Figure 10 a).

87
Figure 9. Setup for simultaneous electrophysiological recordings and microinjections. Connectors and cannulated microelectrode arrays are cemented onto the rat’s skull during surgery. Before each behavioral session, microinjectors connected to the infusion pump via fluid lines are inserted into the guide cannulae. A recording cable connected to the recording system is plugged into the connector. The cable ends in a headstage with two LED lights that allow the experimenter to automatically record the rat’s movements on camera (video tracking data was not examined in this study).
2.2.2.4. Microinfusion schedule

During the training phase, animals were trained in a simple cued approach paradigm (Figure 10a) for either seven (“moderate training”) or 18 (“extended training”) days (Figure 10b). In the “moderate training” group, animals underwent cannulated microarray implantation surgery, a week of postoperative recovery and a week of food deprivation (13 g of chow per day) before pretraining and training began. During training, all animals received bilateral microinjections of vehicle prior to each of the seven daily sessions. Because the quality of recordings can decline over time, surgery and recovery in the “extended training” group took place after 15 days of training. These animals had received a week of food deprivation and pretraining before the beginning of training. Once they recovered from surgery, they were food deprived for a week. After that, they received three additional training sessions to ensure that their levels of responding were back to preoperative levels. Bilateral infusions of a vehicle solution were given prior to the last “reminder” session to habituate animals to the infusion protocol.

The day after the last training day, animals underwent a 2 h test session. The protocol used during the test session was similar to the one used during training, with the only difference that 30 minutes into the session, either a vehicle or an AP5 solution was microinjected into the NAc. In the “moderate training” group, animals were randomly assigned to a group that received bilateral infusions of either saline (“VEH”; N=6) or AP5 (“AP5”; N=5) during the test session. In the “extended training” group (N=5), a within-subject design was implemented instead (i.e., the rats that received bilateral infusions of saline during the first test session were given bilateral AP5 on the second test session and vice versa).
2.2.3. Results

Moderate and extended training groups showed similar S+ vs. S- discrimination prior to the test session (Figure 10c,e, pre-injection period). In the “moderate training” group, AP5 infusions significantly reduced S+ performance index (Figure 10c) and S+ entry probability, while increasing S+ latency and ITI pseudolatency (Supplementary Figure 6a-c). These behavioral changes were accompanied by a decrease in cue-evoked excitations in NAc core neurons (Figure 10d) with no change in baseline firing rates (Supplementary Figure 6g-h).

These effects were not observed in the “extended training” group. After having experienced additional training sessions, cued approach (Figure 10e; Supplementary Figure 6d-f) and cue-evoked excitations (Figure 10f) were unaffected by AP5. Hence, the implication of accumbens NMDAR in performance is transient. Notably, the magnitude of cue-evoked excitations in the absence of the drug was much smaller than that seen in the “moderate training” group (Figure 10d, f).

Importantly, although previous studies have used the resistance of well-learned appetitive responses to NAc NMDAR blockade as evidence that the antagonist acts specifically to inhibit learning during task acquisition (Di Ciano et al., 2001; Hernandez et al., 2005; Kelley et al., 1997; section “1.4.2.2.”), our results indicate that performance is, in fact, dependent on NMDARs during early training. The specific role of accumbens NMDARs in the acquisition of cued approach behavior will be examined in Experiment 3.
Figure 10 NMDARs play a transient role in the expression of NAc cue-evoked excitations and the vigor of the subsequent approach response. (a) Task. ITI: intertrial interval (b) Microinjection schedule. (c) For the “moderate training” group, performance index in S+ (solid) and S- trials (empty squares) before and after saline (VEH, blue) or AP5 (red) microinjection (mean±SEM). The gray rectangle represents the microinjection period. In S+ trials, AP5-treated subjects performed worse than subjects in the control group in every post-injection 30 min bin (***p < 0.001; ** p < 0.01; t-test). (d) Firing rate (FR; mean±SEM) around S+ onset in the "moderate training" group before (gray) or after injection of saline (blue, left) or AP5 (red, right). Insets show activity 100-400 ms after S+ before (“Pre”) or after (“Post”) injection. Boxes show the interquartile range (box height), mean (black line) and median (red/blue line). Activity was unaffected by saline (p = 0.64) but reduced by AP5 injections (***p < 0.001, Wilcoxon). (e-f) Same as (c-d) but for animals in the “extended training” group. In this group, performance was not disrupted by AP5 injections (p > 0.05, t-test). S+-evoked activity was also unaffected by saline (p = 1) or AP5 injections (p = 0.2, Wilcoxon).
2.3. Experiment 3: Effects of intra-accumbens NMDAR blockade on acquisition of cued approach behavior (above and beyond performance effects)

2.3.1. Introduction

Daily bilateral infusions of NMDAR antagonists into the NAc core during early training in naïve animals impair learning (Supplementary Figure 5; Di Ciano & Everitt, 2001; Hernandez et al., 2005; Kelley et al., 1997; Smith-Roe et al., 1999). However, they also affect exploratory behavior (Maldonado-Irizarry & Kelley, 1995; Maldonado-Irizarry & Kelley, 1994) as well as the expression of previously acquired cued approach behavior (Experiment 2). If daily bilateral intra-accumbens AP5 infusions were preventing animals from engaging in normal reward-seeking behavior, their failure to learn might be explained by a reduced number of opportunities to experience the cue-reward contingency. Simultaneously, animals whose engagement in the task is poor would have more nonreinforced exposure to the cue, which would promote latent inhibition of the cue and further delay learning (Lubow & Moore, 1959).

The goal of this experiment was to specifically examine the role of NAc NMDA receptors in learning by using an experimental design that addresses the abovementioned confound. In this study, animals in two groups received either infusions of an NMDA receptor antagonist (AP5 group) or saline infusions (yoked vehicle group) in the NAc core throughout training in a cued approach task. The behavioral paradigm was designed to ensure that both groups received the cue alone, or the cued paired with reward, the same number of times. As a result, any observed differences in learning between these two groups cannot be attributed to a performance deficit. By eliminating this confound, we observed that blockade of NMDA receptors in the NAc interferes with acquisition of cued reward approach behaviors independent of effects on performance.
2.3.2. Materials and methods

2.3.2.1. Cannula surgeries

Bilateral guide cannulae (27-gauge, Plastics One; Roanoke, VA) targeting the NAc core were chronically implanted for microinjection. The coordinates for the tips of the cannulae were, from bregma: AP: +1.2 mm, ML: ±2 mm, DV: -5.8 mm (the tips of the injectors protruded 2 mm from the tips of the cannulae, so the actual DV coordinate of the injection was -7.8 mm). In order to prevent debris from clogging the cannulae, steel obturators and dust caps (Plastics One) were used.

2.3.2.2. Training protocol

Pretraining. After a week of food restriction and regular handling, animals were habituated to 10% sucrose in their home cage for at least 2 h. After that, they received a mock infusion in which two injectors were bilaterally inserted and remained in place for ~30 s. Before training, all rats also received a receptacle training session. Animals in this experiment received receptacle training in an alternative context that consisted of a standard 30 x 25 cm operant chambers with a grid floor and dark metallic walls (Med Associates; St. Albans, VT). During this session, receptacle entries were rewarded on an FR1 schedule with a 10 s post-reinforcement timeout until 20 rewards were delivered. If animals failed to collect all the rewards within ~45 min, they were given a second receptacle training session the next day. On this second session, the receptacle was baited with sucrose before putting the animal inside the chamber.

After receptacle training, animals in this experiment were also given two 30 min “context habituation” sessions in the chambers that would be used for subsequent training. During these sessions, the white noise generator and the house lights were on but no cues or rewards were presented. The goal of these sessions was to extinguish contextual cues and thereby promote quick
acquisition of the cue /reward association (Balsam & Schwartz, 1981).

Training. Each trial consisted of an inter-trial interval (ITI) of 30 s on average (exponentially distributed; min= 15 s, max= 45 s) and an auditory cue (“S+”) that consisted of a “siren” alternating between 4-8 kHz (Figure 11a). A head entry into the reward receptacle during presentation of the cue resulted in delivery of a droplet (~ 150 μl) of 10% sucrose into the same compartment (correct trial) and offset of the cue. If the animal failed to enter the compartment during presentation of the cue (missed trial), the tone would terminate after 10 s and sucrose would not be delivered. Entries into the compartment during the ITI had no programmed consequences. Each daily session ended after 30 min. No other cues were included in this version of the task because doing so would make it very hard to yoke the training experience across groups.

Drug-free extinction test. The day after the training phase ended, all animals were given saline microinjections prior to the test, which consisted of a 30 min extinction session. The task protocol on this session was the same one that was used during training, with the exception that only cues, never rewards, were presented. This prevented within-session learning, allowing us to test the degree to which animals had learned to respond to the reward-predictive cue during the training phase.

2.3.2.3. Yoking procedure

Animals received daily training sessions for six consecutive days and, prior to each session, they were given bilateral microinjections of either AP5 (AP5/VEH group; N=7) or 0.9% saline (VEH/VEH group; N=7) targeted at the NAc core (Figure 11b). During training, both groups were matched (“yoked”) in terms of their experience with the cue regardless of their ability to engage in the task and seek rewards. A sample session explaining the yoking procedure is depicted in Figure
11c. Every day, each animal in the AP5/VEH group was trained for 30 min. The average number of correct trials (“paired cues”) and missed trials (“unpaired cues”) was calculated. Animals in the VEH/VEH and AP5/VEH groups experienced, on average, the same number of paired and unpaired cues on each session. This was achieved by blocking access to the reward receptacle once each VEH/VEH animal had collected the same number of rewards the AP5/VEH group had collected, on average, on that session. Once the receptacle was closed, each animal in the VEH/VEH group continued to experience cues (same ITI). For animals in the VEH/VEH group, the session ended once the total number of unpaired cues they experienced matched the total number of unpaired cues experienced by the AP5/VEH group on that day’s session. The distribution of paired and unpaired cues experienced during each training session by each group is shown in Figure 11d.

2.3.2.3. Microinjections

Prior to each session during the training and test phases, all animals received bilateral intracranial microinjections. During the training phase, animals in the experimental group received daily AP5 injections (1 \( \mu \)g/0.5 \( \mu \)l/side), whereas animals in the “vehicle” group received daily injections of 0.9% saline (0.5 \( \mu \)l/side). On test day, all animals were given saline injections. Drugs were delivered using microinjectors (33 ga, Plastics One) that extended 2 mm below the base of the guide cannulae, targeting the center of the NAc core. Polyethylene tubing filled with mineral oil connected the microinjectors to two 1 \( \mu \)l syringes (Hamilton; Reno, NV) mounted in a microinjection pump (KD Scientific; Holliston, MA). After both microinjectors were inserted into the guide cannulae, the pump was turned on, fluid was infused for 2 min (0.25 \( \mu \)l/side/minute), and microinjectors were left in place for 1.5 min after the end of the injection to allow diffusion of the solution. Animals were then placed immediately in the operant chamber and the session started.
2.3.3. Results

In this experiment, rats were trained in a simple cued approach behavior task. Only if animals made an entry into the reward receptacle before the end of the auditory S+, reward was delivered (Figure 11a). To ensure that both groups had a similar experience with the S+ regardless of their ability to engage in the task, the average number of “reward-paired cues” and “reward-unpaired cues” presentations was experimentally matched across both groups (Figure 11b-d). To facilitate yoking the control group to the experimental group, no S-‘s were introduced in this version of the task.

During training, VEH/VEH rats needed less trials to collect as many rewards (“paired cues”) as the AP5/VEH rats did, on each session, on average (Figure 11d). After the VEH/VEH rats collected the target number of rewards, the reward receptacle was blocked while this group was exposed to the “unpaired cue” as many times as it was needed to match the AP5/VEH group (Figure 11d).

In the drug-free extinction test, despite having a comparable experience with the cue-reward contingency, animals that had been under the influence of AP5 during training showed a reduced likelihood of entering the receptacle during the S+ than control animals, both throughout the whole session (Figure 11e, left) and during the first five minutes (Figure 11e, right). Because VEH/VEH animals also exhibited a higher frequency of receptacle entry during the ITI (not shown), it is important to check whether their entry behavior was modulated by the cue at all. Critically, animals in the VEH/VEH group also exhibited a significantly higher S+ performance index than those in the AP5/VEH group (Figure 11f), indicating that the control group learned to respond specifically to the S+ rather than just increasing its responding irrespective of cue presentation.
Figure 11. NMDARs in NAc are necessary for the acquisition of cued approach behavior. (a) Task. ITI: intertrial interval. (b) Microinjection schedule. (c) ‘Yoking’ procedure (sample session). Each AP5/VEH or VEH/VEH subject is identified by a letter. Bubbles indicate trials during which subjects made a cued entry (black) or not (white). Red boxes indicate trials when the receptacle was closed. (d) Mean±SEM number of ‘Paired’ (black) and ‘unpaired cue’ trials (white) experienced on each session by the AP5/VEH or VEH/VEH group. Red bars indicate trials during which the receptacle was closed. Both groups experienced the same number of ‘paired’ and ‘unpaired cue’ trials daily. (e) Mean±SEM S+ entry probability in the VEH/VEH (blue) and AP5/VEH (red) groups on the test (drug-free) session (whole session, left; first five minutes, right). No rewards were delivered. There were main effects of the drug, bin and drug x bin interaction (Supplementary Table 1). Box plots show median, interquartile range and individual values. S+ entry probability was lower in previously AP5-treated subjects (***p < 0.001, t-test). (f) Same as "(e)" but for S+ performance index. There were main effects of drug and drug x bin interaction. During the first five minutes, animals that received AP5 injections during training exhibited lower S+ performance index (***p < 0.001; t-test).
2.4. Experiment 4: NMDAR contribution to the emergence of training-induced changes in NAc activity.

2.4.1. Introduction

In the NAc core, cue-evoked excitations predictive of cued approach behavior emerge during training (Experiment 1). Neurons in the BLA undergo plasticity during appetitive learning, and they also exhibit cue-evoked excitations after just a few trials of training (Paton et al., 2006; Schoenbaum et al., 1999; Tye et al., 2008). Because BLA neurons send excitatory projections to the NAc (section “1.2.1.”), learning-related signals acquired by NAc neurons during training may not be the result of synaptic strengthening at the level of the NAc. Instead, it is possible that the plasticity that underlies the emergence of cue-evoked excitations takes place upstream of the NAc and that neurons in the NAc just passively relay these signals to downstream motor output structures. LTP in the NAc core depends on NMDARs (section “1.4.1.”) and, in behaving animals, activation of NMDARs during training is necessary for learning cued approach behavior even when controlling for a possible performance effect (Experiment 3). Hence, we hypothesized that NMDARs would mediate the growth of cue-evoked excitations observed in the NAc during training.

Because bilateral AP5 injection blocks not only acquisition of cued approach behavior (Experiment 3), but also its expression (Experiment 2), the impact of the antagonist on NAc unit activity during early cued approach training is likely to reflect some combination of learning and performance deficits. In previous studies (Caref & Nicola, 2018; du Hoffmann & Nicola, 2014), unilateral injections of performance-impairing agents into the NAc only minimally disrupted performance. We therefore examined NAc unit activity in rats microinjected with AP5 unilaterally prior to each daily training session, allowing us to compare learning-related activity in the NAc in
AP5- vs. vehicle-infused hemispheres within individual animals. To ascertain if intra-accumbens NMDAR blockade affects the emergence of cue-evoked excitations by preventing synaptic plasticity and not just by suppressing ongoing excitatory transmission, a group of animals in this group was given a drug-free extinction test after the training phase. The results of this experiment suggest that NMDAR-dependent plasticity within the NAc core is necessary for these neurons to acquire the firing patterns that underlie the acquisition of vigorous cued approach behavior.

2.4.2. Materials and methods

2.4.2.1. Cannulated microelectrode arrays: construction and intracranial implantation

The construction of cannulated microelectrode arrays and the surgery conducted to implant them into the NAc followed the same procedures described in Experiment 1.

2.4.2.2. Training protocol

The pretraining phase in this experiment was similar to the one in Experiment 1 and Experiment 2. The task used during training was similar to the one used in Experiment 2 (Figure 10a). After six days of training, some of the animals received an additional drug-free extinction test session. The protocol used in this test session was similar to the one used during the training sessions with the exception that only cues, never rewards, were presented.

2.4.2.3. Combined microinfusions and electrophysiology

The procedure used to perform local NAc core infusions while recording from this structure was similar to the one described in Experiment 2.
2.4.2.4. Microinfusion schedule

The training phase consisted of six consecutive daily sessions. Animals (N=17) were given AP5 microinjections into the same hemisphere prior to each training session; vehicle was injected into the contralateral hemisphere (Figure 12a). On the seventh day, some of the rats (N=7) were given a test extinction session prior to which bilateral vehicle infusions were administered.

2.4.3. Results

Of 17 animals, 11 acquired cued approach behavior within six days of training as evidenced by the detection of a behavioral CP for S+, but not S- trials (Figure 12b). The average performance aligned to the CP trial (Figure 12c) revealed that the learning profile of these 11 rats was comparable to that of uninjected animals (Figure 6e). The S+ performance index and S+ entry probability were significantly higher after the CP than before, whereas S+ latency was significantly shorter (Figure 12d). The remaining six rats were classified as “non-learners” based on the absence of a detectable CP (Supplementary Figure 9).

In the 11 “learners,” interhemispheric differences in the emergence of cue-evoked excitations cannot be explained by a performance deficit. Therefore, we compared the emergence of cue-evoked excitations in the AP5-injected vs. vehicle-injected NAc (Figure 13; N=313 and 273 neurons from vehicle and AP5 hemispheres, respectively; firing data from non-learners is shown in Supplementary Figure 9). Statistical comparisons across bins and conditions followed the same rationale as in Figure 7 (see “1.5.7. Statistical Analyses”).

In both hemispheres, S+-evoked excitations were significantly larger than those elicited by the S-. This dissociation emerged many trials before a behavioral CP was detected (Figure 13a,b), suggesting that AP5 did not prevent NAc core neurons from encoding the reward-predictive value of the cue in the earliest stage of training (Stage 1, Figure 15). However, just prior to the CP (Stage
2, Figure 15), the proportion of neurons classified as S+-excited became significantly larger in the vehicle side than the AP5 side, and this proportion remained larger in all trial bins until the end of training (Figure 13d). The proportion of S+-inhibited neurons diminished before the CP in the vehicle but not in the AP5 side, and it did not significantly decrease after that in either hemisphere (Figure 13d). Even though AP5-treated neurons were less likely to become cue-excited, when they did, their firing rate in the 100-400 ms window was lower than the firing rate of neurons in the vehicle side – but not significantly so (Figure 13f, left). However, after the CP (during Stage 3, Fig 9), cue-excited neurons in the AP5-treated side were significantly less excited in the 750-2000 ms window after S+ onset than the corresponding neurons in the vehicle side (Figure 13f, right). Thus, in the absence of behavioral effects of AP5, the drug disrupted the increase in S+-evoked excitations occurring just prior to and after the CP.

Neurons in the vehicle-infused side showed growth of S+-evoked excitations after the CP (Stage 3, Figure 15), paralleling the increasing probability and decreasing latency of the S+-elicited approach response. While neurons in the AP5-treated hemisphere of the same animals showed S+-evoked excitations before the CP (Stages 1-2, Figure 15), the magnitude of these excitations did not grow after the CP (Figure 13a,b; Stage 3, Figure 15). Direct comparison between hemispheres confirmed that significant differences in the magnitude of S+-evoked activity caused by AP5 appeared after the CP (Figure 13c). This was true for the post-S+ 100-400 ms window as well as for the 750-2000 ms window, indicating that neurons in the vehicle- but not AP5-treated side remained excited for an extended period after presentation of the reward-predictive cue (Figure 13c; Figure 7e, left). Thus, NMDARs are required for both the growth in peak S+-evoked excitation and for the extended tail of these excitations.

Baseline firing rates did not differ across vehicle- and AP5-infused hemispheres (Supplementary Figure 7), nor did the hemispheres differ in terms of firing rates during non-reward
periods (S- or ITI; Supplementary Figure 8) or after receptacle entries following the S+ (Supplementary Figure 8). In summary, blockade of NMDARs in the NAc core during training disrupted the increase in S+-evoked excitations that normally accompanies increases in the vigor of the cued approach without affecting neuronal firing during other periods of the task.

Because intra-accumbens blockade of NMDARs disrupts the expression of cue-evoked excitations (Figure 10), it is possible that AP5 simply reduced ongoing excitatory transmission without impairing learning-related plasticity. To test whether the drug had long-lasting effects on NAc cue-evoked excitations, we conducted a drug-free test session in a subset (n=7) of animals and compared the magnitude of cue-evoked excitations in the previously AP5-treated hemisphere with those in the previously saline-treated hemisphere. Saline was infused into both hemispheres just prior to the test session and, to prevent cue-reward learning, no rewards were delivered.

Two of the seven rats were “non-learners” during training (Supplementary Figure 9). The remaining five rats (“learners”) showed significantly higher S+ than S- performance in the drug-free extinction test (Figure 14a). Critically, neurons in the hemisphere that had been treated with saline infusions during training showed significantly higher firing rates in response to the S+ compared to neurons in the formerly AP5-treated hemisphere. These intra-hemispheric differences in S+-evoked activity were observed both during the 100-400 and the 750-2000 ms windows after cue onset (Figure 14b,c). This effect was not due to differences in baseline firing rate across hemispheres (Supplementary Figure 7). Therefore, the reduction in cue-evoked excitations produced by AP5 during training outlasted the presence of the drug in the brain. Together, these observations indicate that NMDARs in the NAc are required for the long-lasting plasticity that causes the growth of large-magnitude cue-evoked excitation in NAc neurons. In turn, the NMDAR-mediated emergence of cue-evoked NAc responses drives the development of vigorous cued approach during learning.
Figure 12. Unilateral blockade of NMDARs in NAc core did not interfere with learning in many subjects. (a) Microinjection schedule. VEH: vehicle. (b) Cumulative performance index records on S+ (left) and S- (right) trials in 11 animals that, despite receiving unilateral intra-accumbens AP5 injections during training, were able to acquire cued approach behavior. Each subject’s behavior is represented by a black line. Letters identify different subjects. (c) Left: Mean±SEM entry probability during the S+ (light blue), S- (dark blue) or pre-cue 10 s ITI window (gray) with respect to the trial in which the change point occurred. Before change point, the overall rate of receptacle entry shows a steady increase until, at the change point, the rate of cued entry continues to increase while the rate of uncued entry stabilizes. Right: same as left panel but for latency and ITI pseudolatency. (d) In subjects that received unilateral AP5 injections prior to each training session, S+ performance index and S+ entry probability increased after change point (**p < 0.001; t-test). S+ latency declined (**p < 0.001; t-test) while the ITI pseudolatency did not change after change point (p = 0.241; t-test).
Figure 13. In the AP5-treated hemisphere (but not in the vehicle hemisphere), large and prolonged NAc cue-evoked excitations failed to emerge during training. (a) Mean±SEM excitation 100-400 ms after S+ (solid) or S− (empty) in the saline (VEH, blue) or AP5 (red) hemispheres in five-trial bins around CP. Z sc.: Z scores. (b) Same as “(a)” but median and interquartile range are shown in 35-trial bins. Activity was higher after S+ than S− onset at least 70 trials before CP in both hemispheres (**p < 0.01, Wilcoxon). After CP, S+-evoked activity increased in the vehicle (**p < 0.01) but not the AP5-treated side (p > 0.05, Wilcoxon). Numbers indicate sample size. (c) Comparison of post-S+ activity across hemispheres. After CP, population activity was higher in the vehicle than in the AP5-treated side in the 100-400 ms and 750-2000 ms post-S+ windows (*p < 0.05; ***p < 0.001, Wilcoxon). (d) Proportion of excited (top) or inhibited (bottom) units after S+ onset around the CP. The proportion of S+-excited units was higher in the vehicle (blue) than in the AP5 side (red) right before CP and remained higher thereafter (*p < 0.05; ***p < 0.001, Fisher). (e) Mean±SEM activity around the S+ of all (left) or cue-excited (right) units in the vehicle (blue) or AP5 (red) sides. (f) Same as “(c)” but for cue-excited neurons (*p < 0.05; **p < 0.01; ***p < 0.001, Wilcoxon).
Figure 14. The effects of AP5 on NAc firing outlast the presence of the drug in the brain. Data from drug-free extinction session of rats that received unilateral AP5 infusions during training and learned the task (Learners). VEH: vehicle. (a) Mean±SEM performance on S+ (light blue) and S- (dark blue) trials. Main effects of cue and cue x bin interaction (Supplementary Table 1). (b) Mean±SEM firing activity in Z scores (Z sc.) around S+ onset in NAc neurons treated with saline (blue) or AP5 (red) during training. Inset: proportion of neurons that were classified as cue-excited in the side previously treated with saline (blue) or AP5 (red) (p = 0.31, Fisher). (c) Activity (median and interquartile range) 100-400 ms or 750-2000 ms after S+ onset in the NAc that had been treated with saline (blue) or AP5 (red) during training. All neurons (left) vs. cue-excited neurons (right).
3. DISCUSSION
3.1. Summary of the results.

Learning to expect and seek natural rewards in the presence of predictive environmental cues is essential for survival. NAc neurons receive massive convergent cue- and reward-related information, and project to motor output regions (Mogenson et al., 1980). Excitatory synapses in the NAc can undergo LTP (Floresco et al., 2001; Goto & Grace, 2005; Kombian & Malenka, 1994; LeGates et al., 2018; Pennartz et al., 1993; Popescu et al., 2007), and in vivo studies have implicated the NAc in the acquisition of reward-reinforced behavior (Day & Carelli, 2007; Di Ciano et al., 2001; Gerdjikov et al., 2007; Hernandez et al., 2005; Kelley et al., 1997; Saunders et al., 2018). Because of these findings, it is often assumed that plasticity in the NAc is causally linked to learning of cued reward-seeking behavior. Indeed, addictive behavior is often regarded as a pathological form of conditioned behavior that is acquired because drugs of abuse usurp plasticity mechanisms in the NAc (Lüscher & Malenka, 2011). Yet, the specific NAc-dependent mechanism underlying natural learning of cued reward-seeking behavior has not been identified. This work provides insight into this mechanism by demonstrating that the NAc cue-evoked neuronal firing response that drives cued approach to reward develops prior to expression of the learned behavior, and that this process requires NMDAR-dependent plasticity in the NAc.

The experiments presented here reveal that cued approach task acquisition occurs in distinct stages (Figure 15). In Stage 1, small S+-evoked excitations differentiate the S+ from the S- from the first training trials, well before learning is manifested behaviorally (Supplementary Figure 2b). Intriguingly, these excitations were unaffected by intraaccumbens NMDAR blockade (Figure 13c-f), suggesting that the plasticity necessary for these excitations to appear must be taking place somewhere upstream of the NAc. During Stage 2, S+-evoked excitations in NAc neurons begin to grow a few trials before the CP. This growth continues in Stage 3 as S+ responding becomes more
reliable and vigorous (Figure 15), corresponding with increasing vigor of the cued approach (Figure 7f).

Daily bilateral NMDAR antagonist injections into the NAc prevented the emergence of both cued approach learning and NAc excitations (Supplementary Figure 5). It is likely that these suppressive effects were due, at least in part, to a contribution of NAc NMDARs to the expression of learned behavior (and not merely to its acquisition), as the antagonist reduced both cued approach performance and the cue-evoked excitations of NAc neurons even in animals that had been trained to distinguish the S+ from S- prior to drug infusions (Figure 10c,d). However, S+-evoked excitation and behavioral performance in overtrained animals was no longer dependent on NMDARs (Figure 10), suggesting that overtraining induces a secondary plastic change within the NAc core (Stage 4, Figure 15).

The presence of an antagonist-induced performance deficit does not rule out an additional effect on learning. A performance deficit during early training would likely disrupt learning by limiting the number of paired encounters with the cue and reward. However, antagonist-injected animals exhibited severe learning deficits compared to yoked controls that received no drug but an equivalent number of cue-reward pairings (Figure 11). Therefore, drug-induced reduction in the number of learning trials cannot account for the reduced learning seen in animals under the influence of the NMDAR antagonist. Together, these results suggest that daily NMDAR antagonist treatment induced both a performance deficit and an additional learning deficit.

To isolate the contribution of NAc NMDARs to learning and the emergence of cue-evoked NAc excitations, daily NAc injections of NMDAR antagonists were given unilaterally, which resulted in a normal time course of learning in most animals (Figure 12). Simultaneously, NAc neuronal activity was recorded in both antagonist- and vehicle-injected hemispheres. NMDAR
blockade prevented the growth of cue-evoked excitations that would have otherwise started near the CP (Figure 13). Reduced cue-evoked firing in the drug-injected hemisphere persisted even in the absence of the drug (Figure 14b,c), suggesting that NAc NMDAR transmission mediates the plasticity necessary for NAc neurons to undergo the experience-dependent physiological changes that give rise to cued approach behavior.
Figure 15. Model of the stages of learning cued approach behavior and the mechanisms implicated in the emergence and expression of learning-related S+-evoked excitations in NAc neurons. Top box: likelihood/vigor of approach responses prior to training (left box) and at different points of training (right box) in the presence (light blue) or absence (dark blue) of the S+. Bottom box: conceptual diagrams depicting, at different stages of learning, changes in the strength of S+-evoked excitatory responses in NAc neurons as well as NAc afferents that encode either the S- (left) or the S+ (right). Early in training (Stage 1), small S+-evoked excitations appear and are unaffected by AP5, likely implicating plasticity elsewhere in the circuit in which the NAc is embedded (likely upstream of the NAc). Just before CP and after CP (Stages 2 and 3), NMDAR-mediated plasticity within the NAc is required for the growth of NAc cue-evoked excitations. Between the trials before CP and the point at which animals are showing asymptotic performance, expression of S+-evoked excitations in the NAc requires NMDAR-mediated excitatory transmission, but it becomes independent from it with extended training (Stage 4).
3.2. General discussion

3.2.1. How experience shapes behavior in this paradigm

Cued approach to reward is a simple but essential adaptive behavior. Despite its simplicity, the results presented here illustrate that learning to perform cued approach is a complex, multi-step process. In these experiments, animals were trained in a task in which cues signaling reward availability in a receptacle (S+) and control cues (S-) were presented at long intervals (15-45 s or 20-100 s); receptacle entry during the S+ both triggered reward delivery and terminated the S+, and reward was delivered only during S+ presentation. The most striking indication that animals learned the reward-predictive value of the S+ was the sudden, stable increase in the S+ performance index, a composite measure that indicates the degree to which S+ responding occurs more quickly than predicted by the rate of responding during the ITI. The trial on which this sudden increase (“change point”, CP) occurred varied across animals (Figure 6c,d). CP analysis indicates the point at which an animal’s performance first clearly reflects the learned relationship between the S+ and reward availability, usefully providing a single trial to which electrophysiological data can be aligned in order to investigate the neural changes that underpin learning.

Despite the discrete improvement in performance marked by the CP, animals showed evidence of gradual learning both before and after the CP. Initially, subjects gradually increased their overall frequency of receptacle entry without respect to the presence of cues. This effect is shown in the parallel growth of receptacle entry probability during the S+, S- and ITI, as well as the parallel decreases in the latency to make this response after onset of the S+ and S-, and during the 10 s window that preceded the cue (“ITI pseudolatency”; Figure 6e). This indiscriminate escalation in responding is likely a consequence of learning that reward is intermittently available in the receptacle. In the next stage, the latency to enter the receptacle after S+ presentation continued to
gradually decrease while the S- entry latency and ITI pseudolatency stabilized or slightly increased (Figure 6e). This specific increase in relative responsivity to the S+ is indicative of learning the reward-predictive significance of the S+ and, maybe, the value of responding quickly. Notably, the CP occurred just when S- and ITI responding stabilized. Because only the S+ response latency declined further beginning with the trials near CP, the learning mechanism responsible for this progressive change is likely different from that which drives the growth of indiscriminate responding prior to CP. Thus, these behavioral observations suggest that acquisition of this task is comprised of several processes: prior to the CP, animals learn that reward is often available in the receptacle; and around the time of the CP, animals learn that the S+ reliably predicts reward availability in the receptacle. As evidenced by the decreasing S- performance index after CP (Figure 6c), animals also learn that reward is not available in the presence of the S-.

3.2.2. How experience shapes NAc activity: features, mechanisms and significance

3.2.2.1. S+-evoked excitations begin to grow prior to the appearance of S+-elicited approach

Concurrent measurements of NAc activity and cued approach responses during training revealed the relationship between experience-dependent changes in neuronal activity and the acquisition of appetitive conditioned behavior. In this kind of paradigm, the analysis of group neural data has traditionally been challenged by the extreme variability in learning rates. Previous studies aggregated data by trial or by session —ignoring potential interindividual differences in performance— and/or simply compared neural data across predetermined training stages during which performance is predictably homogeneous across animals (e.g., Day et al., 2007; Roitman et al., 2005; Setlow et al., 2003). In this project, this problem was bypassed by aligning the performance of each animal to the trial after which approach behavior was consistently and
detectably under the control of the S+. This approach led to the characterization, with a high degree of granularity, of the temporal relationship between the emergence of NAc cue-evoked excitations and cued approach behavior.

Large S+-evoked excitations began to emerge in NAc core neurons a few trials before the CP, and the prevalence, magnitude, and duration of these neuronal excitations continued to escalate as the learned response became more vigorous. This increase in NAc responsivity was not explained by the advancing anatomical placement of the probes because S+-evoked responses emerged when the probes were left in place across sessions (Supplementary Figure 4). These growing NAc activations did not reflect a non-specific learning process, such as increasing familiarity with the cue. Rather, they must signal an appetitive association because NAc responses to the S- remained unchanged throughout training (Figure 7a,c). Similarly, the phasic responses to the S+ did not just encode the likelihood of the subsequent approach response because NAc firing rate remained unchanged when the S- was followed by an approach response (Supplementary Figure 2f). Instead, through learning, NAc neurons acquired cue-evoked responses that simultaneously reported the motivational value of the cue and the vigor of the imminent cued response.

Critically, the detected escalation of NAc responsivity to the S+ did not follow the emergence of cued approach behavior. Instead, their growth anteceded the observable expression of cued behavior during training. These findings satisfy the “detectability” test (section “1.3.3.1.”), the first one of Martin et al.’s (2000) logical criteria for causally connecting neuroplasticity and learning. Although this correlation does not prove a causal link between NAc plasticity and acquisition of approach behavior elicited by a cue, it identifies NAc plasticity as an eligible candidate for neural substrate of this kind of learning.
3.2.2.2. Functional relationship between NAc cue-evoked excitations and the initiation of approach

The results from Experiment 1 suggest that growing S+-evoked excitations propelled the increase in cued approach behavior during training. However, for this to be true, NAc cue-evoked responses must actually be driving the approach behavior. Although none of the experiments were specifically designed to address this question, several observations support the causal link between NAc S+-evoked firing and S+-evoked approach responses:

a) In existing literature, NAc cue-evoked excitations are well established as an essential neural trigger of vigorous cued approach responses in animals performing similar tasks. In trained animals, the onset of S+-evoked excitations in the NAc precedes the onset of the approach response, and these neural signals predict the probability, latency and speed of the upcoming movement (McGinty et al., 2013; Morrison et al., 2017). In addition, cued approach behavior is disrupted when cue-evoked excitations are disrupted. This disruption has been achieved by reducing input from the BLA (Ambroggi et al., 2008), PFC (Ishikawa et al., 2008) and VTA (Yun et al., 2004) or by intra-NAc administration of dopamine antagonists (du Hoffmann & Nicola, 2014), µ-opioid antagonists (Caref & Nicola, 2018) or NMDAR antagonists (Experiment 2).

b) During training, the magnitude of NAc cue-evoked excitations was highly predictive of the likelihood and vigor of the cued approach response (Figure 7f). Even after the change point, S+-excitations were smaller in the rare occasion when the cue failed to elicit a receptacle entry compared to trials in which animals responded before cue offset (Supplementary Figure 2f, S+ responded vs. S+ missed columns).

c) NAc S+-evoked excitations were not observed in animals that failed to learn the task (Supplementary Figure 9), even in the vehicle side.

d) Bilateral NMDAR blockade in the NAc core prevented both the emergence of S+-
evoked excitations (Supplementary Figure 5d,e) and learning (Supplementary Figure 5a-c). Although this is at least partially explained by a performance deficit, this manipulation also produced deficits in learning that could not be accounted for by performance deficits (Experiment 3; Figure 11). This is another instance in which the magnitude of NAc activations was tightly associated with the strength of the behavior.

Based on these observations, it seems reasonable to assume that, on any given trial, cue-evoked excitations were not just a correlate or a consequence of performing the approach behavior and, instead, they were driving it. Consequently, the growth of cue-evoked excitations during training must have been what caused the increase in vigor and likelihood of cued approach responses.

These observations do not preclude the possibility that NAc neurons also contributed to increased indiscriminate responding prior to the CP and the stabilization or decrease in such responding at CP. Indeed, bilateral NMDAR antagonist injection in the NAc prevented not only the emergence of S+-evoked approach behavior but also the normal escalation of indiscriminate responding (Supplementary Figure 5a-c). And in moderately trained animals, intra-accumbens AP5 infusions decreased responding during the ITI (Supplementary Figure 6c). On the other hand, NAc encoding of ITI entries or entries during the S- did not show obvious changes during training (Supplementary Figure 3) and was not affected by NMDAR blockade (Supplementary Figure 8). However, firing activity in response to implicit cues perceived by the animal at variable times before receptacle entry could have driven indiscriminate responding, and such firing would be difficult to detect. Therefore, the role of NAc neurons in the control of receptacle approach behavior that is not under control of the S+ remains unclear.
3.2.2.3. Changes in the proportion of S+-inhibited neurons in the NAc

A small population of NAc neurons exhibited S+-elicited inhibitions (Figure 7d). Cue-evoked inhibitions in the NAc have been consistently reported in previous literature (e.g., Day et al., 2006; Krause et al., 2010; Morrison et al., 2017; Nicola et al., 2004a; Taha & Fields, 2006). The percentage of cue-inhibited neurons slightly decreased during training (Figure 7d), and blockade of NMDARs during training disrupted this decrease, but this effect was only significant prior to the change point (Figure 13d). The proportion of cue-inhibited neurons was small (~18% of all recorded neurons), and their temporal parameters (latency, duration) were more diverse than those of excitations, making it harder to characterize them by focusing on the fixed time window that captured the peak of cue-evoked excitations (100-400 ms). Also, cue-evoked inhibitions have a smaller dynamic range because baseline firing of NAc neurons is already considerably low (~0.5-5 Hz). Hence, it is possible that the failure to identify a significant decline after the change point is due to the low statistical power of the dataset and the limited sensitivity of the analyses, and that the results exclusively apply to very large cue-evoked inhibitions.

S+-inhibitions share many features with S+-excitations: they precede the onset of locomotion, they predict the probability and latency of approach and they are larger when animals are closer to a reward-associated lever, so they have also been implicated in initiating appetitive behavior (Morrison et al., 2017; Taha & Fields, 2006). To understand the mechanisms behind their evolution throughout training, it is critical to understand what kind of NAc neuron is cue-inhibited and where these inhibitions may come from. However, the source of these inhibitions is not clear. MSNs send GABAergic projections to each other (Taverna, van Dongen, Groenewegen, & Pennartz, 2004), so S+-inhibitions could be caused by inhibitory projections from other cue-excited MSNs (Morrison et al., 2017). Their longer onset latencies compared to the latency of S+-evoked excitations is consistent with this idea. However, the results do not support this hypothesis. If cue-
evoked inhibitions in some neurons were being driven by cue-evoked excitations in other MSNs, we would expect the proportion of cue-excited and cue-inhibited neurons to change in parallel, but the opposite pattern was observed.

A possible alternative is that cue-inhibited neurons are not MSNs but interneurons (section “1.2.2.2.”). CINs are a likely candidate and, functionally, a critical one, since their activity orchestrates the excitability and plasticity of local striatal circuits (Cachope et al., 2012; Centonze, Gubellini, Bernardi, & Calabresi, 1999). This kind of interneuron is a good candidate because it exhibits phasic inhibitions upon presentation of reward-predictive cues (Aosaki, Graybiel, & Kimura, 1994; Matsumoto, Minamimoto, Graybiel, & Kimura, 2001; Morris et al., 2004). However, CINs only account for about 1-2% of all striatal neurons. It is not impossible that these interneurons were overrepresented in the recordings, their baseline firing rate is high and their somas are large (Kawaguchi, 1993), and thus their activity likely causes large disruptions in the extracellular concentration of ions around them. This might make their fluctuations in voltage relatively easier to detect compared to those of MSNs. But it would be unlikely that about 18% of our recorded units came from this rare population of neurons. Also, throughout learning, cue-inhibitions increase among CINs (Aosaki et al., 1994) via NMDAR-mediated transmission (Oswald, Schulz, Kelsch, Oorschot, & Reynolds, 2015). This is in disagreement with our findings, since our population of cue-inhibited neurons declined, rather than increased, as training progressed. And NMDAR blockade prevented the decrease, not the increase, of cued inhibitions. Finally, pauses in CIN firing require intact dopaminergic transmission in the striatum (Aosaki et al., 1994), but cue-evoked inhibitions similar to the ones we observed are not affected by dopamine antagonists (du Hoffmann & Nicola, 2014). Thus, although training induces changes in the magnitude of cued inhibitions in CINs, it is unlikely that the population of cue-inhibited neurons we found was, at least in its entirety, composed of CINs.
A simple explanation would be that cue-evoked inhibitions in the NAc were driven by transient reductions in excitatory input. Although BLA neurons also exhibit cue-evoked inhibitions, the latency of these responses coincides with that of cue-evoked inhibitions in the NAc, arguing against a direct BLA-to-NAc projection driving these pauses in NAc activity (Ambroggi et al., 2008). Similarly, neurons in the dorsomedial PFC exhibit transient excitations and inhibitions upon presentation of reward-predictive cues (Otis et al., 2017). However, the PFC neurons that project directly to the NAc are mostly cue-excited neurons. Cue-inhibited neurons in the PFC, instead, project to the paraventricular nucleus of the thalamus (Otis et al., 2017), a source of glutamatergic input of NAc neurons (Berendse et al., 1992). Pharmacological inhibition of the dorsomedial PFC disrupts both excitations and inhibitions in the NAc (Ishikawa et al., 2008). This indicates that, even though a direct PFC-to-NAc connection is an unlikely driver of NAc cue-evoked inhibitions, a polysynaptic connection between both structures via the thalamus might be implicated.

Using in vivo Ca^{2+} imaging, Reed et al., (2018) examined task-related changes in the activity of NAc glutamatergic afferent structures. While they were able to capture clear changes—mostly reductions—in upstream activity at the time of reward consumption, they did not find clear cue-evoked inhibitions in these structures. However, these results need to be interpreted with caution. The authors also failed to find cue-evoked excitations in the activity of these NAc-projecting areas, even though these signals have been consistently identified by electrophysiological studies recording from these structures (Ambroggi et al., 2008; Critchley & Rolls, 1996; Jodo et al., 2000; Muramoto, Ono, Nishijo, & Fukuda, 1993; Otis et al., 2017; Schoenbaum et al., 1998, 1999; Tremblay & Schultz, 2000; Wirth et al., 2003). Because Ca^{2+} imaging recordings report overall changes in the activity of a population of neurons, it is possible that cue-evoked signals were undetected if cue-excited and cue-inhibited neurons coexisted within the same pathway. Further work is needed to understand the source of these cue-evoked inhibitions, which may offer some
clues about their behavioral relevance and the mechanisms implicated in their decline during learning.

3.2.2.4. NAc activity during reward consumption: from excitation to inhibition

NAc neurons in trained animals undergo transient excitations and inhibitions, but mostly inhibitions, during reward consumption (Nicola et al., 2004b). Although these inhibitions have been suggested to encode sensory properties of the reward (Carlezon & Thomas, 2009), evidence reveals that they are tightly linked to consummatory motor output. NAc inhibitions precede the initiation of licking for sucrose (Krause et al., 2010; Taha & Fields, 2006), even when sucrose is withheld (Nicola et al., 2004b), and they span the duration of the consummatory phase (Nicola et al., 2004b). Moreover, inactivation of NAc neurons promotes reward consumption (Kelley & Swanson, 1997; Stratford & Kelley, 1997), whereas brief activation of NAc neurons arrests ongoing consummatory behavior (Krause et al., 2010; Millan, Kim, & Janak, 2017). The increase in feeding observed upon NAc inactivation is thought to depend on the disinhibition of VP neurons (Shimura, Imaoka, & Yamamoto, 2006; Smith, Tindell, Aldridge, & Berridge, 2009; Stratford, Kelley, & Simansky, 1999), probably because releasing the VP-imposed inhibitory tone on VTA increases levels of dopamine in the NAc, promoting motivation for food (Soares-Cunha et al., 2018). This could suggest that only VP-projecting neurons express these consumption-related inhibitions.

Our results indicate, however, that NAc firing patterns during the consummatory phase change across training (Figure 8a), adding some complexity to this matter. Before the change point, NAc neurons did not exhibit inhibitions around the time of reward delivery. Instead, some neurons became excited upon S+-entry. This pattern of NAc activity sharply changed around the time cued approach behavior began to emerge. On the session during which the change point occurred and
afterwards, NAc neurons exhibited pronounced inhibitions upon rewarded receptacle entry. Notably, the population of cue-responsive neurons largely overlapped with the population of neurons that exhibited clear responsivity during reward delivery, both before and after the change point (Figure 8a). Interestingly, however, the direction of this relationship flipped during the change point (Figure 8b). Before change point, neurons that were excited by the S+ also became excited after the first subsequent receptacle entry. On the session during which the change point occurred and afterwards, cue-excited units were the same ones that also underwent pronounced inhibitions around the time of reward consumption. This overlap between S+-onset-excited and S+-entry-inhibited neurons was also reported in Nicola et al. (2004a). The authors suggested that this biphasic pattern of NAc activity could contribute to switching between two behavioral strategies at two points of the trial: after cue onset, cue-evoked excitations may promote specific reward-seeking responses and inhibit consummatory behaviors and, upon receptacle entry, a reduction in NAc activity may disinhibit consummatory behaviors while inhibiting competing appetitive responses.

The near absence of inhibitions around the time of reward delivery in naïve rats contrasted with previous literature, which reported that the majority (~75%) of NAc neurons in naïve rats were inhibited upon oral intraoral infusion of sucrose (Roitman et al., 2005). However, both findings support the hypothesis that NAc inhibitions around the time of reward delivery promote efficient consummatory behavior. After the CP, consummatory responses were practically stereotyped: rats were fast to enter the receptacle after S+ onset and their first entry lasted several seconds. Although the operant chambers used in this study did not detect the exact time when the reward was consumed, it is easy to imagine that the sucrose was likely consumed during that first prolonged receptacle entry. Nonetheless, this optimal pattern of consummatory behavior is only made possible by the certainty that fluid will be delivered upon receptacle entry, something that animals need to learn. In Roitman et al. (2005), because sucrose was intraorally infused, rats were forced to engage
in consummatory behavior upon reward delivery. But in this paradigm, that was something that naïve rats had to learn. In fact, before the CP, receptacle entries during the S+ tended to be short and erratic. It was common for rats to exit the receptacle immediately after triggering sucrose delivery just to later engage in a string of shorter randomly-timed ITI entries (during which the reward was presumably consumed). The fact that entry-related inhibitions and efficient reward collection emerged in parallel agrees with the role of these signals as facilitators of consummatory behaviors.

3.2.3. NAc NMDARs (transiently) contribute to the expression of cued approach behavior

If the strengthening of corticostriatal synapses in the NAc is responsible for the emergence of NAc cue-evoked activations and these signals trigger cued approach responses, disruption of NAc activations should impair the expression of the learned behavior (“retrograde alteration” criterion, Martin et al., 2000; section "1.4.2.2."). Models of synaptic plasticity indicate that the expression of LTP depends on the postsynaptic increase in AMPAR-mediated transmission (Herring & Nicoll, 2016; Malinow & Malenka, 2002; section "1.3.2.2."). Although the effects of AMPAR blockade were not directly tested in this project, disruption of excitatory transmission in moderately trained animals had profound neural and behavioral effects.

In Experiment 2, intra-accumbens NMDAR blockade strongly diminished the magnitude of already acquired cue-evoked excitations in the NAc of rats that had only been trained daily for about a week (Figure 10d). Although NMDAR blockade disrupted this aspect of excitatory transmission in the NAc, baseline firing was unaffected, indicating that the effects were not explained by a an overall attenuation of NAc activity (Supplementary Figure 6g,h). Critically, after intra-accumbens AP5 infusions, rats were less likely to approach the reward receptacle during the S+ (Supplementary Figure 6a) and, when they did, it took them longer to do so (Supplementary
Figure 6b). This reduction in NAc activations also increased the ITI pseudolatency (Supplementary Figure 6c), a measure of spontaneous responding. This last observation is in conflict with reports showing that disrupting excitatory transmission in the NAc leads to behavioral disinhibition (Ambroggi et al., 2011). However, an increase in task-irrelevant behavior has been more linked to the inactivation of the NAc shell than the core (Blaiss & Janak, 2009; Floresco, McLaughlin, & Haluk, 2008). On the other hand, these results are actually consistent with previous studies indicating that NMDAR blockade in the NAc interferes with active exploratory behavior (Maldonado-Irizarry & Kelley, 1995; Maldonado-Irizarry & Kelley, 1994), and entries during the ITI could reasonably be conceptualized as exploratory responses. Alternatively, it is possible that ITI entries are not uncued and spontaneous, but driven by weak cues (e.g., the sight, smell or touch of the reward receptacle). These covert cues could possibly trigger NMDAR-dependent cue-evoked excitations in the NAc. However, it would be challenging to capture them because the precise time at which these implicit cues enter the rat’s consciousness is unknown to the observer.

Importantly, however, the reduction in S+-elicited entries caused by NMDAR blockade cannot be explained by a non-specific reduction in overall entry responses because the S+ performance index (a metric of approach behavior specifically driven by the S+) was also diminished by AP5 (Figure 10c). Experiment 2 also reveals that the mechanism by which NAc NMDARs facilitate cued approach at this stage of training is by promoting large cue-evoked excitations in NAc neurons. The results of this experiment are in accordance with the “retrograde alteration” criterion (Martin et al., 2000). Of course, if we considered this finding in isolation, it would still be possible that growing NAc cue-evoked excitations were the result of plasticity upstream of the NAc. In that case, disrupting them at the NAc would just roadblock these activations traveling from corticolimbic areas to motor output structures. But because other results suggest that these signals emerged at the level of the NAc during appetitive conditioning
(Experiment 4; section “3.2.4.1”), the behavioral impairment caused by their disruption provides evidence to satisfy the “retrograde alteration” criterion.

However, interestingly, after enough training (18-19 sessions), the expression of both the approach behavior and the underlying NAc cue-evoked activity became independent of NMDAR activation (Figure 10e,f; Figure 15, Stage 4). A simple explanation of this shift towards NMDAR-independent behavioral and neuronal activity would be an increase in AMPAR expression, a process that mediates the expression of LTP (Herring & Nicoll, 2016). An increase in AMPAR-mediated transmission would likely make the relative contribution of NMDARs to NAc excitations less relevant. Alternatively, it is possible that overtraining induced an upregulation of NMDAR expression in NAc synapses. In that case, although the dose of AP5 used in this experiment was high, it may not have been high enough to have an effect on cue-evoked NAc activity and approach behavior. However, several observations contradict these two hypothesis.

First, if extensively training animals exhibited a postsynaptic increase of glutamatergic receptors (either AMPARs or NMDARs) relative to moderately trained animals, we would have expected the magnitude of NAc activations to be larger in the extended vs. the moderate training group, but we actually found the opposite pattern (Figure 10d vs. 10f, preinfusion period). This reduction in the magnitude of cue-evoked excitations must have been mediated by an additional plasticity mechanism activated by extended training. For example, the concentration of dopamine release in the NAc upon presentation of a reward-associated cue decreases after asymptotic performance is achieved (Clark, Collins, Sanford, & Phillips, 2013). Because dopaminergic transmission regulates the excitability (du Hoffmann & Nicola, 2014) and plasticity (Horvitz, 2002; Reynolds & Wickens, 2002) of corticostriatal synapses (section “1.2.3.3.”), it is possible that a decline in the concentration of NAc dopamine at the time of cue presentation attenuated the strength of NAc excitatory response. Alternatively, it is possible that the reduction in NAc activations was
caused by a reduction in cue-evoked input from glutamatergic afferents, although the evidence to support that possibility is sparse.

It is possible that even if NAc activations eventually become smaller and independent of NMDARs, their expression during extended training relies on other kinds of receptors within the NAc. An emergent body of literature suggests otherwise, indicating that, with overtraining, behaviors that once depended on NAc participation can eventually emancipate from NAc function altogether. Infusions of D1 receptor antagonists into the NAc core impair the expression of Pavlovian cued approach after 4-5 but not after 10-11 sessions of training (Dobrovitsky et al., 2019). This did not merely reflect a transient involvement of mesolimbic dopaminergic transmission in the expression of the response because blockade of ionotropic glutamate receptors had a similar effect (Dobrovitsky et al., 2019), suggesting the recruitment of a non-NAc core substrate. The dorsal striatum is also thought to gradually disengage from the expression of goal-directed actions (Barnes, Kubota, Hu, Jin, & Graybiel, 2005; Carelli, Wolske, & West, 1997; Smith & Graybiel, 2013; Tang, Pawlak, Prokopenko, & West, 2007). This has motivated the proposal that, with extended training, cortico-cortical pathways are formed and, consequently, sensory cortices can access motor cortical outputs directly, in the absence of basal ganglia mediation (Ashby, Turner, & Horvitz, 2010).

On the other hand, other studies have found that, after many training sessions, NAc cue-evoked excitations and/or goal-directed responses are still sensitive to D1 and D2 antagonists (du Hoffmann & Nicola, 2014), and combined AMPA/NMDAR antagonists infused into the NAc core (Ambroggi et al., 2011). The potential effects of D1 or AMPAR antagonists in the group of extensively trained rats in Experiment 2 remains an open question. It is possible that the amount of training required for the expression of appetitive responses to become NAc independent varies across task protocols and across studies. It is not clear that the rats in the extended training group in
Experiment 2 were in the same stage of learning as those whose behavior became NAc-independent in Dobrovitsky et al. (2019). The rats’ level of performance is not a good indicator, as suggested by the fact that rats in Experiment 2 displayed comparable performance levels after moderate vs. after extended training (Figure 10c,e, preinfusion period). Small variations in the task paradigm, however, such as the introduction of an S- or the temporal predictability of the cue, may influence the rate of acquisition and the critical period of NAc involvement. For example, in Clark et al. (2013), an overtrained reward-predictive cue recovered the ability to elicit a strong dopaminergic signal in the NAc when the temporal structure of the task changed, making cue onset more unpredictable. Finally, because data from the moderate vs. extended conditions comes from different populations of animals, it is possible that the difference in NAc activity across groups was due to a sampling bias. Examining changes in NAc activity as moderately trained rats become overtrained is another promising avenue of future research.

3.2.4. What these results say about the role of NAc plasticity on appetitive learning

3.2.4.1. Plasticity in the NAc is necessary for learning cued approach: what is the evidence?

Daily bilateral NMDAR antagonist injection in the NAc profoundly impaired cued approach learning (Supplementary Figure 5), suggesting that NMDAR-dependent plasticity within the NAc is essential for task acquisition. Previous pharmaco-behavioral studies drew similar conclusions (Di Ciano & Everitt, 2001; Hernandez et al., 2005; Kelley et al., 1997; Smith-Roe et al., 1999) but faced the confound that NMDAR antagonists could have impaired behavioral performance rather than plasticity. This confound has historically been addressed by showing that performance was unaffected by bilateral intra-accumbens NMDAR antagonist microinjections in well-trained animals. However, Experiment 2 found that single bilateral NMDAR antagonist injections impair
task performance and reduce the magnitude of cue-evoked excitations after a week of daily training (Figure 10c,d), suggesting that blockade of NMDARs on days 1-6 does indeed impair learning at least in part by impairing performance. To test whether impaired performance is sufficient to account for learning deficits on our task, **Experiment 3** imposed a limit on performance of vehicle-injected animals (“yoked animals”) such that the number of cue-reward pairings they experienced was similar to that of daily bilateral NMDAR antagonist-injected animals. Antagonist-injected animals exhibited severe learning deficits compared with their yoked controls, despite having experienced the same cue-reward contingency during training (Figure 11).

We cannot rule out the possibility that other performance-related factors, such as the timing of rewards within the session or the interval between cue onset and reward, contributed to impaired learning. In addition, although both groups experienced, on average, the same number of “unpaired cue” trials, the control group experienced many of these trials in the presence of an additional discrete stimulus (i.e., the gate preventing access to the receptacle). The gate may have acted as a conditioned inhibitor for the control group, helping maintain the cue-reward association intact during those trials (Rescorla, 2003). Although the most parsimonious interpretation of the results of is that daily NMDAR antagonist treatment induced both a performance deficit and an additional learning deficit, these alternative possibilities will have to be addressed in future experiments.

How does this finding fit into the framework laid out by Martin et al. (2000)? **Experiment 3** provides evidence to satisfy the “anterograde alteration criterion” by showing that a treatment that is known to block LTP in the NAc (i.e., NMDAR blockade, section “1.4.1.”) disrupts learning, even when controlling for performance effects. **Experiment 1** found experience-dependent increases in NAc cue and reward-consumption encoding associated with the emergence of the learned response (“detectability criterion”). **Experiment 2** revealed that, once acquired, disruption of NAc cue-evoked excitations attenuated the expression of recently acquired cued approach behavior (“retrograde
alteration criterion”). Hence, we hypothesized that the physiological mechanism by which NMDAR blockade anterogradely disrupted learning in Experiment 3 was by obstructing the emergence of NAc cue-evoked excitations during training. To address this question, in Experiment 4, animals received daily unilateral NAc NMDAR antagonist injections into the same hemisphere from which single unit activity was being monitored. This revealed the effects of the antagonist on NAc neuronal activity during learning in the absence of a behavioral deficit in either learning or performance. AP5 did not affect the emergence of training-induced changes in NAc activity during reward consumption (Supplementary Figure 8) but it did in the time window after cue onset (Fig. 13). Critically, the impact of AP5 on S+-evoked excitations depended on the stage of training relative to the CP (Figure 15).

As early in training as after 10 trials (much before the CP), some neurons in the NAc core exhibited brief, small S+-evoked excitations (Supplementary Figure 2b). This observation is in agreement with previous reports (Atallah et al., 2014; Roitman et al., 2005; Setlow et al., 2003). Remarkably, the emergence of these early, cue-encoding excitations was unaffected by NMDAR blockade in the NAc (Figure 13), suggesting that the plasticity necessary for these excitations takes place upstream of the NAc (Figure 15, Stage 1).

Areas upstream of the NAc also respond to reward-predictive cues (Ambroggi et al., 2008; Critchley & Rolls, 1996; Jodo et al., 2000; Muramoto et al., 1993; Otis et al., 2017; Schoenbaum et al., 1998, 1999; Tremblay & Schultz, 2000; Wirth et al., 2003). Neurons in the BLA, in particular, encode the motivational value of stimuli after just a few trials (Belova, Paton, & Salzman, 2008; Paton et al., 2006; Schoenbaum et al., 1999) via rapid NMDAR-mediated plasticity (Tye et al., 2008). Because these areas send excitatory projections to the NAc, NAc neurons at this early stage of training likely passively relay cue-evoked signals from the upstream structures from where they originate to downstream motor output areas. A finding that seemingly contradicts this hypothesis is
that these early excitations were absent in bilaterally or unilaterally AP5-treated animals that failed to acquire cued approach behavior (Supplementary Figure 5, 9). If the origin of these signals lies upstream of the NAc, why would a manipulation at the level of the NAc interfere with their emergence? However, because intra-accumbens NMDAR blockade in naïve rats interferes with exploratory (Maldonado-Irizarry & Kelley, 1995) and cued behavior (Experiment 2), the emergence of early cue-evoked activations in intact areas upstream of the NAc (and anywhere in the brain) could have been disrupted by a performance deficit.

S+-evoked excitations grew robustly in prevalence, magnitude, and duration beginning with the trials just prior to CP and continuing to grow until the end of training (Figure 15, Stages 2 and 3). In contrast to the tenuous, brief, and NMDAR-independent cue-evoked excitations observed during early trials, these later cue-evoked excitations were strongly attenuated in magnitude, duration and prevalence by daily unilateral NMDAR antagonist injection in the NAc (Figure 13c-f, trials after -35). Because unilateral AP5 produced no detectable behavioral deficit, it is unlikely that the AP5-induced reductions in cue-evoked neuronal responses were due to reductions in performance.

The observation that S+-evoked excitations in the AP5-injected NAc were smaller than those observed in the vehicle-injected side is consistent with the acute reduction in S+-evoked excitation observed when a single injection was made after seven training sessions (Figure 10). This suggests that perhaps the antagonist limited the magnitude of ongoing NAc cue-evoked activity. Importantly, however, in a drug-free extinction test conducted after training, neurons in the previously antagonist-injected NAc had much smaller S+-evoked excitations than neurons in the previously vehicle-injected NAc (Figure 14). Therefore, daily NMDAR antagonist injections do not merely produce acute reductions in the magnitude of cue-evoked excitation but prevent the emergence of S+-evoked excitation from just prior to the CP onward, even when the excitation is
monitored during a drug-free test session. Intra-NAc NMDA receptor blockade prevented the plasticity necessary for long-term changes in neuronal responsiveness to the S+ at the level of the NAc. Because these excitations are essential for cued approach behavior, it is reasonable to conclude that learning to approach reward in response to reward-predictive cues is the result of NMDA receptor-dependent plasticity within the NAc.

3.2.4.2. Main caveats

The conclusion that NAc plasticity promotes cued approach learning via the facilitation of cue-evoked increases in NAc neuronal firing rate should be taken with the following limitations in mind.

a) **Plasticity or increased excitatory input?** After LTP induction, the postsynaptic response to a presynaptic input of fixed strength increases. Hence, a definitive proof of LTP induction would require evidence that the input-output relationship has been strengthened after the presumed plasticity-inducing experience. The present study identified experience-dependent increases in NAc postsynaptic activity, but changes in the strength of the synapse were not directly monitored. This is, this study did not track the activity of any of the structures that project to the NAc while simultaneously recording from the NAc, as doing so *in vivo* in behaving animals poses a huge methodological challenge.

Throughout training, the cue parameters and the cue-reward relationship remained constant. Thus, the observed increase in cue encoding in NAc neurons must necessarily reflect that plasticity is taking place at least somewhere in the circuit in which the NAc is embedded. Nonetheless, technically, the approach used here would not be able to differentiate between the strengthening taking place at NAc synapses and an increase in afferent excitatory drive.
The claim that plasticity in the NAc facilitated the emergence of cued approach learning is, however, a strong inference based on the following observations:

- *In vitro* studies, which afford great control over synaptic input-output relationships, strongly identify NMDAR activation as the primary mechanism for LTP induction in the NAc (*section “1.4.1.”*). If training induced plasticity in the NAc, it would be improbable that it would be spared by NMDAR antagonist injections. *In vivo* studies have also found increased AMPAR/NMDAR ratio in BLA-to-NAc (Namburi et al., 2015) and hippocampus-to-NAc (LeGates et al., 2018) synapses after acquisition of appetitive behaviors.

- NMDAR activation within the NAc is causal to the emergence of the observed increases in cue-evoked excitations in the NAc (*Experiment 4*). The disruptive effects of AP5 on learning-related cue-evoked excitations outlasted the presence of the drug in the organism (Figure 14), arguing against a mere reduction in routine presynaptic activity.

- NAc NMDAR blockade during training prevented the acquisition of cued approach behavior, even when accounting for its role in performance (*Experiment 3*).

In summary, it is unlikely that the effects of AP5 on the emergence of NAc activations and cued approach behavior can be solely explained by a reduction of increasing excitatory drive coming from afferent structures. The most parsimonious explanation is that intra-accumbens NMDAR blockade disrupted training induced signals in the NAc (and the acquisition of the behavior) by preventing plasticity at the level of the NAc.

b) **Downstream effects can become upstream effects in basal ganglia loops.** It is possible that the antagonist-induced disruption of excitatory transmission in the NAc interfered with plasticity in downstream structures. Because the NAc is embedded in a recurrent
corticostriatothalamocortical loop, intraaccumbens AP5 injections that diminish cue-evoked excitations could have reduced the excitatory drive onto neurons that are both downstream and upstream of the NAc (e.g., in the thalamus or prefrontal cortex). Consequently, NAc AP5 injections could have impaired plasticity in these structures, but not within the NAc itself, resulting in diminished cue-evoked excitations even when AP5 is no longer present (Figure 14). Indeed, plasticity within other corticostriatothalamic structures may well have contributed to learning, as shown by the ability of NAc neurons to differentiate the S+ from the S- very early in training in a NAc NMDAR-independent way (Figure 13c-f). However, given the well-established role of NAc NMDARs in LTP in the NAc (section “1.4.1.”), it is unlikely that intra-NAc AP5 injections, which reduced the magnitude of NAc neurons’ cue-evoked excitations, impaired plasticity exclusively elsewhere and not in the NAc. Therefore, the most likely explanation for our results is that AP5 disrupted learning by preventing synaptic potentiation at the level of the NAc, which does not negate the contribution of plasticity in other regions to reinforcement learning (Kelley, Andrzejewski, Baldwin, Hernandez, & Pratt, 2003).

c) This interpretation hinges on a juxtaposition of neural and behavioral effects. Experiment 3 demonstrates the connection between NMDARs in the NAc and learning cued approach. Experiment 4 establishes a connection between NMDARs in the NAc and the acquisition of NAc cue-evoked excitations, a learning-associated neural signal (Experiment 1). However, the present work did not establish, within the same animals, the following connection: NMDAR blockade → effects on neural signaling in the NAc during training → effects on the acquisition of cued approach behavior. Because bilateral infusions of AP5 in the NAc induce a performance deficit early in training (Experiment 2), this was not a possibility.
Nonetheless, cue-evoked NAc activations have been well-established as the neural trigger of vigorous reward-seeking approach behaviors (section “1.2.4.”). It is hard to conceive that a manipulation that disrupts the emergence of cue-evoked excitations impaired the acquisition of cued approach behavior without these two events being connected.

The choice of the NMDAR antagonist in this project was motivated by the desire to integrate these findings within those of other groups in the field, most of which have used AP5 to probe the link between NAc plasticity and appetitive conditioning (section “1.3.3.2.”). Validating these results by implementing manipulations that disrupt plasticity while sparing normal transmission remains a goal for future research. This kind of approach would allow to simultaneously test the cellular, neural and behavioral effects of disrupting NAc plasticity. One possibility would be using the selective antagonist ifenprodil, which, unlike AP5, does not block the entire NMDAR complex. Ifenprodil targets the NR2B subunit of NMDARs, which is most strongly associated with plasticity and less implicated in normal channel function (Williams, 2001). This selective antagonist has already been used to identify the effects of disrupting lateral amygdala plasticity on the acquisition of fear conditioning (Rodrigues, Schafe, & LeDoux, 2001). An alternative approach would be to disrupt plasticity in the NAc by interrupting, after each training session, some step of the plasticity-inducing intracellular cascade initiated by NMDAR activation. A suitable choice would be anisomycin, a protein synthesis inhibitor that prevents the acquisition of reward-reinforced behavior when infused into the NAc core after each training session (Hernandez et al., 2002).

3.2.4.3. What kind of associative learning is mediated by NAc plasticity?
Because the delivery of reward was contingent on receptacle entry, the associative structure of the task was instrumental. This does not mean that what animals learned was an instrumental contingency. Even purely Pavlovian protocols impose a similar operant requirement: unless the reward is intraorally infused, animals have to enter the receptacle soon after the cue to experience the cue-reward association. It is a formal possibility then, that in this paradigm, instrumental learning was not involved at all in the acquisition of the cued approach response. However, unlike Pavlovian cues, the S+ in this task sets a limited temporal window of opportunity that incentivizes quick action: the faster animals respond, the sooner they obtain the reward. These temporal requirements are absent in Pavlovian paradigms and, possibly, they determine the extent to which the NAc is implicated in the acquisition and expression of the response.

NAc neurons acquire phasic responses to both Pavlovian (Day et al., 2006; Roitman et al., 2005) and instrumental (Nicola et al., 2004a; Setlow et al., 2003) cues. However, NAc neurons may be differentially implicated in controlling the expression of the learned response depending on the associative structure of the task. Although the present results cannot directly speak to this hypothesis, a recent study tested this possibility in the VP—the main efferent of NAc indirect pathway neurons—. Richard, Stout, Acs, & Janak (2018) trained animals using either an instrumental or Pavlovian version of the same cued approach task that was employed in this study. After training, neurons in the VP came to encode the reward-predictive value of the cue regardless of the associative structure of the task. However, interestingly, VP activity predicted the latency of the cued approach response in rats that received instrumental but not Pavlovian training, although in both cases the topography of the response was identical. Moreover, disruption of VP activity increased the latency of the instrumental but not the Pavlovian responses (Richard et al., 2018). It is likely, then, that despite the quasi-Pavlovian structure of the task used in this project, the instrumental temporal requirements imposed on the animals’ behavior crucially recruited NAc
In fact, tenuous NMDAR-independent cue-evoked excitations appear very early in training in the NAc (Supplementary Figure 2a,b; Figure 13c-f). These early signals encode the cue-reward association (i.e., they discriminate S+ from S-), but they do not represent the subsequent behavior. This form of cue-reward signal matches the kind of activity exhibited by BLA neurons (Belova et al., 2008; Paton et al., 2006), which is consistent with the hypothesis that it emerges upstream of the NAc (Figure 15, Stage 1). In well-trained animals, however, cue-evoked excitations in the NAc simultaneously encode the reward-predictive value of the cue and the vigor of the imminent response (McGinty et al., 2013; Morrison et al., 2017). Therefore, at some point during training, NAc neurons begin to connect cue-reward information with locomotor output. Our results suggest that this limbic-motor integration does not occur elsewhere in the circuit but, rather, at the level of the NAc (Figure 15, Stage 2). Around the time when cue-evoked approach responses clearly emerged, NAc S+-evoked excitations began to increase (Figure 7) via NMDAR-mediated transmission (Figure 13). Hence, although plasticity within NAc neurons may not be necessary for endowing stimuli with motivational value, it is necessary for learning to translate that information into overt reward-seeking behavior.

It is important to note that, although NAc neurons “learn” to jointly encode cue and response-related information, this by no means suggests that NAc activity exclusively represents instrumental associations. Pavlovian associations are known to promote operant behaviors (Lovibond, 1983; Rescorla & Solomon, 1967). This phenomenon is evident, for example, in Pavlovian-to-instrumental transfer (for a review, Cartoni, Balleine, & Baldassarre, 2016). Yin et al. (2008) suggest that the acquisition of Pavlovian stimulus-outcome associations during training depends on NAc plasticity mechanisms. They also sustain that Pavlovian associations formed in the NAc can invigorate the expression of instrumental actions (Corbit & Balleine, 2011), but that
learning about action-outcome associations depends on the dorsal striatum. The present results contradict such a clear-cut division of labor between striatal subregions.

Let us consider that the response acquired by animals in our study is controlled by an instrumental response-outcome association, and that presentation of the S+ activates a Pavlovian cue-reward association which in turn promotes the vigor of an operant response. According to Yin et al., (2008), plasticity within the NAc would be exclusively recruited at the time when the cue-reward association is formed. Subsequently, NAc activity would only contribute to boost the expression of the approach response, which would be driven by an operant contingency formed at the level of the dorsal striatum. However, the present results conflict with that interpretation in two critical aspects: a) NAc encoding of the cue-reward association was supplied by other structure in the circuit (i.e., NAc neurons acquired discrimination between the S+ and the S- even under conditions of NMDAR blockade) and, b) plasticity in the NAc was necessary for the acquisition of the cued response, which would not be the case if the response relied on an association formed at the level of the dorsal striatum. These two findings would contradict the model in Yin et al., (2008) even if the underlying representational structure driving the response was a cue-reward association. Our results suggest that, during training, NAc neurons acquire the ability to initiate approach responses in the presence of cues that possess incentive value, but we cannot speak to the existence or the nature of the underlying response-outcome contingency.

In summary, a combination of instrumental and Pavlovian processes likely drove the cued approach responses acquired by animals in this study. Manipulations of the cue-reward and approach-reward contingencies may have shed light on this matter, but characterizing the relative contribution of Pavlovian/instrumental processes was beyond the scope of this study. At this point, it was decided to remain agnostic about the underlying contingencies controlling the acquired approach behavior, but this also constitutes an interesting question to be explored in future research.
3.2.4.4. NAc plasticity mechanisms that may be implicated in cued approach learning

The precise nature of the NMDAR-dependent plasticity underlying the increase in cue-evoked excitations in the NAc also remains a topic for further research. One obvious candidate is NMDAR-dependent LTP, which has been observed in NAc neurons (section “1.4.1.”). Models of NMDAR-mediated neuroplasticity suggest that postsynaptic activation of these receptors causes a rush of Ca$^{2+}$ into the neuron and this event, in turn, sets in motion the intracellular machinery that leads to the strengthening of the synapse by promoting an increase in postsynaptic AMPAR-mediated transmission (Citri & Malenka, 2008; Kerchner & Nicoll, 2008). However, the approach employed in this project leaves room for additional possibilities.

For example, (Shindou et al., 2019) argue that NMDAR antagonists could exert an indirect effect on NAc plasticity by regulating presynaptic release of dopamine, a key mediator of striatal plasticity (Horvitz, Choi, Morvan, Eyny, & Balsam, 2007; Reynolds & Wickens, 2002). Activation of glutamatergic inputs into the NAc can increase the firing rate of VTA neurons (Floresco, Todd, & Grace, 2001), and striatal glutamate can locally regulate dopamine release from VTA terminals (Cheramy et al., 1991; Desce et al., 1992; Krebs et al., 1991). However, it is unlikely that NMDARs in VTA terminals are implicated in regulating striatal dopamine release in vivo (for a review, Cachope & Cheer, 2014). This is an interesting possibility that awaits to be examined in future research. Definitely establishing whether the locus of NMDAR-mediated plasticity in the NAc is postsynaptic involves manipulations that would be very hard to implement in behaving animals (e.g. loading postsynaptic cells with the NMDAR channel blocker MK-801). However, examining potential effects of NMDAR antagonists on dopamine release in this task is a feasible endeavor. It could be achieved by combining intracranial infusions with recordings of dopamine in the NAc.
using, for example, fast-scan cyclic voltammetry. A similar approach has been successfully used to test the effects of µ-opioids on phasic dopamine release onto the NAc core (Gómez-A et al., 2019).

Alternatively, the observation that AP5 reduced the magnitude of peak cue-evoked excitation even in the absence of plasticity (Figure 10d) is consistent with a role for NMDARs in basal synaptic transmission (Daw, Stein, & Fox, 1993). Because strong excitation is a prerequisite for associative plasticity (Hebb, 1949; Wigström & Gustafsson, 1986), NMDARs could have contributed to plasticity simply by facilitating strong postsynaptic excitation upon S+ presentation. If this was the case, strong photostimulation of NAc neurons should induce plasticity by merely increasing depolarization via the opening of channelrhodopsin cation channels and the activation of voltage-gated Ca\(^{2+}\) channels (this is, in an NMDAR-independent fashion). Although it has not been tested in NAc core neurons, induction of LTP using high-frequency photostimulation of ChR2-expressing pyramidal or dorsal striatal cells is disrupted by NMDAR blockade (Ma et al., 2018; Zhang & Oertner, 2007). This suggests that the contribution of NMDARs to the potentiation of NAc synapses is likely not limited to the mediation of strong postsynaptic depolarization.

Another clue to a possible plasticity mechanism is the finding that cue-evoked excitations in NAc core neurons were of remarkably long duration (Figure 7a; Supplementary Fig. 2d,e). Beginning prior to the CP, NAc neurons remained excited two seconds after S+ onset compared to the same time period after the S- onset (Supplementary Figure 2d). In the two-second window that preceded S+-elicited entries, NAc neurons were still significantly excited, even when the entry happened five or more seconds after the cue (Supplementary Figure 2e). The prolonged duration of these incipient cue-evoked excitations was considerably shortened by NMDAR blockade (Figure 13c, e-f). NMDARs could contribute to plasticity by facilitating this long excitation. Interestingly, mesolimbic dopamine encodes a value prediction error (Schultz, 1998) that is thought to contribute to striatal plasticity (Horvitz, 2002; Reynolds et al., 2001; Wickens et al., 2007) but only when
dopamine release onto NAc neurons occurs within ~2 s of glutamatergic input (Fisher et al., 2017; Shindou et al., 2019; Yagishita et al., 2014). Therefore, it is possible that, by keeping NAc neurons in a state of elevated activity, NMDARs allow the same neurons that were excited by the cue to maintain their excitation until the time of reward delivery, providing a neural eligibility trace for the reinforcing effects of dopamine release.

Several pieces of evidence are consistent with this eligibility trace hypothesis. First, the slow kinetics of NMDAR-mediated transmission (Dale & Roberts, 1985; Forsythe & Westbrook, 1988) facilitate robust temporal summation, which increases spike output upon repeated stimulation (Hunt & Castillo, 2012). Second, before the CP, neurons that were S+-excited also tended to exhibit excitations after receptacle entry (presumably during reward consumption; Figure 8). Early in training, unexpected reward delivery likely triggered phasic excitations in VTA neurons (Schultz, 1998), causing the release of both dopamine (Day et al., 2007) and glutamate (Stuber, Hnasko, Britt, Edwards, & Bonci, 2010; Tecuapetla et al., 2010; Wang et al., 2017) from dopamine terminals onto NAc neurons. Some glutamatergic afferents to the NAc respond with excitations to reward delivery (e.g., Muramoto et al., 1993; Paton et al., 2006). Hence, the convergence of reward-related signals from glutamatergic afferents and from VTA terminals —co-releasing dopamine and glutamate—may have selectively strengthened synapses onto striatal neurons that were still active at the time of reward delivery, resulting in the escalation of S+-evoked responses in those neurons. This mechanism, although speculative, would explain how this circuit solves the “credit assignment problem” – i.e., how two temporally disconnected events, whose neural representations always occur at a delay, can become associated within the same synapse.

Recent evidence may pose a challenge to this mechanism by showing that, in animals learning a cue-reward association for the first time, dopaminergic signals do not encode a reward prediction error (Coddington & Dudman, 2018). The authors argue that, in naïve animals, VTA
excitations around the time of reward consumption encode the self-initiated consummatory behavior, not the unexpected delivery of reward. Voltammetric recordings of dopamine release onto the NAc of naïve rats undergoing Pavlovian conditioning are consistent with this finding (Day et al., 2007). This raises several alternatives. Let us concentrate in the stage at which NAc plasticity is recruited during training, which is prior to the CP but after cue-reward associations forged in upstream regions start to be supplied to the NAc (Figure 15, Stage 2). At this point, S+ onset begins to elicit NMDAR-dependent prolonged depolarizations in NAc output neurons likely responsible for driving approach behavior. Because of these prolonged cue-evoked excitations, S+-responsive neurons in the NAc may become primed to respond upon further stimulation. Thus, reward-evoked glutamatergic input could further excite those NAc neurons and promote plasticity in the S+ encoding synapses via a Hebbian mechanism (i.e., it would facilitate that inputs carrying information about the S+ became good predictors of strong NAc depolarization). However, at that point in training, those inputs are already good predictors of NAc firing activity, it is improbable that a weaker reward-related excitatory input would have such a big potentiating effect on the already strong cue-encoding synapses.

Although dopamine may not encode a reward prediction error early in training (Coddington & Dudman, 2018), that does not preclude its involvement in the excitability and plasticity of S+-encoding synapses. Because of the prolonged excitations elicited by S+ onset in some NAc neurons, dopamine released during self-initiated consummatory behavior could still contribute to selectively potentiate those synapses. It could also selectively increase their excitability (Gerfen & Surmeier, 2011; Nicola et al., 2000), which is consistent with the fact that reward-related excitations were only observed in neurons that had already been excited by the cue. Also, animals in this project underwent a pretraining phase during which they received free rewards in the receptacle. Thus, it is also possible that, by the time training starts, events that precede reward delivery such as the sound
of the reward pump have become Pavlovian cues. In that case, phasic release of dopamine in the NAc could be occurring upon receptacle entry, even if it was not the elicited by the reward itself. Although suggestive, all these possibilities are speculative. More work is needed to elucidate the dynamics of NAc dopamine during training in this task and the glutamate-dopamine interplay in the facilitation of NAc cue-evoked activity during early stages of training.

The particular projection (or projections) whose synaptic contacts onto NAc core neurons are strengthened during training also remains unknown. Tetanic stimulation of the three main glutamatergic afferents to the NAc core (BLA, PFC, and hippocampus) leads to NMDAR-dependent potentiation of their excitatory synapses (Floresco, Blaha, et al., 2001a; Floresco, Blaha, Yang, & Phillips, 2001b; LeGates et al., 2018; Pennartz et al., 1993; Popescu et al., 2007). After appetitive learning, the AMPAR/NMDAR ratio—a proxy for glutamatergic synaptic strength—increases in hippocampal and amygdalar synapses onto NAc neurons (LeGates et al., 2018; Namburi et al., 2015). In trained animals, neurons in the all of these regions exhibit phasic excitations upon presentation of reward-predictive cues (Jodo et al., 2000; Miyashita et al., 1989; Muramoto et al., 1993; Paton et al., 2006), and intact input from the BLA and prefrontal cortex is required for cue-evoked excitations in NAc core neurons as well as for the performance of cued approach behavior (Ambroggi et al., 2008; Ishikawa et al., 2008). Further research to characterize the role of the NAc’s afferents in the development of cue-evoked excitations is necessary to better understand how these excitations emerge and drive cued approach learning.

3.2.5. Final considerations

The results of this project suggest that, as a result of NMDAR-mediated plasticity within the NAc core, neurons in this structure acquire responsivity to a reward-predictive cue during training.
These experience-dependent neuroadaptations in the NAc core are necessary for animals to learn approach responses to the predicted location of reward at the time the cue is presented. These results also suggest that accumbens NMDARs mediate the expression of cued approach behavior, although their involvement—and perhaps that of the NAc core—diminishes with overtraining.

These findings raise many additional intriguing questions. Some of these have been mentioned in previous sections, such as the identity of the cells that undergo the observed learning-related changes, the upstream structures whose NAc projections become potentiated with training, the nature of the underlying contingencies driving the incipient behavior, the plasticity mechanisms recruited by NMDAR activation, etc. However, to determine whether NAc LTP is causally linked to learning, one big question remains to be answered: is the long-lasting strengthening of excitatory NAc synapses sufficient for the acquisition of cued approach responses? Of course, this question poses a huge experimental challenge, but some recently developed techniques offer some hope. For example, it is now possible to express light-sensitive receptors in pathways that are particularly active in behaving animals (Guenthner, Miyamichi, Yang, Heller, & Luo, 2013). Hence, one could imagine an experiment in which, for example, during early training, BLA-to-NAc fibers encoding the cue-reward association come to express the excitatory opsin channelrhodopsin. Then, by applying low frequency photostimulation of this pathway, the synapses between those terminals and the NAc could be depotentiated (Kim & Cho, 2017; Pascoli et al., 2012). Based on these results, weakening those synapses should attenuate the previously acquired responses. If that happened, could optogenetically induced LTP in this pathway, in the absence of experience, reinstate cued approach behavior?
4. APPENDIX

4.1. Supplementary Figures
Supplementary Figure 1. Representative individual neurons at different points of training.

(a) Sample peri-event time raster plots (top) and histograms (bottom) aligned to the time of S+ onset. Each row of graphs shows three representative neurons of the same animal, one recorded on the day before change point (left), another one recorded on the change point session (middle) and the last one recorded on the sixth day of training (right). Dots in the raster plots represent action potentials fired by the recorded neuron and trials are sorted from earliest to latest from top to bottom. Histograms were converted to firing rate using 50 ms bins. The y-axis of histograms is capped at 15 Hz to facilitate comparison across neurons. “Day” numbers refer to the training day.

(b) Same as “(a)” but with neuronal data aligned to the time of receptacle entry during the S+.
Post-cue (100-400 ms) firing by session and breakdown of session 1

Duration of cue-evoked excitations

Post-cue (100-400 ms) firing by sensory modality of the cue

Post-cue (100-400 ms) firing by cue/response combination
Supplementary Figure 2. Additional graphs showing NAc core firing activity during training.

(a) Population firing rate (median and interquartile range) in the 100-400 ms window after S+ (light blue) or S- (dark blue) onset by session. Numbers indicate sample size. The gray line indicates the cumulative percentage of units recorded from animals that exhibited a behavioral change point on or before that session. Post-cue firing was higher in S+ than S- trials in most sessions (*p < 0.05; **p < 0.01; ***p < 0.001, Wilcoxon).

(b) Same as “(a)” but only for the first session in 10-trial bins. Firing rate was higher after S+ than S- onset after only 10 trials (*p < 0.05; **p < 0.01; ***p < 0.001, Wilcoxon).

(c) Population firing rate (median and interquartile range) in the 100-400 ms window after S+ presentation when the S+ was a tone (blue) or a light (red) in the sessions before (left, “Pre CP”) or after change point (right, “Post CP”). There was no main effect of the sensory modality of the cue (Tone vs. Light; F1,145 = 0.006, p = 0.9403).

(d) Same as Figure 7c but for the 750-2000 ms post-cue window. Starting just before behavioral change point, firing rate after S+ onset was higher than after S- onset in this window (**p < 0.01; ***p < 0.001, Wilcoxon).

(e) To test whether the firing rate of NAc neurons was elevated prior to receptacle entry in S+ trials even when the latency to enter was long, we calculated the firing rate during the pre-entry 2 s window in trials during which it took animals 5 s or more to make a receptacle entry. Starting before behavioral change point, pre-entry NAc firing rate was higher in S+ than S- trials even when the latency to enter the reward receptacle was over 5 s (*p < 0.05; **p < 0.01, Wilcoxon).

(f) Each line depicts the average firing rate of each recorded neuron in the post-cue 100-400 ms window after S+ and S- cues that subjects responded to (resp.) or missed. Units are divided into three blocks depending on whether the session in which they were recorded was before the behavioral change point (Before CP), the session during which the change point took place (“CP session”) or after the change point (After CP). Within each block, neurons are sorted from top to bottom in descending order according to the magnitude of their activity in the 100-400 ms post-S+ window. The legend on the right shows the correspondence between colors and firing rate values.
Supplementary Figure 3. Activity in the NAc around the time of receptacle entry.

(a) From left to right, four heat maps represent average neuronal activity around the time of S+ onset, S+ entry, S- entry and ITI entry. Across heat maps, each line represents the same neuron. Units are divided into three blocks depending on whether the session during which they were recorded took place before the behavioral change point (Before CP), on the session during which the change point took place (CP session) or after the change point (After CP). Within each block, neurons are sorted from top to bottom in descending order according to the magnitude of their activity in the 100-400 ms post-S+ window. The legend on the right shows the correspondence between colors and firing rate values (in Z scores).

(b-c) Black dots represent each neuron’s firing rate in the 100-400 ms window after S+ onset plotted against the same neuron’s firing rate in the 0-1500 ms window after S- (b) or ITI (c) entry before (top) and after (bottom) behavioral change point. The regression line is shown in gray and the outliers are depicted in red (a few outliers fall outside the graph axes). Outliers are excluded from the analyses that yielded the results shown in these graphs. Including those outliers did not substantially change the results (Supplementary Table 1). Firing rate after S- or ITI entry was not significantly correlated with S+-evoked firing rate before or after change point (p > 0.05).
Supplementary Figure 4. NAc cue-evoked excitations emerged during training regardless of whether the electrodes were driven down in between sessions or not.

(a) For animals whose arrays were not driven down after each session, comparison of average S+ performance index (**t = -6.84, p < 0.001), entry probability (**t = -3.9, p = 0.0059, latency (**t = 5.24, p = 0.0018) and ITI pseudolatency (t = -1.72, p = 0.0059) before change point (Pre CP) vs. after change point (Post CP).

(b) When electrode arrays are not driven down in between sessions, the resulting data set includes recordings of some neurons that are the same across days, and others that are not. This means that data collected across days contains a mixture of repeated and non-repeated measures. This precludes the comparison between sessions using statistical inference tests, since these tests require that observations across conditions are comprised of either repeated measures samples (within-subjects comparisons) or different samples (across-subjects comparisons). Driving the electrodes down in between sessions to sample a new population of neurons each day avoids this confound, but it also introduces a potential anatomical confound when comparing neuronal activity across sessions. In order to assess whether advancement of the probes had an effect on the learning-related increase in S+-evoked firing, we compared post-S+ firing in the group of subjects whose arrays were maintained in the same location during training with those subjects whose arrays were advanced in between sessions (Figure 7), both before and after the change point. The graph shows firing rate (median and interquartile range) in the 100-400 ms post-S+ window before change point (Pre CP) and after change point (Post CP) in cue-excited neurons of rats whose arrays were driven down (blue) or not (gray) after each session. S+ evoked activity before or after the change point is similar across groups (p > 0.05, Wilcoxon).

(c) Average activity per channel (in channels that captured firing rate from two or more units) on the day before (left) and the day after (right) behavioral change point during the 100-400 ms window after S+ (light blue) or S- (dark blue). Within-channel comparisons showed that activity evoked by the S+ was higher than activity evoked by the S- in both sessions. They also revealed that S+-evoked activity was higher on the day after behavioral change point compared to the day before behavioral change point (**p < 0.01; ***p < 0.001, Wilcoxon). These results suggest that the emergence of cue-evoked excitations observed in Figure 7 are not accounted for by the dorsoventral location of the recording electrodes.
Supplementary Figure 5. Bilateral blockade of NMDARs during training disrupts the emergence of cue-evoked excitations in NAc as well as the acquisition of cued approach behavior.

(a) Mean±SEM entry probability during the S+ (light blue), S- (dark blue) or pre-S+ ITI window (gray) in animals that received daily bilateral AP5 injections prior to training.

(b-c) Same as ‘(a)’ but for latency and ITI pseudolatency (b), and performance index (c).

(d) Firing rate (median and interquartile range) in the 100-400 ms window after presentation of S+ (light red) or S- (dark red) in 35-trial bins (each bin corresponds to a session) in animals that received daily bilateral AP5 injections. During the first session, activity elicited by the S- was higher than activity elicited by the S+ (**p < 0.001, Wilcoxon). Post-S+ firing was comparable to post-Sfiring in subsequent sessions (p > 0.05, Wilcoxon). Numbers indicate sample size.

(e) Proportion of significantly excited (solid bars) or inhibited (white bars) NAc units upon presentation of the S+ in subjects that received daily bilateral AP5 injections. The proportion of neurons significantly excited or inhibited by the cue was independent of the amount of training animals had received (excitations: p = 0.2718; inhibitions: p = 0.9478, Fisher).
Supplementary Figure 6. Additional graphs showing how behavior and NAc core activity were affected by bilateral AP5 microinjections after moderate or extended training.

(a) Probability of entry during the S+ (left) or S- (right) before (Pre) or after (Post) infusion of vehicle (blue, n = 6) or AP5 (red, n = 5) in moderately trained animals. In S+ trials, entry probability was significantly diminished by microinjection of AP5 (*t = -3.504, p = 0.0248) but not vehicle (t = -0.445, p = 0.6624).

(b-c) Same as “(a)” but for cued latency (b) or ITI pseudolatency (c). In S+ trials, microinjections of AP5 increased the latency to make an entry during both the S+ (*t = -3.085, p = 0.0367) and the ITI period (*t = -2.916, p = 0.0434), whereas vehicle injections did not have that effect (S+: t =0.709, p = 0.7450; ITI: t = -0.229, p = 0.3881).

(d-f) Same as “(a-c)” but for animals that received extended (n =5) instead of moderate training.
A two-factor ANOVA using drug and time as within-subject factors revealed no main or interactive effects in S+ or S- entry probability/latency and ITI pseudolatency (all effects: p > 0.05).

(g-h) Baseline firing rate before injection plotted against baseline firing rate after saline (g) or AP5 (h) injection. In both cases, the 99% confidence interval (CI) around the slope of the regression line (vehicle: 0.46-1.38; AP5: 0.59-1.06) did not significantly differ from the unity line (i.e. the confidence interval contained the value “1”), suggesting that baseline firing rate was not affected by either injection.

(i-j) Same as “g-h” but for animals that received extended training prior to the saline (CI: 0.92-1.16) or AP5 (CI: 0.84, 1.11) injection. The baseline firing rate in these animals was also unaffected by the injections.
Supplementary Figure 7. Baseline firing rate in NAc core was not affected by unilateral infusions of AP5.

(a) Raw firing rate (median and interquartile range) in the 2 s window before S+ onset in the saline (blue) or AP5-treated (red) side in 35-trial bins around the trial in which the behavioral change point took place. Numbers represent the number of neurons recorded on each bin on the vehicle (blue) or the AP5-treated hemisphere (red). There was no difference in baseline firing rate across hemispheres in any of the bins (p > 0.05, Wilcoxon; Holm-Sidak adjusted).

(b) During the extinction test, learners’ firing rate (median and interquartile range) in the 2 s window before S+ onset in the hemisphere that had been treated with saline (blue) or AP5-treated (red) during training. There was no difference in baseline firing rate across hemispheres during this session (p > 0.05, Wilcoxon).
Supplementary Figure 8. NAc activity around the time of S+, S- or ITI entry in animals treated with unilateral AP5 microinjections.

(a) Heat maps representing firing rate in 50 ms bins around the time of S+ entry (top), S- entry (middle) or ITI entry (bottom) in the vehicle (left) or AP5-treated side (right) of subjects that received unilateral AP5 microinjections during training. Each line on each heat map represents a neuron. Neurons are divided into two blocks depending on whether the animal learned the task during training (learner) or not (non-learner). In the learners block, neurons are further divided into three blocks: units recorded before the change point (Before CP), during the session in which the CP took place (CP) or after the CP (After CP). Within each one of these blocks, units are sorted from top to bottom in descending order based on their average firing rate in the 0-500 ms window after the event the data is aligned to (i.e., S+ entry, S- entry or ITI entry respectively). The magnitude of the firing rate on each bin is color-coded according to the legend in the right.

(b) Firing rate during the pre-entry 2 s window in the vehicle (blue) or AP5-treated side (red) in S+ trials during which it took animals 5 s or more to make a receptacle entry. Starting before behavioral change point, pre-S+-entry firing rate was higher in the vehicle than in the AP5-treated side even when the latency to enter the reward receptacle was long (*p < 0.05; **p < 0.001, Wilcoxon).

(c) The proportion of excited (solid) or inhibited (empty) units upon S+ entry before (left) or after (right) CP across hemispheres (vehicle: blue; AP5: red) was comparable (p > 0.05, Fisher). The magnitude of the post-S+-entry response of these units (insets: median and interquartile range) was also similar (p > 0.05, Wilcoxon).
Supplementary Figure 9. Cue-evoked excitations did not emerge in the NAc core neurons of animals that failed to learn the task under daily unilateral AP5 injections (“non-learners”).

(a) Individual cumulative performance index records on S+ (left) and S- (right) trials in animals that received unilateral AP5 injections and did not learn the task. Each line represents a different animal. A positive change point was not identified in their S+ performance.

(b-d) Mean±SEM performance index (b), latency (c) and entry probability (d) of non-learners in 5-trial bins throughout training. S+ trials are represented in light blue, S- trials in dark blue and, in gray, the 10 s ITI window that preceded the S+.

(e) For animals that failed to learn the task, population firing rate in NAc neurons in the vehicle (left) or AP5 (right) side in S+ trials (light blue/red) and S- trials (dark blue/red) in the 100-400 ms window after the cue. S+-evoked excitations did not emerge throughout training in any of the sides (p > 0.05, Wilcoxon; Holm-Sidak adjusted).

(f) The proportion of significantly S+-excited (top) or inhibited (bottom) units in the vehicle (blue) and AP5-treated (red) side of non-learners. Throughout training, the percentage of neurons whose activity was significantly modulated by the cue did not differ across hemispheres (p > 0.05, Fisher; Holm-Sidak adjusted). Only in the last session, there was a significant increase in the percentage of cue-excited units (*p = 0.0465, Fisher; Holm-Sidak adjusted).

(g) Performance index in S+ (light blue) and S- (dark blue) trials during the drug-free extinction test in the two non-learners that were given this test.

(h) Firing rate around the time of S+ onset in 50 ms bins in the vehicle (blue) and AP5 (red) sides during the drug-free extinction test in animals that failed to learn the task during training. The inset represents the percentage of units that were excited by the S+ during the drug-free extinction test in the hemispheres that, during training, received either vehicle (blue; n = 26) or AP5 (red; n =15) injections. There were no differences in the percentage of cue-excited units across hemispheres in these animals (p = 1, Fisher).
Electrophysiological recordings only

Recordings/microinjections (AP5 effects after moderate vs. extended training)

+1.56 mm
+1.44 mm
+1.32 mm
+1.20 mm

Microinjections only

Recordings/microinjections (Bilateral or unilateral AP5 injections during training)

+1.68 mm
+1.56 mm
+1.44 mm
+1.32 mm
+1.20 mm
+1.08 mm

Saline
AP5
No infusion
Saline or AP5
Supplementary Figure 10. Anatomical location of injection and recording sites.

For each experiment, diagrams of coronal sections of rat brain at different anteroposterior coordinates (Paxinos & Watson, 2007). In animals that received no infusions, empty blue circles mark the tips of the electrode arrays. Solid dots mark the sites where the injectors delivered saline (blue), AP5 (red) or either one depending on the session (purple).
4.2. Supplementary Table

The tables below contain information about the statistical tests performed on the data depicted in the main figures (4.2.1.) and supplementary figures (4.2.2.).

4.2.1. Statistical tests. Main figures.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Independent variable(s)</th>
<th>Dependent variable</th>
<th>Test</th>
<th>Result</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 6f</td>
<td>Before vs. after change point</td>
<td>S+ perf. index</td>
<td>Paired t-test (Holm-Sidak corrected)</td>
<td>t_{5} = -11.968, p = 0.0003</td>
<td>n = 6</td>
</tr>
<tr>
<td>Figure 6f</td>
<td>Before vs. after change point</td>
<td>S+ entry probability</td>
<td>Paired t-test (Holm-Sidak corrected)</td>
<td>t_{5} = -6.069, p = 0.0035</td>
<td>n = 6</td>
</tr>
<tr>
<td>Figure 6f</td>
<td>Before vs. after change point</td>
<td>S+ latency</td>
<td>Paired t-test (Holm-Sidak corrected)</td>
<td>t_{5} = 6.849, p = 0.0030</td>
<td>n = 6</td>
</tr>
<tr>
<td>Figure 6f</td>
<td>Before vs. after change point</td>
<td>ITI pseudolatency</td>
<td>Paired t-test (Holm-Sidak corrected)</td>
<td>t_{5} = 0.1855, p = 0.8601</td>
<td>n = 6</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (-120 to -81 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0053</td>
<td>n = 45</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (-80 to -41 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.4649</td>
<td>n = 55</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (-40 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0075</td>
<td>n = 68</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (0 to 39 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 63</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (40 to 79 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 37</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>S+ vs. S- (80 to 119 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0004</td>
<td>n = 17</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>-120 to -81 vs. -40 to -1 trials from CP</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.2957</td>
<td>n = 45 / n = 68</td>
</tr>
<tr>
<td>Figure 7c</td>
<td>-40 to -1 vs. 40 to 79 trials from CP</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0218</td>
<td>n = 68 / n = 37</td>
</tr>
<tr>
<td>Figure 7d</td>
<td>-120 to -81 vs. -40 to -1 trials from CP</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0321</td>
<td>n = 45 / n = 68</td>
</tr>
<tr>
<td>Figure 7d</td>
<td>-40 to -1 vs. 40 to 79 trials from CP</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0229</td>
<td>n = 68 / n = 37</td>
</tr>
<tr>
<td>Figure 7d</td>
<td>-120 to -81 vs. -40 to -1 trials from CP</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.3424</td>
<td>n = 45 / n = 68</td>
</tr>
<tr>
<td>Figure 7d</td>
<td>-40 to -1 vs. 40 to 79 trials from CP</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0153</td>
<td>n = 68 / n = 37</td>
</tr>
<tr>
<td>Figure 7e</td>
<td>-120 to -81 vs. -40 to -1 trials from CP</td>
<td>Cue-excited units Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0473</td>
<td>n = 17 / n = 39</td>
</tr>
<tr>
<td>Figure 7e</td>
<td>-40 to -1 vs. 40 to 79 trials from CP</td>
<td>Cue-excited units Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0473</td>
<td>n = 39 / n = 30</td>
</tr>
<tr>
<td>Figure 7f</td>
<td>Average S+ latency on each session</td>
<td>Average firing 100-400 ms after S+ (Z sc.) of all cue-excited units on each session</td>
<td>Simple linear regression</td>
<td>Without outliers: r = -0.8, p &lt; 0.0001; R² = 0.65; β=-0.85, p &lt; 0.0001</td>
<td>n = 24</td>
</tr>
<tr>
<td>Figure 7f</td>
<td>Average S+ entry probability on each session</td>
<td>Average firing 100-400 ms after S+ (Z sc.) of all cue-excited units on each session</td>
<td>Simple linear regression</td>
<td>Without outliers: r = 0.71, p &lt; 0.001; R² = 0.50; β=-5.4, p &lt; 0.001</td>
<td>n = 24</td>
</tr>
<tr>
<td>Figure 7f</td>
<td>Average S+ performance index on each session</td>
<td>Average firing 100-400 ms after S+ (Z sc.) of all cue-excited units on each session</td>
<td>Simple linear regression</td>
<td>Without outliers: r = 0.78, p &lt; 0.001; R² = 0.60; β=-0.96, p &lt; 0.001</td>
<td>n = 24</td>
</tr>
<tr>
<td>Figure 10b</td>
<td>Before change point Firing 100-400 ms after S+ onset (Z sc.)</td>
<td>Before change point Firing 0-1500 ms after S+ entry (Z sc.)</td>
<td>Simple linear regression</td>
<td>Without outliers: r = 0.5, p &lt; 0.0001; R² = 0.25; β=0.62, p &lt; 0.0001</td>
<td>n = 97</td>
</tr>
<tr>
<td>Figure 10b</td>
<td>On or after change point session Firing 100-400 ms after S+ onset (Z sc.)</td>
<td>On or after change point session Firing 0-1500 ms after S+ entry (Z sc.)</td>
<td>Simple linear regression</td>
<td>Without outliers: r = -0.38, p = 0.0003; R² = 0.15; β=-1.29, p = 0.0003</td>
<td>n = 89</td>
</tr>
<tr>
<td>Figure 10c</td>
<td>AP5 vs. VEH group 30 min bins Moderate training</td>
<td>S+ performance index</td>
<td>Mixed two-factor ANOVA</td>
<td>Drug: F(1, 9) = 12.119, p = 0.0069</td>
<td>n = 11</td>
</tr>
<tr>
<td>Figure 10c</td>
<td>AP5 vs. VEH group 1 to 30 min (baseline) Moderate training</td>
<td>S+ performance index</td>
<td>Welch’s t-test (Holm-Sidak corrected)</td>
<td>t_{(22.7)} = 0.304, p = 0.3819</td>
<td>n = 5 / n = 6</td>
</tr>
<tr>
<td>Figure 10c</td>
<td>AP5 vs. VEH group 31 to 60 min Moderate training</td>
<td>S+ performance index</td>
<td>Welch’s t-test (Holm-Sidak corrected)</td>
<td>t_{(25.93)} = 4.292, p = 0.0004</td>
<td>n = 5/ n = 6</td>
</tr>
<tr>
<td>Figure 10c</td>
<td>AP5 vs. VEH group 61 to 90 min Moderate training</td>
<td>S+ performance index</td>
<td>Welch’s t-test (Holm-Sidak corrected)</td>
<td>t_{(25.39)} = 4.021, p = 0.0007</td>
<td>n = 5/ n = 6</td>
</tr>
<tr>
<td>Figure 10c</td>
<td>AP5 vs. VEH group 91 to 120 min Moderate training</td>
<td>S+ performance index</td>
<td>Welch’s t-test (Holm-Sidak corrected)</td>
<td>t_{(27.76)} = 3.553, p = 0.0013</td>
<td>n = 5/ n = 6</td>
</tr>
<tr>
<td>Figure 10d</td>
<td>VEH group (S+): before vs. after infusion Moderate training</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.6406</td>
<td>n = 8</td>
</tr>
<tr>
<td>Figure 10d</td>
<td>AP5 group (S+): before vs. after infusion Moderate training</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.00001</td>
<td>n = 30</td>
</tr>
<tr>
<td>Figure 10e</td>
<td>AP5 vs. VEH infusion 30 min bins Extended training</td>
<td>S+ performance index</td>
<td>Two-factor rep. measures ANOVA. Within-subject: Drug</td>
<td>Drug: F(1, 4) = 2.251, p = 0.06729</td>
<td>n = 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within-subject: Time</td>
<td>Time: F(1, 4) = 0.207, p = 0.2079</td>
<td></td>
</tr>
<tr>
<td>Figure</td>
<td>Drug</td>
<td>Time bin</td>
<td>Drug x time: $F_{(1, 4)} = 0.211$, $p = 0.6701$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10e</td>
<td>AP5 vs. VEH infusion 1 to 30 min (baseline) Extended training</td>
<td>S+ performance index</td>
<td>Unpaired t-test (Holm-Sidak corrected) $t_{(4)} = 0.188$, $p = 1$ n = 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10e</td>
<td>AP5 vs. VEH infusion 31 to 60 min Extended training</td>
<td>S+ performance index</td>
<td>Unpaired t-test (Holm-Sidak corrected) $t_{(4)} = 1.744$, $p = 0.206$ n = 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10e</td>
<td>AP5 vs. VEH infusion 61 to 90 min Extended training</td>
<td>S+ performance index</td>
<td>Unpaired t-test (Holm-Sidak corrected) $t_{(4)} = -0.314$, $p = 1$ n = 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10e</td>
<td>AP5 vs. VEH infusion 91 to 120 min Extended training</td>
<td>S+ performance index</td>
<td>Unpaired t-test (Holm-Sidak corrected) $t_{(4)} = 0.139$, $p = 1$ n = 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10f</td>
<td>VEH infusion (S+): before vs. after infusion Extended training</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p = 1$ n = 47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10f</td>
<td>AP5 infusion (S+): before vs. after infusion Extended training</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p = 0.2041$ n = 59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11e</td>
<td>AP5 vs. VEH infusion</td>
<td>S+ entry probability</td>
<td>Mixed two-factor ANOVA: Drug: $F_{(1, 12)}=33.26$, $p &lt; 0.001$; Bin: $F_{(1, 12)}=38.4$, $p &lt; 0.001$; Drug x bin: $F_{(1, 12)}=22.35$, p &lt; .001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11e</td>
<td>AP5 vs. VEH during first 5 min</td>
<td>S+ entry probability</td>
<td>Welch’s t-test $t_{(11,55)}=9.72$, $p &lt; 0.001$ n = 7/ n = 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11f</td>
<td>AP5 vs. VEH</td>
<td>S+ performance index</td>
<td>Mixed two-factor ANOVA: Drug: $F_{(1, 12)}=24.11$, $p &lt; 0.001$; Bin: $F_{(1, 12)}=1.11$, $p = 0.31$; Drug x bin: $F_{(1, 12)}=10.25$, $p=0.007$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11f</td>
<td>AP5 vs. VEH during first 5 min</td>
<td>S+ performance index</td>
<td>Welch’s t-test $t_{(11,47)} = 5.41$, $p &lt; 0.001$ n = 7/ n = 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12d</td>
<td>Before vs. after change point</td>
<td>S+ performance index</td>
<td>Paired t-test (Holm-Sidak corrected) $t_{(10)} = -10.21$, $p &lt; 0.00001$ n = 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12d</td>
<td>Before vs. after change point</td>
<td>S+ Entry probability</td>
<td>Paired t-test (Holm-Sidak corrected) $t_{(10)} = -5.061.., p = 0.001$ n = 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12d</td>
<td>Before vs. after change point</td>
<td>S+ latency</td>
<td>Paired t-test (Holm-Sidak corrected) $t_{(10)} = 5.938$, $p = 0.0004$ n = 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12d</td>
<td>Before vs. after change point</td>
<td>ITI pseudolatency</td>
<td>Paired t-test (Holm-Sidak corrected) $t_{(10)} = 1.245$, $p = 0.241$ n = 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>VEH side: S+ vs. S- (-105 to -71 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p = 0.3360$ n = 75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>VEH side: S+ vs. S- (-70 to -36 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p &lt; 0.0001$ n = 137</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>VEH side: S+ vs. S- (-35 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p &lt; 0.0001$ n = 129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>VEH side: S+ vs. S- (0 to 34 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected) $p &lt; 0.0001$ n = 131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13b</td>
<td>VEH side: S+ vs. S-</td>
<td>Firing 100-400 ms after cue</td>
<td>Wilcoxon signed- $p &lt; 0.0001$ n = 92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure</td>
<td>Group 1</td>
<td>Group 2</td>
<td>Test</td>
<td>Corrected p-value</td>
<td>n (Group 1)</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>VEH side: S+ vs. S- (70 to 104 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 36</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>VEH side: bin 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p=0.0019</td>
<td>n = 75/n = 129</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>VEH side: bin 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p=0.0013</td>
<td>n = 129/n = 92</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (-105 to -71 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0004</td>
<td>n = 66</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (-70 to -36 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 103</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (-35 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 96</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (0 to 34 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 102</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (35 to 69 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 85</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: S+ vs. S- (70 to 104 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 75</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: bin 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p=0.0071</td>
<td>n = 66/n = 96</td>
</tr>
<tr>
<td>Figure 13b</td>
<td>AP5 side: bin 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.7631</td>
<td>n = 96/n = 75</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (-105 to -71 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0839</td>
<td>n = 75/n = 66</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (-70 to -36 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5941</td>
<td>n = 137/n = 103</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (-35 to -1 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5941</td>
<td>n = 129/n = 96</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (0 to 34 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0181</td>
<td>n = 131/n = 102</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (35 to 69 trials from CP)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0008</td>
<td>n = 92/n = 85</td>
</tr>
<tr>
<td>Figure 13c</td>
<td>S+: VEH vs. AP5 side (70 to 104 trials</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0364</td>
<td>n = 36/n = 75</td>
</tr>
<tr>
<td>Figure</td>
<td>S+: VEH vs. AP5 side (from CP)</td>
<td>Firing 750-2000 ms after cue (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>13c</td>
<td>(-105 to -71 trials from CP)</td>
<td></td>
<td></td>
<td>0.9541</td>
<td>75/66</td>
</tr>
<tr>
<td>13c</td>
<td>(-70 to -36 trials from CP)</td>
<td></td>
<td></td>
<td>0.9541</td>
<td>137/103</td>
</tr>
<tr>
<td>13c</td>
<td>(-35 to -1 trials from CP)</td>
<td></td>
<td></td>
<td>0.1461</td>
<td>129/96</td>
</tr>
<tr>
<td>13c</td>
<td>(0 to 34 trials from CP)</td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
<td>131/102</td>
</tr>
<tr>
<td>13c</td>
<td>(35 to 69 trials from CP)</td>
<td></td>
<td></td>
<td>0.0002</td>
<td>92/85</td>
</tr>
<tr>
<td>13c</td>
<td>(70 to 104 trials from CP)</td>
<td></td>
<td></td>
<td>&lt;0.0001</td>
<td>36/75</td>
</tr>
<tr>
<td>13d</td>
<td>VEH side: trial bins 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.0014</td>
<td>75/129</td>
</tr>
<tr>
<td>13d</td>
<td>VEH side: trial bins 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.0004</td>
<td>129/92</td>
</tr>
<tr>
<td>13d</td>
<td>VEH side: trial bins 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.0006</td>
<td>75/129</td>
</tr>
<tr>
<td>13d</td>
<td>VEH side: trial bins 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.3471</td>
<td>129/92</td>
</tr>
<tr>
<td>13d</td>
<td>AP5 side: trial bins 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.2403</td>
<td>66/96</td>
</tr>
<tr>
<td>13d</td>
<td>AP5 side: trial bins 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.2403</td>
<td>96/75</td>
</tr>
<tr>
<td>13d</td>
<td>AP5 side: trial bins 1 vs. 3 (-105 to -71 vs. -35 to -1 trials from CP)</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.0689</td>
<td>66/96</td>
</tr>
<tr>
<td>13d</td>
<td>AP5 side: trial bins 3 vs. 5 (-35 to -1 vs. 35 to 69 trials from CP)</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.5356</td>
<td>96/75</td>
</tr>
<tr>
<td>13d</td>
<td>VEH vs. AP5 sides (-105 vs. -71 from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.7807</td>
<td>75/66</td>
</tr>
<tr>
<td>13d</td>
<td>VEH vs. AP5 sides (-70 vs. -36 from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.1395</td>
<td>137/103</td>
</tr>
<tr>
<td>13d</td>
<td>VEH vs. AP5 sides (-35 vs. -1 from CP)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p =</td>
<td>n =</td>
</tr>
<tr>
<td>13d</td>
<td></td>
<td></td>
<td></td>
<td>0.0363</td>
<td>129/96</td>
</tr>
<tr>
<td>13d</td>
<td>VEH vs. AP5 sides</td>
<td>Proportion of cue-excited</td>
<td>Fisher’s exact test</td>
<td>p =</td>
<td>n =</td>
</tr>
</tbody>
</table>
| 13d | | | | 0.0009 | 131/
<p>| Figure 13d | VEH vs. AP5 sides (0 vs. 34 from CP) | Proportion of cue-excited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.0002 | n = 92/85 |
| Figure 13d | VEH vs. AP5 sides (35 vs. 69 from CP) | Proportion of cue-excited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.0460 | n = 75/66 |
| Figure 13d | VEH vs. AP5 sides (-105 vs. -71 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.7076 | n = 137/103 |
| Figure 13d | VEH vs. AP5 sides (-70 vs. -36 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.7076 | n = 92/85 |
| Figure 13d | VEH vs. AP5 sides (0 vs. 34 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.7076 | n = 36/75 |
| Figure 13d | VEH vs. AP5 sides (35 vs. 69 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.7076 | n = 129/96 |
| Figure 13d | VEH vs. AP5 sides (-70 vs. -36 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.6126 | n = 131/102 |
| Figure 13d | VEH vs. AP5 sides (0 vs. 34 from CP) | Proportion of cue-inhibited neurons | Fisher’s exact test for count data (Holm-Sidak corrected) | p = 0.7076 | n = 36/75 |
| Figure 13f | S+: VEH vs. AP5 side (-105 to -71 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.4073 | n = 19/14 |
| Figure 13f | S+: VEH vs. AP5 side (-70 to -36 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.4073 | n = 45/24 |
| Figure 13f | S+: VEH vs. AP5 side (-35 to -1 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.4835 | n = 63/31 |
| Figure 13f | S+: VEH vs. AP5 side (0 to 34 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.4835 | n = 87/43 |
| Figure 13f | S+: VEH vs. AP5 side (35 to 69 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.4835 | n = 56/37 |
| Figure 13f | S+: VEH vs. AP5 side (70 to 104 trials from CP) | Firing 100-400 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.2173 | n = 23/36 |
| Figure 13f | S+: VEH vs. AP5 side (-105 to -71 trials from CP) | Firing 750-2000 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.5289 | n = 19/14 |
| Figure 13f | S+: VEH vs. AP5 side (-70 to -36 trials from CP) | Firing 750-2000 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.5289 | n = 45/24 |
| Figure 13f | S+: VEH vs. AP5 side (-35 to -1 trials from CP) | Firing 750-2000 ms after cue (Z sc.) | Wilcoxon rank sum test (Holm-Sidak corrected) | p = 0.2173 | n = 63/31 |</p>
<table>
<thead>
<tr>
<th>Figure</th>
<th>S+ vs. AP5 side</th>
<th>Firing 750-2000 ms after cue (Z sc.)</th>
<th>Wilcoxon rank sum test (Holm-Sidak corrected)</th>
<th>p</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>13f</td>
<td>(0 to 34 trials from CP)</td>
<td>Cue-excited units</td>
<td></td>
<td>0.0162</td>
<td>87/43</td>
</tr>
<tr>
<td>13f</td>
<td>(35 to 69 trials from CP)</td>
<td>Cue-excited units</td>
<td></td>
<td>0.0049</td>
<td>56/37</td>
</tr>
<tr>
<td>13f</td>
<td>(70 to 104 trials from CP)</td>
<td>Cue-excited units</td>
<td></td>
<td>0.0002</td>
<td>23/36</td>
</tr>
<tr>
<td>14a</td>
<td>S+ vs. S- 5 trial bin “Learners”</td>
<td>Performance index</td>
<td>Two-factor rep. measures ANOVA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Within-subject:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Cue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Bin.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Cue:</strong> $F_{1,5} = 119.926$, $p = 0.0001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Bin:</strong> $F_{1.06, 5.3} = 2.38$, $p = 0.1170$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Cue x bin:</strong> $F_{0.86, 3.35} = 10.527$, $p = 0.0023$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14b</td>
<td>VEH vs. AP5 side “Learners”</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data</td>
<td>0.3073</td>
<td>38/39</td>
</tr>
<tr>
<td>14c</td>
<td>VEH vs. AP5 side “Learners”</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>&lt;0.0001</td>
<td>38/39</td>
</tr>
<tr>
<td>14c</td>
<td>VEH vs. AP5 side “Learners”</td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>&lt;0.0001</td>
<td>38/39</td>
</tr>
<tr>
<td>14c</td>
<td>VEH vs. AP5 side “Learners”</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>&lt;0.0001</td>
<td>38/39</td>
</tr>
<tr>
<td>14c</td>
<td>VEH vs. AP5 side “Learners”</td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>&lt;0.0001</td>
<td>38/39</td>
</tr>
</tbody>
</table>
### 4.2.2. Statistical tests. Supplementary figures.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Independent variable(s)</th>
<th>Dependent variable</th>
<th>Test</th>
<th>Result</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1a</td>
<td>S+ vs. S- Session (1 to 6)</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Mixed two-factor ANOVA</td>
<td>Session: $F_{(5, 358)} = 4.87, p = 0.0003$&lt;br&gt;Between-subject: session&lt;br&gt;Within-subject: kind of cue</td>
<td>n = 186</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 1</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0013</td>
<td>n = 35</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 2</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0146</td>
<td>n = 32</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 3</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.4025</td>
<td>n = 32</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 4</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0002</td>
<td>n = 39</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 5</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 22</td>
</tr>
<tr>
<td>S2a</td>
<td>S+ vs. S- Session 6</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 26</td>
</tr>
<tr>
<td>S2b</td>
<td>S+ vs S- Session 1. Trials 1-10.</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.3320</td>
<td>n = 35</td>
</tr>
<tr>
<td>S2b</td>
<td>S+ vs S- Session 1. Trials 11-20.</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0400</td>
<td>n = 35</td>
</tr>
<tr>
<td>S2b</td>
<td>S+ vs S- Session 1. Trials 21-30.</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0076</td>
<td>n = 35</td>
</tr>
<tr>
<td>S2b</td>
<td>S+ vs S- Session 1. Trials 31-40.</td>
<td>Firing 100-400 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0081</td>
<td>n = 35</td>
</tr>
<tr>
<td>S2c</td>
<td>S+ sensory modality (light vs. tone) Before vs. After CP</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Two-factor ANOVA</td>
<td>S+ modality: $F_{(1, 145)} = 0.006, p = 0.9403$&lt;br&gt;Pre/post CP: $F_{(1, 145)} = 29.355, p &lt; 0.0001$&lt;br&gt;Interaction: $F_{(1, 145)} = 0.0778, p = 0.7805$</td>
<td>n = 149</td>
</tr>
<tr>
<td>S2c</td>
<td>Light S+: pre vs. post CP</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 70/ n = 6</td>
</tr>
<tr>
<td>S2c</td>
<td>Tone S+: pre vs. post CP</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.0247</td>
<td>n = 27/ n = 46</td>
</tr>
<tr>
<td>S2c</td>
<td>Pre CP: tone vs. light S+</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.2718</td>
<td>n = 27/ n = 70</td>
</tr>
<tr>
<td>S2c</td>
<td>Post CP: tone vs. light S+</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5278</td>
<td>n = 46/ n = 6</td>
</tr>
<tr>
<td>S2d</td>
<td>S+ vs. S- (-120 to -81 trials from CP)</td>
<td>Firing 750-2000 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.2815</td>
<td>n = 45</td>
</tr>
<tr>
<td>S2d</td>
<td>S+ vs. S- (-80 to -41 trials from CP)</td>
<td>Firing 750-2000 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.2815</td>
<td>n = 55</td>
</tr>
<tr>
<td>S2d</td>
<td>S+ vs. S- (-40 to -1 trials from CP)</td>
<td>Firing 750-2000 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 68</td>
</tr>
<tr>
<td>S2d</td>
<td>S+ vs. S-</td>
<td>Firing 750-2000 ms after cue (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p &lt; 0.0001</td>
<td>n = 63</td>
</tr>
<tr>
<td>S2d</td>
<td>Trial bins 1 vs. 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-120 to -61 vs. -40 to 0 trials from CP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p = 0.1151</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 45 /</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2d</th>
<th>Trial bins 3 vs. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-40 to -1 vs. 40 to 79 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.086</td>
</tr>
<tr>
<td></td>
<td>n = 68 /</td>
</tr>
<tr>
<td></td>
<td>n = 37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2d</th>
<th>Trial bins 2 vs. 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-80 to -41 vs. 0 to 39 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>n = 55 /</td>
</tr>
<tr>
<td></td>
<td>n = 63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2d</th>
<th>Trial bins 4 vs. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0 to 39 vs. 80 to 119 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing 750-2000 ms after S+ (Z sc.)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.0481</td>
</tr>
<tr>
<td></td>
<td>n = 63 /</td>
</tr>
<tr>
<td></td>
<td>n = 17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 1 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-120 to -61 vs. -40 to 0 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.9989</td>
</tr>
<tr>
<td></td>
<td>n = 42 /</td>
</tr>
<tr>
<td></td>
<td>n = 31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 3 vs. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-40 to -1 vs. 40 to 79 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.6142</td>
</tr>
<tr>
<td></td>
<td>n = 38 /</td>
</tr>
<tr>
<td></td>
<td>n = 51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 2 vs. 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-80 to -41 vs. 0 to 39 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.0285</td>
</tr>
<tr>
<td></td>
<td>n = 68 /</td>
</tr>
<tr>
<td></td>
<td>n = 47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 4 vs. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0 to 39 vs. 80 to 119 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.0177</td>
</tr>
<tr>
<td></td>
<td>n = 63 /</td>
</tr>
<tr>
<td></td>
<td>n = 63</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 1 vs. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-120 to -61 vs. -40 to 0 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.029</td>
</tr>
<tr>
<td></td>
<td>n = 12 /</td>
</tr>
<tr>
<td></td>
<td>n = 22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2e</th>
<th>Trial bins 3 vs. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-40 to -1 vs. 40 to 79 trials from CP)</td>
</tr>
<tr>
<td></td>
<td>Firing -2000-0 ms (Z sc.) before entry</td>
</tr>
<tr>
<td></td>
<td>(when latency &gt; 5 s)</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
</tr>
<tr>
<td></td>
<td>p = 0.0062</td>
</tr>
<tr>
<td></td>
<td>n = 11 /</td>
</tr>
<tr>
<td></td>
<td>n = 17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3b</th>
<th>Before change point session.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average firing 100-400 ms after S+ onset (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Before change point session.</td>
</tr>
<tr>
<td></td>
<td>Average firing 0-1500 ms after S- entry (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Simple linear regression</td>
</tr>
<tr>
<td></td>
<td>Without outliers: r = -0.22, p = 0.0582; R^2 = 0.05; β=-0.64, p = 0.0542</td>
</tr>
<tr>
<td></td>
<td>n = 97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3b</th>
<th>Before change point session.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average firing 100-400 ms after S+ onset (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Before change point session.</td>
</tr>
<tr>
<td></td>
<td>Average firing 0-1500 ms after ITI entry (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Simple linear regression</td>
</tr>
<tr>
<td></td>
<td>Without outliers: r = -0.09, p = 0.3993; R^2 = 0.01; β=-0.15, p = 0.3993</td>
</tr>
<tr>
<td></td>
<td>n = 97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3c</th>
<th>On or after change point session.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average firing 100-400 ms after S+ onset (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>On or after change point session.</td>
</tr>
<tr>
<td></td>
<td>Average firing 0-1500 ms after S- entry (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Simple linear regression</td>
</tr>
<tr>
<td></td>
<td>Without outliers: r = -0.19, p = 0.0905; R^2 = 0.03; β=0.15, p = 0.0905</td>
</tr>
<tr>
<td></td>
<td>n = 89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S3c</th>
<th>On or after change point session.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average firing 100-400 ms after S+ onset (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>On or after change point session.</td>
</tr>
<tr>
<td></td>
<td>Average firing 0-1500 ms after ITI entry (Z sc.).</td>
</tr>
<tr>
<td></td>
<td>Simple linear regression</td>
</tr>
<tr>
<td></td>
<td>Without outliers: r = -0.07, p = 0.5357; R^2 = 0; β=-0.2, p = 0.5357</td>
</tr>
<tr>
<td></td>
<td>n = 89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S4a</th>
<th>Before vs. after change S+ performance index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paired t-test</td>
</tr>
<tr>
<td></td>
<td>t_{12} = -7.94,</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
</tr>
<tr>
<td>point</td>
<td>(Holm-Sidak corrected)</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------</td>
</tr>
<tr>
<td>S4a</td>
<td>Before vs. after change point</td>
</tr>
<tr>
<td>S4a</td>
<td>Before vs. after change point</td>
</tr>
<tr>
<td>S4a</td>
<td>Before vs. after change point</td>
</tr>
<tr>
<td>S4b</td>
<td>Pre CP: driving arrays down vs. not driving arrays down</td>
</tr>
<tr>
<td>S4b</td>
<td>Post CP: driving arrays down vs. not driving arrays down</td>
</tr>
<tr>
<td>S4b</td>
<td>Driving arrays down: pre CP vs. post CP</td>
</tr>
<tr>
<td>S4c</td>
<td>S+ vs. S- firing. Channel average, day before CP. Not driving arrays down.</td>
</tr>
<tr>
<td>S4c</td>
<td>S+ vs. S- firing. Channel average, day after CP. Not driving arrays down.</td>
</tr>
<tr>
<td>S4c</td>
<td>S+ firing. Channel average. Day before vs. after CP. Not driving arrays down.</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 1 to 35)</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 36 to 70)</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 71 to 105)</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 106 to 140)</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 141 to 175)</td>
</tr>
<tr>
<td>S5d</td>
<td>S+ vs. S- (Trial 176 to 210)</td>
</tr>
<tr>
<td>S5e</td>
<td>% excited vs. % non-excited by bin (5 session)</td>
</tr>
<tr>
<td>S5e</td>
<td>% inhibited vs. % non-inhibited by bin (5 session)</td>
</tr>
</tbody>
</table>

<p>| S6a   | AP5s vs. VEH group Pre vs. post infusion Moderate training | S+ entry probability | Mixed two-factor ANOVA: Between-subject: drug Within-subject: pre/post inf. | Drug: F_{(1, 9)} = 50.991, p &lt; 0.0001 | n = 6/ n = 5 |
| S6a   | VEH group: pre vs. post infusion Moderate training. | S+ entry probability | Paired t-test (Holm-Sidak corrected) | t_{(5)} = -0.445, p = 0.6624 | n = 6/ n = 5 |
| S6a   | AP5 group: pre vs. post infusion Moderate training. | S+ entry probability | Paired t-test (Holm-Sidak corrected) | t_{(5)} = -3.504, p = 0.0248 | n = 6/ n = 5 |
| S6a   | AP5 vs. VEH group Pre vs. post infusion Moderate training | S- entry probability | Mixed two-factor ANOVA: Between-subject: drug Within-subject: pre/post inf. | Drug: F_{(1, 9)} = 0.517, p = 0.4905 | n = 6/ n = 5 |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Group</th>
<th>Condition</th>
<th>Measure</th>
<th>ANOVA Details</th>
<th>Example Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>S6b</td>
<td>AP5 vs. VEH group</td>
<td>Pre vs. post infusion</td>
<td>S+ latency</td>
<td>Mixed two-factor ANOVA: Between-subject: drug</td>
<td>Drug: $F_{(1,9)} = 88.274, p &lt; 0.0001$ Pre/post infusion: $F_{(1,9)} = 11.002, p = 0.009$ Interaction: $F_{(1,9)} = 12.038, p = 0.007$</td>
</tr>
<tr>
<td>S6b</td>
<td>VEH group</td>
<td>Pre vs. post infusion</td>
<td>S+ latency</td>
<td>Paired t-test</td>
<td>$t_{(6)} = 0.709, p = 0.7450$</td>
</tr>
<tr>
<td>S6b</td>
<td>AP5 group</td>
<td>Pre vs. post infusion</td>
<td>S+ latency</td>
<td>Paired t-test</td>
<td>$t_{(4)} = -3.085, p = 0.0367$</td>
</tr>
<tr>
<td>S6b</td>
<td>AP5 vs. VEH group</td>
<td>Pre vs. post infusion</td>
<td>S- latency</td>
<td>Mixed two-factor ANOVA: Between-subject: drug</td>
<td>Drug: $F_{(1,9)} = 0.054, p = 0.8217$ Pre/post infusion: $F_{(1,9)} = 4.408, p = 0.0651$ Interaction: $F_{(1,9)} = 2.7, p = 0.14$</td>
</tr>
<tr>
<td>S6c</td>
<td>VEH group</td>
<td>Pre vs. post infusion</td>
<td>ITI pseudolatency</td>
<td>Paired t-test</td>
<td>$t_{(6)} = -0.299, p = 0.3881$</td>
</tr>
<tr>
<td>S6c</td>
<td>AP5 group</td>
<td>Pre vs. post infusion</td>
<td>ITI pseudolatency</td>
<td>Paired t-test</td>
<td>$t_{(4)} = -2.916, p = 0.0434$</td>
</tr>
<tr>
<td>S6d</td>
<td>AP5 vs. VEH infusion</td>
<td>Pre vs. post infusion</td>
<td>S+ entry probability</td>
<td>Two-factor rep. measures ANOVA: Within-subject: Drug</td>
<td>Drug: $F_{(1,4)} = 0.4490, p = 0.5395$ Pre/post infusion: $F_{(1,4)} = 5.943, p = 0.0713$ Interaction: $F_{(1,4)} = 0.4490, p = 0.5395$</td>
</tr>
<tr>
<td>S6d</td>
<td>AP5 vs. VEH infusion</td>
<td>Pre vs. post infusion</td>
<td>S- entry probability</td>
<td>Two-factor rep. measures ANOVA: Within-subject: Drug</td>
<td>Drug: $F_{(1,4)} = 0.062, p = 0.8161$ Pre/post infusion: $F_{(1,4)} = 0.859, p = 0.4063$ Interaction: $F_{(1,4)} = 0.033, p = 0.8640$</td>
</tr>
<tr>
<td>S6e</td>
<td>AP5 vs. VEH infusion</td>
<td>Pre vs. post infusion</td>
<td>S+ latency</td>
<td>Two-factor rep. measures ANOVA: Within-subject: Drug</td>
<td>Drug: $F_{(1,4)} = 0.448, p = 0.5401$ Pre/post infusion: $F_{(1,4)} = 4.153, p = 0.1112$ Interaction: $F_{(1,4)} = 1.325, p = 0.3138$</td>
</tr>
<tr>
<td>S6e</td>
<td>AP5 vs. VEH infusion</td>
<td>Pre vs. post infusion</td>
<td>S- latency</td>
<td>Two-factor rep. measures ANOVA: Within-subject: Drug</td>
<td>Drug: $F_{(1,4)} = 0.002, p = 0.9698$ Pre/post infusion: $F_{(1,4)} = 1.743, p = 0.2572$ Interaction: $F_{(1,4)} = 0.116, p = 0.7499$</td>
</tr>
<tr>
<td>S6f</td>
<td>AP5 vs. VEH infusion</td>
<td>Pre vs. post infusion</td>
<td>ITI pseudolatency</td>
<td>Two-factor rep. measures ANOVA: Within-subject: Drug</td>
<td>Drug: $F_{(1,4)} = 0.002, p = 0.9694$ Pre/post infusion: $F_{(1,4)} = 1.324, p = 0.3140$ Interaction: $F_{(1,4)} = 2.783, p = 0.1706$</td>
</tr>
</tbody>
</table>

<p>| S7a | VEH vs. AP5 sides | Baseline firing rate | Wilcoxon rank sum test | $p = 1$ | $n = 75$ |</p>
<table>
<thead>
<tr>
<th>(-105 vs. -71 from CP)</th>
<th>2000-0 ms pre S+</th>
<th>(Holm-Sidak corrected)</th>
<th>n = 66</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S7a</strong> VEH vs. AP5 sides (-70 vs. -36 from CP)</td>
<td>Baseline firing rate (-2000-0 ms pre S+)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5617</td>
</tr>
<tr>
<td><strong>S7a</strong> VEH vs. AP5 sides (-35 vs. -1 from CP)</td>
<td>Baseline firing rate (-2000-0 ms pre S+)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5617</td>
</tr>
<tr>
<td><strong>S7a</strong> VEH vs. AP5 sides (0 vs. 34 from CP)</td>
<td>Baseline firing rate (-2000-0 ms pre S+)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 0.5617</td>
</tr>
<tr>
<td><strong>S7a</strong> VEH vs. AP5 sides (35 vs. 69 from CP)</td>
<td>Baseline firing rate (-2000-0 ms pre S+)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 1</td>
</tr>
<tr>
<td><strong>S7a</strong> VEH vs. AP5 sides (70 vs. 105 from CP)</td>
<td>Baseline firing rate (-2000-0 ms pre S+)</td>
<td>Wilcoxon rank sum test (Holm-Sidak corrected)</td>
<td>p = 1</td>
</tr>
</tbody>
</table>

**S7b** Previously VEH vs. AP5 sides (Extinction test, "learners") Baseline firing rate (-2000-0 ms pre S+) Wilcoxon rank sum test p = 0.5185 n = 38/ n = 39

**S8b** VEH vs. AP5 (-105 to -71 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.681 n = 21/ n = 15

**S8b** VEH vs. AP5 (-70 to -36 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.2246 n = 22/ n = 20

**S8b** VEH vs. AP5 (-35 to -1 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.0329 n = 25/ n = 15

**S8b** VEH vs. AP5 (0 to 34 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.0404 n = 28/ n = 11

**S8b** VEH vs. AP5 (35 to 69 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.0009 n = 15/ n = 10

**S8b** VEH vs. AP5 (70 to 105 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.0001 n = 4/ n = 9

**S8b** VEH side. Trial bins 3 vs. 3 (-105 to -71 trials to -35 to -1 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.1732 n = 21/ n = 25

**S8b** VEH side. Trial bins 3 vs. 3 (-35 to -1 vs. 35 to 69 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.0088 n = 25/ n = 15

**S8b** APS side. Trial bins 3 vs. 3 (-105 to -71 vs. 35 to -1 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.7568 n = 15/ n = 15

**S8b** APS side. Trial bins 3 vs. 3 (-35 to -1 vs. 35 to 69 trials from CP) Firing -2000-0 ms (Z sc.) before S+ entry (when latency > 5 s) Wilcoxon rank sum test (Holm-Sidak corrected) p = 0.7694 n = 15/ n = 10

**S8c** Before change point session. VEH vs. AP5 Proportion of S+-entry-excited units Fisher’s exact test for count data (Holm-Sidak corrected) p = 0.2762 n = 160/ n = 110

**S8c** Before change point session. VEH vs. AP5 Proportion of S+-entry-inhibited units Fisher’s exact test for count data (Holm-Sidak corrected) p = 0.1455 n = 160/ n = 110

**S8c** Before change point session. Entry-excited units. VEH vs. AP5 Firing rate (0-1500 ms after S+ entry) Wilcoxon rank sum test p = 0.9864 n = 33/ n = 29

**S8c** Before change point session. Entry-inhibited units. VEH vs. AP5 Firing rate (0-1500 ms after S+ entry) Wilcoxon rank sum test p = 0.4064 n = 62/ n = 37

**S8c** After change point session. VEH vs. AP5 Proportion of S+-entry-excited units Fisher’s exact test for count data (Holm-Sidak corrected) p = 0.7939 n = 160/ n = 110

**S8c** After change point session. VEH vs. AP5 Proportion of S+-entry-inhibited units Fisher’s exact test for count data (Holm-Sidak corrected) p = 0.6250 n = 160/ n = 110

**S8c** After change point session. Entry-excited units. VEH vs. AP5 Firing rate (0-1500 ms after S+ entry) Wilcoxon rank sum test p = 0.6102 n = 35/ n = 40

**S8c** After change point session. Entry-inhibited Firing rate (0-1500 ms after S+ entry) Wilcoxon rank sum test p = 0.5586 n = 63/ n = 69
<table>
<thead>
<tr>
<th>S9e</th>
<th>VEH side: S+ vs. S- (0-35 trial bin)</th>
<th>Firing 100-400 ms after S+ (Z sc.)</th>
<th>Wilcoxon signed-rank test (Holm-Sidak corrected)</th>
<th>p = 0.9788</th>
<th>n = 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>S9e</td>
<td>VEH side: S+ vs. S- (36-70 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 81</td>
</tr>
<tr>
<td>S9e</td>
<td>VEH side: S+ vs. S- (71-105 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0122</td>
<td>n = 86</td>
</tr>
<tr>
<td>S9e</td>
<td>VEH side: S+ vs. S- (106-140 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 38</td>
</tr>
<tr>
<td>S9e</td>
<td>VEH side: S+ vs. S- (141-175 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 18</td>
</tr>
<tr>
<td>S9e</td>
<td>VEH side: S+ vs. S- (176-210 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.0276</td>
<td>n = 26</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (1-35 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.6820</td>
<td>n = 38</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (36-70 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.6776</td>
<td>n = 24</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (71-105 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.6776</td>
<td>n = 36</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (106-140 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.6776</td>
<td>n = 25</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (141-175 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.5404</td>
<td>n = 20</td>
</tr>
<tr>
<td>S9e</td>
<td>AP5 side: S+ vs. S- (176-210 trial bin)</td>
<td>Firing 100-400 ms after S+ (Z sc.)</td>
<td>Wilcoxon signed-rank test (Holm-Sidak corrected)</td>
<td>p = 0.1238</td>
<td>n = 15</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 35 trial bins (sessions)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0301</td>
<td>n = 304</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 1-35 vs. 36-70 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 55/81</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 36-70 vs. 71-105 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 81/86</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 71-105 vs. 106-140 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.7165</td>
<td>n = 86/38</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 106-140 vs. 141-175 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 38/18</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 141-175 vs. 176-210 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0465</td>
<td>n = 18/26</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH side: 35 trial bins (sessions)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.2596</td>
<td>n = 304</td>
</tr>
<tr>
<td>S9f</td>
<td>AP5 side: 35 trial bins (sessions)</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.5149</td>
<td>n = 158</td>
</tr>
<tr>
<td>S9f</td>
<td>AP5 side: 35 trial bins (sessions)</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.164</td>
<td>n = 158</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 1-35 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 55/38</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 36-70 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 81/24</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 71-105 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher's exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 86/36</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 106-140 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 38/ n = 25</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>----------------</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 141-175 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 18/ n = 20</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 176-210 trial bin.</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 26/ n = 15</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 1-35 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 55/ n = 38</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 36-70 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.1671</td>
<td>n = 81/ n = 24</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 71-105 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.0579</td>
<td>n = 86/ n = 36</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 106-140 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 38/ n = 25</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 141-175 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 1</td>
<td>n = 18/ n = 20</td>
</tr>
<tr>
<td>S9f</td>
<td>VEH vs. AP5 side: 176-210 trial bin.</td>
<td>Proportion of cue-inhibited neurons</td>
<td>Fisher’s exact test for count data (Holm-Sidak corrected)</td>
<td>p = 0.7728</td>
<td>n = 26/ n = 15</td>
</tr>
<tr>
<td>S9h</td>
<td>VEH vs. AP5 side</td>
<td>Proportion of cue-excited neurons</td>
<td>Fisher’s exact test for count data</td>
<td>p = 1</td>
<td>n = 26/ n = 15</td>
</tr>
</tbody>
</table>
5. REFERENCES


https://doi.org/10.1523/JNEUROSCI.2246-11.2011


https://doi.org/10.1146/annurev.ne.09.030186.002041


https://doi.org/10.1016/j.neuron.2008.07.004

https://doi.org/10.1126/science.8023166

https://doi.org/10.1016/j.tics.2010.02.001


core and shell in the “Accumbens” part of the rat ventral striatum: Immunohistochemical detection of retrogradely transported fluoro-gold. *Journal of Comparative Neurology*, 338(2), 255–278. https://doi.org/10.1002/cne.903380209


https://doi.org/10.1016/S0306-4522(99)00214-6


https://doi.org/10.1016/j.neuron.2016.09.025

https://doi.org/10.1152/jn.1996.75.4.1659

https://doi.org/10.1152/jn.00473.2002


https://doi.org/10.1007/BF00711085


https://doi.org/10.1021/acschemneuro.6b00307


Horvitz, J. C. (2002). Dopamine gating of glutamatergic sensorimotor and incentive motivational input signals to the striatum. *Behavioural Brain Research, 137*(1), 65–74. https://doi.org/10.1016/S0166-


Localization of D1 and D2 dopamine receptors in brain with subtype-specific antibodies. 

*Proceedings of the National Academy of Sciences of the United States of America*, 90(19), 8861–8865.


https://doi.org/10.1016/0306-4522(83)90010-6


Maldonado-Irizarry, Carmen S., & Kelley, A. E. (1994). Differential behavioral effects following...


https://doi.org/10.1152/jn.00658.2003

https://doi.org/10.1016/j.neuron.2016.12.015


https://doi.org/10.1126/science.1094285

https://doi.org/10.1016/S0896-6273(03)00169-7


https://doi.org/10.1038/nature04490


https://doi.org/10.1523/JNEUROSCI.4402-07.2008

https://doi.org/10.3389/fnsyn.2010.00146


transmission in the nucleus accumbens only in rats having developed behavioral sensitization.


https://doi.org/10.1126/science.167.3926.1740


https://doi.org/10.1126/science.1239073


https://doi.org/10.1523/JNEUROSCI.12-12-04595.1992


Sesack, S. R., & Pickel, V. M. (1990). In the rat medial nucleus accumbens, hippocampal and catecholaminergic terminals converge on spiny neurons and are in apposition to each other. Brain Research, 527(2), 266–279. https://doi.org/10.1016/0006-8993(90)91146-8


Temporal Specificity of Reward
Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum.

*Neuron, 91*(1), 182–193. https://doi.org/10.1016/j.neuron.2016.05.015


https://doi.org/10.3389/fnana.2010.00150


