4-22-2024

Inferring Function from Form: Subcortical Projection Cell Types in Rat Orbitofrontal Cortex

Suelynn Ren
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST.LOUIS

Division of Biology and Biomedical Sciences
Neurosciences

Dissertation Examination Committee:
Adam Kepecs, Chair
Joseph Dougherty
Tamara Hershey
Bo Li
Ilya Monosov

Inferring Function from Form: Subcortical Projection Cell Types in Rat Orbitofrontal Cortex

by
Suelynn Ren

A dissertation presented to
Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2024
St. Louis, Missouri

Table of Contents

List of Figures

Acknowledgements

I am forever grateful for the support of my scientific mentors and collaborators. First, I would like to thank my thesis committee members Tamara Hershey, Ilya Monosov, Joseph Dougherty, Bo Li, Alfredo Fontanini, and Anthony Zador for their support and invaluable advice. Second, most of the experiments and analyses presented in this thesis would not be possible without our collaborators: Thomas Klausberger and Ben-Orli Nathanson, Pavel Osten and Rodrigo Munoz-Castaneda, Longwen Huang, Melody Wu and Jessica Tollkuhn, and Alexander R. Nectow. Last, I would like to thank the scientific mentors who were present at the very start of my scientific adventure, Paul Clarke and Eugene Kiyatkin.

I was incredibly lucky enough to be in the right place at the right time. I thank Adam Kepecs for fostering an exciting and ambitious scientific environment. Without his scientific and professional support, I would never have developed the breadth and depth of knowledge necessary to tackle the questions laid out in this thesis. Further, I have been fortunate to train in multiple supportive and collaborative environments, at the mesoscale in the Kepecs Laboratory and the macroscale at Cold Spring Harbor Laboratory and Washington University in St Louis. I thank the entire Marks community and members of the Kepecs Laboratory, both past and present. In specific, I would like to thank my scientific partner and friend, Torben Ott, who stood with me throughout many scientific and personal pitfalls.

The efforts and work presented in this thesis were partially supported by the National Institute of

Mental Health and the Ruth L. Kirschstein National Research Service Award (F30 MH120935-01A1).

Suelynn Ren

Washington University in St. Louis

May 2024

This work is dedicated to the ducks paddling frantically and my family.

ABSTRACT OF THE DISSERTATION

Inferring Function from Form: Subcortical Projection Cell Types in Rat Orbitofrontal Cortex

by

Suelynn Ren

Doctor of Philosophy in Biology and Biomedical Sciences

Neurosciences

Washington University in St Louis, 2024

Professor Adam Kepecs, Chair

Frontal cortex supports sophisticated behaviors by controlling phylogenetically older subcortical brain regions. Within the deep layers of cortex, subcortically projecting pyramidal neurons integrate local and long-range inputs along the entire depth of the cortical column. How subcortically projecting neurons might be fractionated into cell types and how their unique features constrain top-down communication is unknown. In this dissertation, I first established the output circuit architecture of subcortically projecting neurons in rat orbitofrontal cortex at a cellular resolution. I found that subcortically projecting neurons preferentially innervated a single target, demonstrating a one-neuron-one-target projection logic. Such target-defined projection neurons were molecularly distinct and spatially segregated into previously unappreciated, intermediate sublayers of L5b. These anatomical results suggest that subcortically projecting neurons represent connectivity defined cell types which are positioned to act as specialized information channels. To test the functional roles of OFC subcortical projections, I next developed two task variants that enable testing for (1) static and dynamic learning rate representations by manipulating the outcome distribution and (2) subjective value representations that could drive dopaminergic signaling and

in turn, momentary choice preferences. Taken together, this work reveals the higher-order organizational principles of rat orbitofrontal cortex, namely a “one neuron one target” projection logic and a “deeper farther” spatial logic. Such a highly structured circuit architecture bolters the hypothesis that subcortically projecting neurons might serve as the final arbiters, routing critical information that can drive both healthy and maladaptive behaviors.

Chapter 1

Finding the right level of granularity in frontal cortex

1.1 Introduction: a circuit-driven understanding of choice behaviors

It is increasingly clear that there is no magic bullet for psychiatric disease (Fernandes et al., 2017; Insel, 2014; Insel and Quirion, 2005). Advanced technologies have pushed many fields of biomedicine towards precision medicine and targetable cell types. While the same technical advances enabled neuroscientists to uncover new neuronal cell types and characterize the genetic architecture of psychiatric diseases, the translational relevance of this progress has been less visible. First, this is because behavior arises from the complex interactions of networks of neuron and second, there is no true pathological cell type in psychiatric disease. Rather, maladaptive behaviors from poor decision making to anhedonia are likely the result of multiple circuit

dysfunctions that corrupt core computations within the brain (Gottesman and Gould, 2003; Montague et al., 2012).

Aberrant decision-making is a core feature of nearly all psychiatric diseases (Bechara and Damasio, 2002; Bechara et al., 2002; Brand et al., 2005; Chamberlain et al., 2007; Ersche et al., 2006; Heerey et al., 2008; Lawrence et al., 2006). Behavioral manifestations of mental dysfunctions can be incredibly varied and are thought to arise from subtle alterations of cortical circuits as well as impaired cortical-to-subcortical communication (i.e., top-down control) (Dalley et al., 2011; Groman and Jentsch, 2012; Gueguen et al., 2021; Petrovic and Castellanos, 2016; Renteria et al., 2018). OFC dysfunction is implicated in numerous psychiatric disorders and orbitofrontal cortex has recently been conceptualized as the hub of the valuation system. OFC is densely interconnected with critical subcortical valuation and associative learning centers (Gremel and Costa, 2013; Gremel et al., 2016; Harada et al., 2021; Hirokawa et al., 2019; Lichtenberg et al., 2017; Malvaez et al., 2019; Pascoli et al., 2018; Sias et al., 2021; Takahashi et al., 2009a). Neuroimaging studies point to metabolic and structural abnormalities and impaired communication with regions such as the basal ganglia but lack the resolution to pinpoint neural populations (Altshuler et al., 2005; Beucke et al., 2013; Bremner et al., 2002; Girgis et al., 2007; Levy and Dubois, 2006; Volkow et al., 2003, 2011). A relatively small population of projection neurons directs cortical-to-subcortical communication (Gabbott and Stewart, 1987; Zhang et al., 2021). Residing within layer 5b, these subcortically projecting neurons constitute an information bottleneck that might influence behavior and if corrupted, might result in cascading dysfunction (Adesnik and Naka, 2018; Shepherd and Rowe, 2017; Sherman and Usrey, 2021; Tosches and Laurent, 2019). Subcortically projecting neurons are highly diverse in terms of their anatomy,

physiology and subcortical targets, it is unclear how many distinct types there are and how these might differentially contribute to choice behavior.

The goals of this work are to first, establish a blueprint of projection cell types within orbitofrontal cortex and second, to develop a rich family of choice behaviors that can be used to determine if such cell types encode stable representations and contribute to different behavioral processes. Here, I establish the output architecture of orbitofrontal cortex and its higher-order organizational principles, revealing a previously unappreciated "deeper further" sublaminar structure that might reflect "inside out" development. Further, I developed a family of choice behaviors and reinforcement learning models that can be used to fractionate choice behaviors into their underlying quantifiable processes and parameters (i.e., decision variables), which in future, can be mapped to cell type specific representations. Such a descriptive language constitutes a computational fingerprint that can be tracked across learning stages, task contexts, and healthy and diseased states. Together, this work lays the foundation to generate a functional cell type map that can used to both make and test mechanistic hypotheses about orbitofrontal cortex.

1.2 Projection patterns as the first plane of dissection for functional cell types

Technical advances in single cell technologies have revealed immense diversity within developmental neuron classes, raising the question as to what represents a meaningful neuron type (or subtype) (Fishell and Heintz, 2013; Gouwens et al., 2019; Muñoz-Castañeda et al., 2020; Usoskin et al., 2015; Zhang et al., 2021). Indeed, the criteria to subdivide cell types is open. From an evolutionary perspective, a cell type can be defined by its regulatory independence (Arendt et al., 2016; Tosches et al., 2018). That is, a cell type has unique transcriptional responses that

required the recruitment of an additional transcription factor. Importantly, while this definition is the useful for cross-species comparisons, it does not assume a unified function. On the other hand, from a utilitarian and systems neuroscience perspective, a cell type's circuit function is the defining feature that should arise from the others. This functional definition of a cell type brings together fields from neuroanatomy and developmental biology to systems neuroscience as it aims to establish an evolutionary prior for the distribution of information in the brain. Here, we argue that projection patterns are a principled means for dissecting functionally relevant cell types.

The six-layered structure of mammalian cortex is its most striking architectural feature and was the first evidence of distinct neuronal types (DeFelipe, Javier, and Edward G. Jones., 1988). Pyramidal neurons are the main excitatory neurons in cortex with distinct neuron types constituting each layer. First-born neurons are destined for the deepest layer and later born neurons migrate sequentially to more superficial layers in an inside-out manner (McConnell and Kaznowski, 1991a; O'Leary and Koester, 1993; Rakic, 1974). The environmental factors present at the time of a neuron's birth (or more precisely, its final mitotic division) are thought to drive neurons down a pre-set developmental path that restricts its gene expression patterns and laminar fate. For instance, Fezf2 regulates the differentiation of extra-telencephalic neurons in deep layer 5b and prevents the expression of intra-telencephalic transcription factors like Satb2, but beyond these early developmental branchpoints, it is unclear how subtype identity might be further refined (Alcamo et al., 2008; Britanova et al., 2008; Chen et al., 2005, 2008; McConnell and Kaznowski, 1991a; Yao et al., 2021). Indeed, broad classes of projection neurons can be differentiated by their stereotyped laminar position, dendritic morphology, and electrophysiological properties. Less is known, however, about how many subtypes exist across cortex.

The advent of single cell sequencing technologies rapidly propelled the discovery of novel neuronal subtypes. High-throughput approaches such as droplet-based and nuclear single cell sequencing made it possible to capture the transcriptomes of tens of thousands of neurons across species, from jellyfish to mice to humans (Kebschull et al., 2016a; Lake et al., 2016a; Tasic et al., 2016; Weissbourd et al., 2021; Yao et al., 2021; Zeng et al., 2012; Zhang et al., 2021). Yet, it is not obvious how to faithfully map such genetically identified subtypes to their circuit functions. This is in part methodological as there are ambiguous decision boundaries in clustering, but a more difficult problem, is what to do when there are conflicting sources of information. What does it mean when transcriptomic and projection identities clash? Lui et al. reported no one-to-one mapping from transcriptomic to projection neuron types in mouse prefrontal cortex, such that most projection defined cells included multiple transcriptomic types (Lui et al., 2021). In contrast, Murugan et al. identified largely non-overlapping, anatomically and molecularly distinct projection populations to some of the same targets in Lui et al (Murugan et al., 2017). One of the major differences between these studies is the primary definition of a cell type. That is, cell types were identified first by their transcriptomic identities in Lui et al, and their projection identities in Murugan et al. Though this might seem to be subtle distinct, structure is often more visible in one plane versus another, a basic concept that underlies dimensionality reduction. As such, finding the right plane that aligns multiple information modalities from genetics to projection patterns to electrophysiological properties is critical to identify functional cell types. From these distinct vantage points, Lui et al. found that most behaviorally relevant representations were mixed and distributed across cell types, whereas Murugan et al. found cell type specific tuning for spatial-social information. Such findings point to different hypotheses for how single neurons in prefrontal

cortex perform their computations but are only interpretable if the definition of cell type is reliable and not artifactual.

Temporally regulated gene expression might underlie conflicting genetic and projection identity (Klingler et al., 2021; Sahni et al., 2021a, 2021b). Transcription factors and cell-surface molecules that mediate cell-environment interactions are thought to shape neuronal fate (Frantz and McConnell, 1996; McConnell, 1988). However, such genetic heterogeneity peaks during circuit assembly and rapidly drops-off once mature connections are established (Li et al., 2017). Such reduction in genetic variation presents a challenge for using transcriptome clustering analyses to identify reliable clusters, particularly as most single cell sequencing experiments are in adults. Further, it is possible that there are simply no unique genes that shape projection identity. For instance, Klinger et al. (2021) recently demonstrated that some distinct projection populations cannot be identified as molecularly distinct during postnatal development. Rather, such subtypes arise from a single generic program that unfolds at different paces. Together, these findings argue that rather than subdivide neuron types by their transcriptomic differences, their outputs or projection patterns might lend insight into mature circuit functions.

Mesoscale projection patterns might be principled means to segregate functionally relevant neuron subtypes in adults (Harris et al., 2019; Muñoz-Castañeda et al., 2020; Oh et al., 2014; Zingg et al., 2014). Projection mapping experiments reveal reproducible, highly complex innervation patterns that provide insight into both function and developmental origins (Gergues et al., 2020; Kebschull et al., 2016a; Ren et al., 2019; Senn et al., 2014). For instance, within the basolateral amygdala, projections to the nucleus accumbens, medial central amygdala, and ventral hippocampus mediate distinct aspects of motivated behavior (Beyeler et al., 2018). These anatomical projection populations are topologically organized and form a valency map. Even basal forebrain cholinergic

neurons that historically, were thought function in a more diffuse manner, can be subdivided into types based on their projections to superficial and deep cortical laminae, a pattern that reflects their birth order (Allaway et al., 2020). Such organization seems to defy random connectivity, but the rules of target selection vary considerably by neuron type and region (Dahmen et al., 2022; Sosulski et al., 2011; Yuste, 2011).

At the extremes, projection neurons can either "broadcast" information to multiple targets or act as dedicated lines that preferentially route information to a single target (Han et al., 2018a). Though there are abundant broadcasting cells in primary motor and visual cortices, it is not obvious how this mode of information distribution shapes circuit function (Han et al., 2018a; Kita and Kita, 2012; Muñoz-Castañeda et al., 2020; Shepherd, 2013; Winnubst et al., 2019). That is, it is unclear if a subtype's circuit function is derived from their connections in toto or from a subset of branches; it remains unknown to what degree branching patterns are stochastic and a higher-order projection logic has yet to be established. Amongst broadcasting cells, there is some evidence to suggest that there might be a "primary" target that drives a cell type's function. In primary motor cortex, projection neurons to the thalamus and medulla were both spatially and molecularly distinct, and had specialized roles in motor control, despite their varied connections to cortex and striatum (Economo et al., 2018a; Muñoz-Castañeda et al., 2020; Yao et al., 2021; Zhang et al., 2021). Likewise, in primary visual cortex, projection neurons to higher order visual areas such as AL (anterolateral) and PM (posteromedial) were largely non-overlapping in both their local and long-range connections (Han et al., 2018a; Kim et al., 2018). Though these projection subtypes had similar stimulus response properties, they formed functional subcircuits that would allow for the independent transmission of sensory information. It is a challenge to determine the right degree of granularity from 'every neuron is unique' to broad projection classes (Migliore and Shepherd,

2005; Muñoz-Castañeda et al., 2020; Zeng and Sanes, 2017). Large-scale projection mapping might reveal an overarching logic that can be used to classify functional cell types.

1.3 Subcortical projection neuron types: the bottleneck layer between the new and the old

As in other cortical areas, neurons in orbitofrontal cortex are organized into distinct layers. Neurons projecting subcortically are enriched in cortical layer 5 (L5) and layer 6 (L6) (Angevine and Sidman, 1961; Caviness and Rakic, 1978; Frantz and McConnell, 1996; Gilbert and Kelly, 1975; McConnell and Kaznowski, 1991a). Layer 5 pyramidal neurons are thought to integrate inputs from multiple sources, including intra-columnar and long-range inputs, and constitute the sole output to subcortical areas , outside of the cortico-thalamic loop (Bourassa and Descheˆnes, 1995; Bourassa et al., 1995; Chen et al., 2005; Constantinople and Bruno, 2013; Deschênes et al., 1994; Economo et al., 2018; Harris and Mrsic-Flogel, 2013; Kim et al., 2015; Koester and O'Leary, 1993; Lai et al., 2008; Li et al., 2015; Lur et al., 2016; Mao et al., 2011). Projection neurons in layer 5 are typically categorized into two major classes: intratelencephalic (IT) neurons in upper sublayer 5a and extratelencephalic (ET) neurons in deep layer 5b. Projection neurons that target the midbrain and brainstem reside in layer 5b and projection neurons that target the striatum reside in layer 5a. While these broad projection classes are often treated as categorical cell types, their diverse projection targets defy this simplistic view (Bourassa and Descheˆnes, 1995; Chen et al., 2019; Deschênes et al., 1994; Ghosh et al., 2011; Han et al., 2018; Koester and O'Leary, 1993; Muñoz-Castañeda et al., 2020; Oh et al., 2014; O'Leary and Stanfield, 1985; Wilson, 1987; Winnubst et al., 2019). First, layer 5b neurons project to an array of highly

specialized subcortical brain structures and their anatomical and molecular diversity remains little explored (Gabbott et al., 2005; Lui et al., 2021; Murugan et al., 2017). Second, a lack of single-cell tracing studies leaves open the question whether subcortical projection neurons can be defined by their target and constitute distinct neuronal types.

Making up only ~5% of the projection neurons, L5b subcortically projecting neurons are morphologically distinct, representing the largest neurons in cortex with extensive dendritic arbors that span all layers of cortex (Brown and Hestrin, 2009; Kasper et al., 1994; Larkum et al., 2009; Mason and Larkman, 1990; Petreanu et al., 2009; Ramaswamy and Markram, 2015; Zhang et al., 2021). Their highly branched, thick-tufted apical dendrite that reaches to the cortical surface underlies characteristic electrophysiological properties, such as bursting (Chagnac-Amitai et al., 1990; Guan et al., 2015; Hattox and Nelson, 2007; Shai et al., 2015; Wang and McCormick, 1993). These neurons are not only connected across cortical layers, but also are heavily interconnected within L5 itself (Beul and Hilgetag, 2015; Brown and Hestrin, 2009; Dani and Nelson, 2009; Otsuka and Kawaguchi, 2011). Thus, positioned to integrate diverse inputs from all layers of cortex, these neurons act as an information bottleneck and mediate all cortical-subcortical communication, outside of the cortico-thalamic system (Arlotta et al., 2005; Lai et al., 2008; O'Leary and Koester, 1993). These subcortical projections must be precisely organized to route information relevant to the appropriate downstream regions. One possible mechanism is to directly constrain the flow of information through different projection patterns that serve as specialized information channels. Such a division-of-labor across distinct types of subcortical projection neurons has been previously proposed, but immense projection pattern diversity obscures the logic of target selection.

The advent of single cell technologies enabled the first large-scale interrogation of projection neurons (Muñoz-Castañeda et al., 2020; Yao et al., 2021; Zhang et al., 2021). Most studies on projection patterns and target selection are on a subtype of subcortically projecting neurons within motor cortex, pyramidal tract (PT) neurons. There is incredible diversity in their projection patterns, with neurons sending collaterals to structures in seemingly random combinations (Kita and Kita, 2012; Lévesque et al., 1996a). However, large-scale circuit mapping in mouse motor cortex looking at hundreds of single neuron reconstructions demonstrated that pyramidal tract neurons follow a one-to-many projection logic, distributing information to diverse, but non-random combinations of targets (Economo et al., 2018a; Muñoz-Castañeda et al., 2020; Winnubst et al., 2019). These data suggest the existence of brain-wide organizational principles that guide target selection, but the rules and the degree to which these are shared across cortex are unknown.

1.4 From form to function: a cell type specific approach to parsing complexity in prefrontal cortex

Neurons in frontal cortex correlate with a seemingly endless number of features - from sensory and spatial information to decisions and economic value to abstract concepts and rules (Bicks et al., 2015; Fleming and Dolan, 2012; Fuster, 2001; Miller, 2000). It is unknown, however, whether these diverse neural activity patterns might be organized into a set of conserved, core functions. Neurons are the basic computational units in the brain, representing vast inter-connected networks. Determining how complex behavior arises from such networks is a major goal in neuroscience and biological psychiatry, and one that has largely been driven by the concept of cell types (Gandal et al., 2016; Gordon, 2016; Steinberg et al., 2015). A cell type – a group of neurons that share similar anatomic, molecular, and physiological properties – represents not only a repeatable circuit component that constrains the flow of information across the brain, but also one that can be

targeted to probe internal computations and behavioral processes. Recent technical advances have enabled large-scale circuit mapping at a cellular resolution (Han et al., 2018a; Kebschull et al., 2016a; Muñoz-Castañeda et al., 2020; Winnubst et al., 2019), but the path from form to function is precarious. First, because the recent surge of anatomical studies has raised questions as to the definition of a cell type (Kepecs and Fishell, 2014; Miller et al., 2020a; Mukamel and Ngai, 2019). Second, because questions of form and function are largely addressed in disparate literatures at different resolutions, with the former focused more on broad developmental neuronal classes and the latter on anonymous neural recordings in complex behaviors. As such, the seemingly intimate link from a cell's identity to its functional circuit role has only just begun to be explored.

Across prefrontal cortex, neurons are incredibly diverse in their behavioral correlates, which is thought to reflect varied cognitive functions from long-term planning to self-reflection (Bicks et al., 2015; Fleming and Dolan, 2012; Fuster, 2001; Miller, 2000). This diversity, however, has made it difficult to understand how neurons carry out specific computations. For instance, neurons often respond to behavioral variables in different combinations (e.g., mixed selectivity) (Rigotti et al., 2013). Sophisticated, population-based approaches have been used to decode behaviorally relevant variables (Duncker and Sahani, 2021; Jazayeri and Ostojic, 2021; Vyas et al., 2020); recent studies suggest that such diverse neural responses could arise from neuron subtypes with distinct connectivity patterns, a possibility that is beginning to be understood in motor cortex (Economo et al., 2018a; Hirokawa et al., 2019; Lui et al., 2021; Spellman et al., 2021; Terra et al., 2020). As such, diverse representations might be organized across core conserved subnetworks and recruited to support varied behavioral processes. The contribution of distinct cell types, however, has not been widely explored in frontal cortex. This is because first, most neurons are recorded agnostic

to their anatomic and molecular features and second, there is no census of projection neuron types in prefrontal cortex.

Part of prefrontal cortex, orbitofrontal cortex contains heterogeneous neurobiological cell types that may partly account for its striking functional response diversity (Gabbott et al., 2005a). Within orbitofrontal cortex, cell type specific recordings in rats reveal an intimate relationship between a cell's identity and its functional role (Groman et al., 2019; Hirokawa et al., 2019; Pascoli et al., 2018). In a recent study, Hirokawa et al found that single neuron responses in orbitofrontal cortex were highly structured and fell into distinct clusters, with each encoding decision variables, such as reward magnitude and previous trial outcome. Recording from an anatomically defined projection population, these authors further found that this population was highly uniform and encoded past trial outcome, with markedly sustained activity from the receipt of reward to the onset of the next trial. Recordings from other subcortically projecting populations demonstrate that OFC outputs encode behaviorally relevant information and mediate dissociable functions for learning and decision-making (Bariselli et al., 2020; Groman et al., 2019; Hirokawa et al., 2019; Izquierdo, 2017; Lui et al., 2021; Malvaez et al., 2019; Namboodiri et al., 2019; Pascoli et al., 2018). Further, such target defined subcortically projecting neurons can be mapped to separable reinforcement learning computations and causally linked to maladaptive behaviors that maintain psychiatric diseases like drug addiction (Groman et al., 2019). This hints at an underlying logic wherein neural representations are sorted to distinct subcortical structures, a form of information distribution that requires distinct output tracts and ordered subcortical target selection. These findings are suggestive of a circuit module that can be recruited across different tasks. This functional division-of-labor, however, needs anatomically distinct output channels as a structural prerequisite.

1.5 Running the table from cell types to reinforcement learning models

From the premise that the brain has evolved to solve computational problems, reinforcement learning models have provided mechanistic insight into healthy and aberrant choice behaviors (Daw et al., 2006; Dorris and Glimcher, 2004; Groman et al., 2019; Keramati and Gutkin, 2014; Lak et al., 2020; Rutledge et al., 2009). Reinforcement learning models provide a proven, principled framework both for explaining behavior and relating it to neural activity (Sutton and Barto, 2018); paired with high-dimensional, well-controlled behaviors, such an approach enables fractionating choice behavior into their underlying quantifiable processes. For instance, in both mice and humans, hallucination-like perception was established as a quantitative behavior in a two-alternative forced choice task that included single trial time investments (Schmack et al., 2021). Within a reinforcement learning framework, these hallucination-like percepts were cast as high-confidence false alarms and this phenotype could be described via a limited set of parameters (i.e., decision variables) across individuals and species. Importantly, such quantifiable processes and parameters can be mapped to neural representations, lending insight as to how such computations are instantiated in neural circuits.

Hirokawa et al. identified highly structured representations in orbitofrontal cortex that encoded reinforcement learning components such as decision confidence and reward magnitude (Hirokawa et al., 2019). Such a finding suggests that the right analytical lens can organize diverse frontal cortical representations. Armed with a foundational understanding of anatomical cell types, an approach that combines high-dimensional behaviors and computational modelling enables mapping decision variables to their underlying neural substrate. This is critical for future experiments to determine if such representations are stable across a variety of task contexts and

for bi-directional causal manipulations to both induce and rescue maladaptive behaviors. To test the hypothesis that OFC projection types encode distinct decision variables, we developed a novel behavioral task that includes perceptual and outcome uncertainty, with a focus on ventral tegmental area projection types (Chapter 4). The ventral tegmental area projection type was particularly intriguing because OFC input (Takahashi et al., 2009a, 2011) is thought to be required for a critical function of ventral tegmental area dopamine neurons, namely reward prediction error (Schultz et al., 1997). However, precisely what messages the OFC sends to the ventral tegmental area, and how this impacts choice behavior is unknown.

1.6 Ventral striatal dopamine release and its relationship to value

Dopamine is famously related to value (Berridge and Robinson, 1998; Hamid et al., 2016; Kim et al., 2020; Mohebi et al., 2019; Schultz et al., 1997; Wise and Rompre, 1989), but precisely what information is encoded and its relationship to choice remains widely debated. Advances in high resolution and high throughput techniques from genetically encoded dopamine sensors to voltammetry allowed for measuring dopamine release across multiple timescales in behaving animals (Hamid et al., 2016; Patriarchi et al., 2018, 2020; Robinson et al., 2003). Classically, dopamine is thought to reflect reward prediction error, an important signal in temporal difference learning (Schultz et al., 1997; Sutton and Barto, 2018). It is often challenging, however, to disentangle value from reward prediction error signals. In other words, motivational signals that "look forward" versus prediction error that "looks backwards" (Berke, 2018). Specifically, observations of ramping dopamine signals in the ventral striatum have been proposed to reflect either value related signals related to motivation or prediction error signals (Hamid et al., 2016;

Howe et al., 2013; Kim et al., 2020; Mikhael et al., 2022). Distinguishing these possibilities and the role of dopamine in different types of decisions is further complicated by regional differences in the striatum and heterogeneous dopaminergic cell types (Beier et al., 2015a, 2019; Clatworthy et al., 2009; Di Ciano et al., 2008; Gokce et al., 2016; Hamid et al., 2016; Lynd-Balta and Haber, 1994; Märtin et al., 2019).

There is substantial evidence that dopamine release in the ventral striatum reflects expected value for an upcoming reward and invigorates approach behavior (Bromberg-Martin et al., 2010; Schultz et al., 2015; Wise, 2004). Such expected value signals are modulated by subjective factors such as effort, delay to reward and even conspecific distress (Hollon et al., 2014; Lak et al., 2014a; Lichtenberg et al., 2018; Saddoris et al., 2015; Varazzani et al., 2015). Such findings make it a strong neural candidate for subjective value coding.

Subjective value is derived from the inherent and fluctuating preferences of the chooser. The determination of subjective value is commonly assessed via a series of choices across inherently different reward types, such as apples and oranges. As most reward types are not directly comparable and there is no "right" answer, it is thought that such economic decisions are made by comparing a "common currency" value. This "common currency" allows one to rank different rewards relative to a common reference point. However, precisely how neurons might integrate and transform objective information about a reward into subjective value remains unclear. A better understanding of the neural representation of subjective might disentangle these processes.

Classically, subjective value is inferred through a series of choices, such as progressive ratio break points or psychophysical indifference points (Gardner et al., 2018, 2017; Padoa-Schioppa and Assad, 2006a). This is often framed as an entirely separate process from the value described in

learning theory (which in some cases, is identical to prediction error) (Padoa-Schioppa and Schoenbaum, 2015). In learning theory, single trial estimates can be inferred from reinforcement learning models that quantify how past information is integrated and updates current value estimates (Delgado et al., 2011; Sutton and Barto, 2018). Such value estimates are then pit against one another and mapped to a choice. In Chapter 5, we propose a novel framework to bring together these distinct measures of subjective value, revealed choice preference and model inferred value.

1.7 Purpose and organization of this thesis

The purpose of this thesis is to bridge anatomy, neural circuits, and behavioral modelling to drive a biologically grounded and mechanistic understanding of decision making. The overarching goal is to fractionate choice behavior into discrete processes and decision variables that can be mapped to distinct cell types. The findings in this thesis represent the first blueprint of subcortically projecting neurons in orbitofrontal cortex and demonstrate that these neurons are distinct cell types, as described in Chapter 2. The developmental implications of such an organization are outlined in Chapter 3. Importantly, an output architecture in which subcortically projecting neurons can be subdivided into target defined subtypes, is suggestive of segregated functional circuits that are positioned to selectively route information to downstream structures. Such a subcortical sorting principle has long been hypothesized but has yet to be demonstrated in any forebrain region.

From this anatomical framework, I next developed a set of high-dimensional behaviors that can be used to test the generalizability of decision variables in future. These include a classic matching task with trial-by-trial time investment that captures fluctuations in subjective value (Chapter 6) and a variance task (Chapter 4) that can de-couple sensory and value uncertainty, and test for

dynamic representations of learning rate. Such choice paradigms combined with reinforcement learning models can fractionate choice behaviors into their underlying quantifiable processes, generating a vocabulary of decision variables that can be used to quantify individual behavioral variability and in turn, track parameters over time and across healthy and disease states. Such an approach can generate a "computational fingerprint" which can be used as a proof-of-principle for novel behavioral diagnostics and cell type driven therapeutics.

Chapter 2

Target defined subcortical projection neuron types of rat orbitofrontal cortex

Suelynn Ren[1*], Torben Ott[1,2*], Ben-Orli Nathanson[3], Rodrigo Munoz-Castaneda[4], Steven Ryu[1], Shujing Li[4], Melody Wu[4], Thomas Klausberger[3], Junya Hirokawa[5], Jessica Tollkuhn[4], Longwen Huang[4,7], Alexander R. Nectow[6], Pavel Osten[4], Adam Kepecs[1]

1: Department of Neuroscience and Department of Psychiatry, Washington University in St. Louis, St. Louis, MO 63110, USA

2: Current address: Bernstein Center for Computational Neuroscience Berlin, Humboldt University of Berlin, Berlin, Germany

3: Center for Brain Research, Dept. Cognitive Neurobiology, Medical University of Vienna, 1090 Vienna, Austria

4 Cold Spring Harbor Laboratory, Cold Spring Harbor, NY11724, USA

5: Doshisha University, Kyotanabe, Kyoto 610-0394, Japan

6: Columbia University, New York City, New York, NY10032, USA

7: Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

*equal contribution

Abstract

Frontal cortex supports sophisticated behaviors by controlling evolutionarily more ancient brain regions through subcortically projecting neurons. There are diverse subcortically projecting pyramidal neurons, but the organizational principles of their projection logic remains unknown. Here, we combined viral tracing, single-cell resolution projection mapping and RNA sequencing to comprehensively map the subcortical outputs of rat orbitofrontal cortex to establish a blueprint for the organization of its long-range projecting neuron types. At a single neuron resolution, we characterize the connectivity of ~450,000 neurons and demonstrate that most neurons project to single subcortical targets, representing dedicated lines of communication. Neurons projecting to multiple target structures are rare, but represent non-random projection motifs. These single-target projection neurons are organized into intermediate sublayers beyond the classical layer 5a and 5b distinction, and express unique sets of genes, including transcription factors and surface receptors. These refined sublayers are ordered according to the distance of their projection target in a "deeper further" pattern, which may reflect the inside-out developmental sequence of cortex. Taken together, subcortically projecting neurons in orbitofrontal cortex represent distinct neuron types that are positioned as segregated information channels, an architecture that hints at differential information routing from frontal cortex.

Highlights

- Extra-telencephalic projection neurons in rat orbitofrontal cortex constitute target-defined neuron types
- Brainwide single-neuron projection maps reveal mostly one-neuron-one-target connectivity

- Laminar position in refined sublayers of layer 5b correlates with distance to projection target in a “deeper further” pattern
- Target-defined subcortical projection populations represent molecularly distinct subtypes

Introduction

Neocortex must communicate with evolutionarily older subcortical structures to provide the additional layer of control that enables greater behavioral repertoire and flexibility in mammals. Communication from the neocortex to subcortex takes place through large pyramidal cells located in the deep layers of cortex. Positioned at the interface between the new and old brain, these long-range projection neurons make up cortical layer 5 (L5) and comprise two broad classes of pyramidal cells: intra-telencephalic (IT) neurons that project within the telencephalon, located mostly in superficial layer 5a, and extra-telencephalic (ET) neurons in deep layer 5b, which project to many subcortical structures including the superior colliculus and various brainstem nuclei (Bates and Killackey, 1984; Killackey et al., 1989; McGeorge and Faull, 1989; Morishima et al., 2011; Muñoz-Castañeda et al., 2020; Oh et al., 2014; O’Leary and Koester, 1993; Winnubst et al., 2019; Wise and Jones, 1977). Long-range projection neurons are specialized to integrate and transform diverse cortical input sources. For instance, ET neurons have large dendritic arbors that span across all cortical layers (Constantinople and Bruno, 2013; Fletcher and Williams, 2019; Larkum et al., 2009; Markram et al., 1997; Petreanu et al., 2009; Zarrinpar and Callaway, 2016). These long-range projection neurons are well positioned to coordinate the flow of information to subcortex. The organizational principles, or projection patterns, that govern these cortical outputs are, however, unclear. For instance, it is unknown whether individual projection neurons target single subcortical structures, or target a random mixture of downstream areas.

Our understanding of subcortical projection patterns has been mostly derived from pyramidal tract (PT) neurons, an ET subtype within motor cortex that typically sends collateral branches to multiple structures (Akintunde and Buxton, 1992; Hirai et al., 2012; Kita and Kita, 2012; Muñoz-Castañeda et al., 2020; Rojas-Piloni et al., 2017; Winnubst et al., 2019). The projection patterns of ET neurons might be a principled means by which to segregate neuronal subtypes, a lens which brings anatomy, genetics, and circuit function into focus on a single plane. For instance, some PT subtypes selectively target discrete spinal segments and can be anatomically differentiated, even at the earliest stages of axon extension (Sahni et al., 2021a, 2021b). Such a foundation is critical for identifying hodological molecular controls, beyond the specification of broad projection classes (Arlotta et al., 2005; Chen et al., 2005; Lai et al., 2008; Molyneaux et al., 2007; Paolino et al., 2018; Weimann et al., 1999; Woodworth et al., 2016; Yao et al., 2021). It remains an open question how subcortical target selection and genetic identity are related, and how such attributes shape mature circuit functions. Mapping the subcortical projectome might uncover the rules of target selection that could emerge from distinct neuronal subtypes.

Part of prefrontal cortex, the orbitofrontal cortex (OFC) is thought to underlie varied cognitive functions including economic decision-making, valuation and learning (Gallagher et al., 1999; Gottfried, 2003; Gremel and Costa, 2013; Lak et al., 2014b; Miller et al., 2020b; O'Doherty et al., 2001; Padoa-Schioppa and Assad, 2006a; Tremblay and Schultz, 1999). These complex behavioral functions are reflected in the striking diversity of neural activity that is observed in OFC. Recent studies suggest that diverse neural responses in prefrontal cortex could arise from neuron subtypes with distinct projection targets (Groman et al., 2019; Hirokawa et al., 2019; Lui et al., 2021; Spellman et al., 2021; Terra et al., 2020). Such connectivity defined subtypes contribute to different valuation processes and can be causally linked to maladaptive choice behaviors and

psychological phenotypes, such as blunted, adaptive choice biases and compulsivity (Groman et al., 2019; Harada et al., 2021; Hirokawa et al., 2019; Malvaez et al., 2019; Pascoli et al., 2018). However, due to a lack of single cell tracing studies in prefronal cortex, it is unknown whether projection neurons can be defined by their target(s) and constitute distinct neuronal subtypes.

Here, we set out to understand the architectural logic of subcortical projection cell types in the rat OFC. We used a combination of viral and classic tracing, single-cell resolution projection mapping, and RNA sequencing to comprehensively map subcortical outputs from OFC (Kebschull et al., 2016a; Li et al., 2018; Nectow et al., 2017; Soudais et al., 2001). We systematically categorize projection patterns in OFC at a single neuron resolution and find that most subcortical projection neurons target a single region, with rare multi-target projection motifs. These target-defined projection populations are organized into different but overlapping sublayers, suggesting distinct developmental origins. We also demonstrate that these anatomically distinct projection populations are molecularly distinct, with different patterns of transcription factors and membrane proteins. Together, these multi-modal data reveal that projection-defined populations represent distinct cell types. This anatomical framework in OFC is well suited to support the selective distribution of information to different downstream structures. Such selective routing of information to specific subcortical targets could enable frontal cortex to exert precise control over primordial processes served by subcortex.

Results

Broad and patchy brain-wide projections from orbitofrontal cortex

We mapped major OFC outputs across the entire rat brain to identify key subcortical target structures (Figure 1). We studied the ventro-lateral, lateral, and dorsolateral subregions of OFC based on previous physiology and inactivation studies implicating this area in valuation, learning and confidence (Constantinople et al., 2019a; Gremel and Costa, 2013, 2013; Groman et al., 2019; Hirokawa et al., 2019; Lak et al., 2014; Masset et al., 2020; Miller et al., 2020;; Takahashi et al., 2009). Here, we primarily focused on the segment that lies between approximately +4.5 mm and +3.5 mm anterior of bregma in rats and refer to all these subregions together as OFC (Figure 1A). To identify subcortical projection targets, we performed both classic histology and image analysis in adult rats bilaterally injected with an anterograde viral tracer (AAV1-CAG-tdTomato).

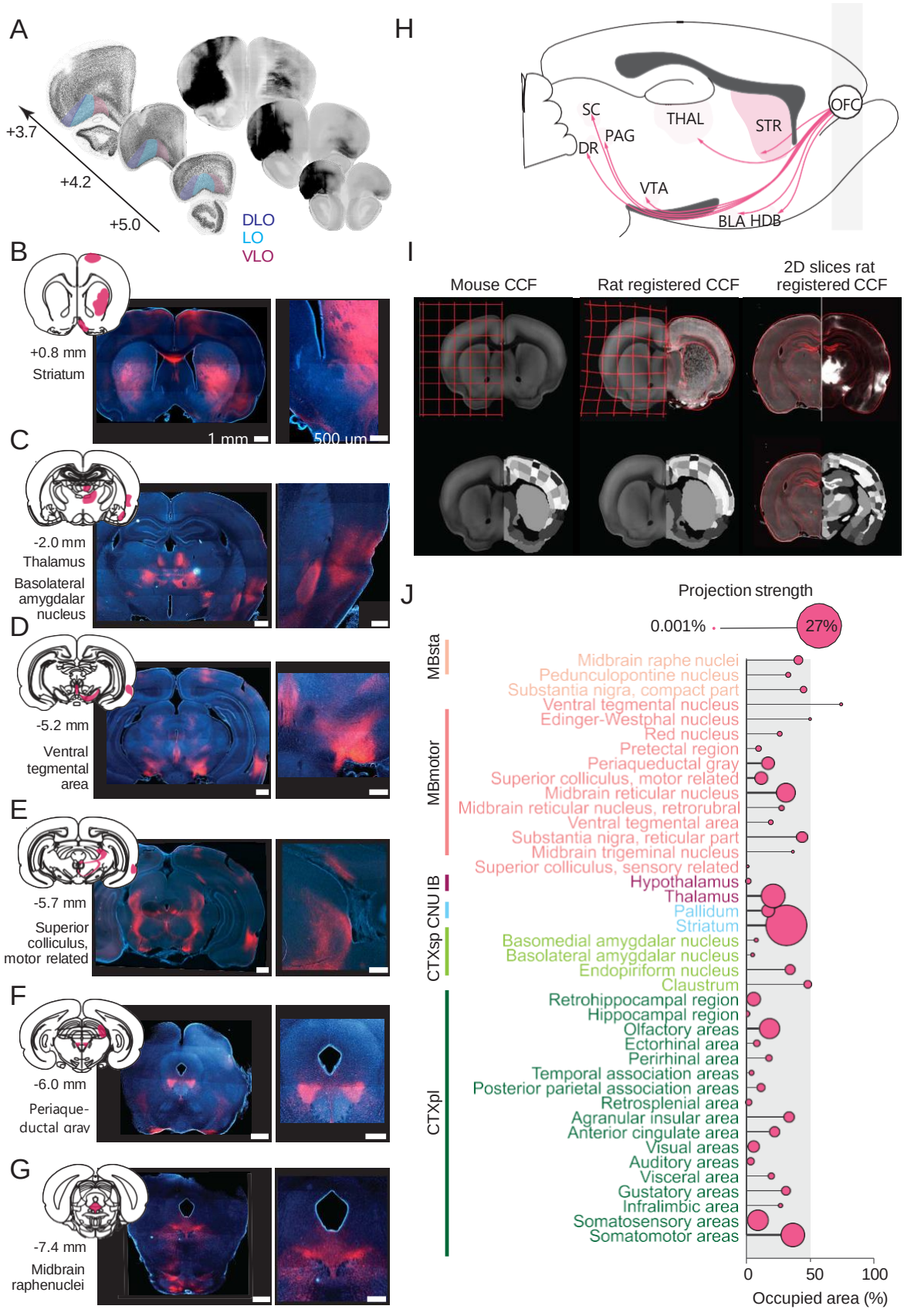

A
+3.7
+4.2
+5.0
DLO
LO
VLO
B
+0.8 mm
Striatum
1 mm
500 um
C
-2.0 mm
Thalamus
Basolateral amygdalar nucleus
D
-5.2 mm
Ventral tegmental area
E
-5.7 mm
Superior colliculus, motor related
F
-6.0 mm
Periaque-ductal gray
G
-7.4 mm
Midbrain raphenuclei
H
SC
PAG
DR
THAL
STR
OFC
VTA
BLA HDB
I
Mouse CCF
Rat registered CCF
2D slices rat registered CCF
J
Projection strength
0.001%
27%
MBsta
MBmotor
CTXsp CNU IB
CTXpl
Midbrain raphe nuclei
Pedunculopontine nucleus
Substantia nigra, compact part
Ventral tegmental nucleus
Edinger-Westphal nucleus
Red nucleus
Pretectal region
Periaqueductal gray
Superior colliculus, motor related
Midbrain reticular nucleus
Midbrain reticular nucleus, retrorubral
Ventral tegmental area
Substantia nigra, reticular part
Midbrain trigeminal nucleus
Superior colliculus, sensory related
Hypothalamus
Thalamus
Pallidum
Striatum
Basomedial amygdalar nucleus
Basolateral amygdalar nucleus
Endopiriform nucleus
Claustrum
Retrohippocampal region
Hippocampal region
Olfactory areas
Ectorhinal area
Perirhinal area
Temporal association areas
Posterior parietal association areas
Retrosplenial area
Agranular insular area
Anterior cingulate area
Visual areas
Auditory areas
Visceral area
Gustatory areas
Infralimbic area
Somatosensory areas
Somatomotor areas
0
50
100
Occupied area (%)

Figure 1 OFC projections to subcortex are broad and patchy.

(A) NeuN stain of representative OFC slices and shaded dorso-lateral, lateral, and ventro-lateral subregions (left). Injection sites for anterograde tracer, AAV1-CAG-TdTomato (right). **(B-G)** Coronal sections of OFC input to different subcortical targets: striatum (B), thalamus and basolateral amygdala (C), ventral tegmental area (D), superior colliculus (E), periaqueductal gray (F), and dorsal raphe (G). OFC input is largely restricted to specific subregions such as the ventro-lateral periaqueductal gray and lateral superior colliculus. **(H)** Overview of the major projection targets from OFC. **(I)** Workflow for manual registration of selected slices onto the Common Coordinate Framework in rats. **(J)** Quantitative analysis of projection strength and the % area innervated across all registered structures. Projection strength was represented by the size of the circular marker, ranging from 27% for the most abundant projection to the striatum to 0.001% for the rarest projections to the deep, brainstem nuclei. % Area occupied is represented along the x-axis and demonstrated that most regions receive restricted input from OFC

OFC projects to diverse subcortical structures across the brain. From wide-field imaging of individual brain slices, we manually selected a subset that had strong fluorescence. These brain slices were manually registered to a rat Common Coordinate Framework (Figure 1). We first quantified the strength of OFC input to its downstream projection targets from pixel counts (Figure 1C). OFC input was found strongly in the striatum and thalamus, but also in the amygdala, ventral tegmental area, periaqueductal gray and the hitherto unidentified projection to the deep lateral subregion of the superior colliculus that includes the nucleus of optic tract.

Projection pathway targets were highly localized to distinct subregions with high spatial specificity. We quantified the area occupied by OFC efferent fibers, ranging from 0 to 100%, with 100% indexing a region that receives either diffuse innervation across the entire structure or very small nuclei (Figure 1C). We found that OFC output tends to occupy less than 40% of the area, reflecting a more restricted output to subcortical structures. Visual inspection demonstrated that this restricted input is, indeed, localized to specific subregions of subcortical structures (Figure 1D-I). For instance, OFC output to the amygdala is restricted to a ~300 um anterior-posterior segment, output to the periaqueductal gray is confined to a ventro-lateral subregion, and projections are restricted to the ventro-lateral part of superior colliculus. Retrograde tracing

confirmed that OFC indeed projected to these areas and that the observed output patterns were not caused by spillover of anterograde viral tracer to neighboring frontal areas (see below Figure 5). Taken together, these widespread, but highly targeted projections suggest that the OFC sends specific input to diverse subcortical structures, representing a densely connected hub.

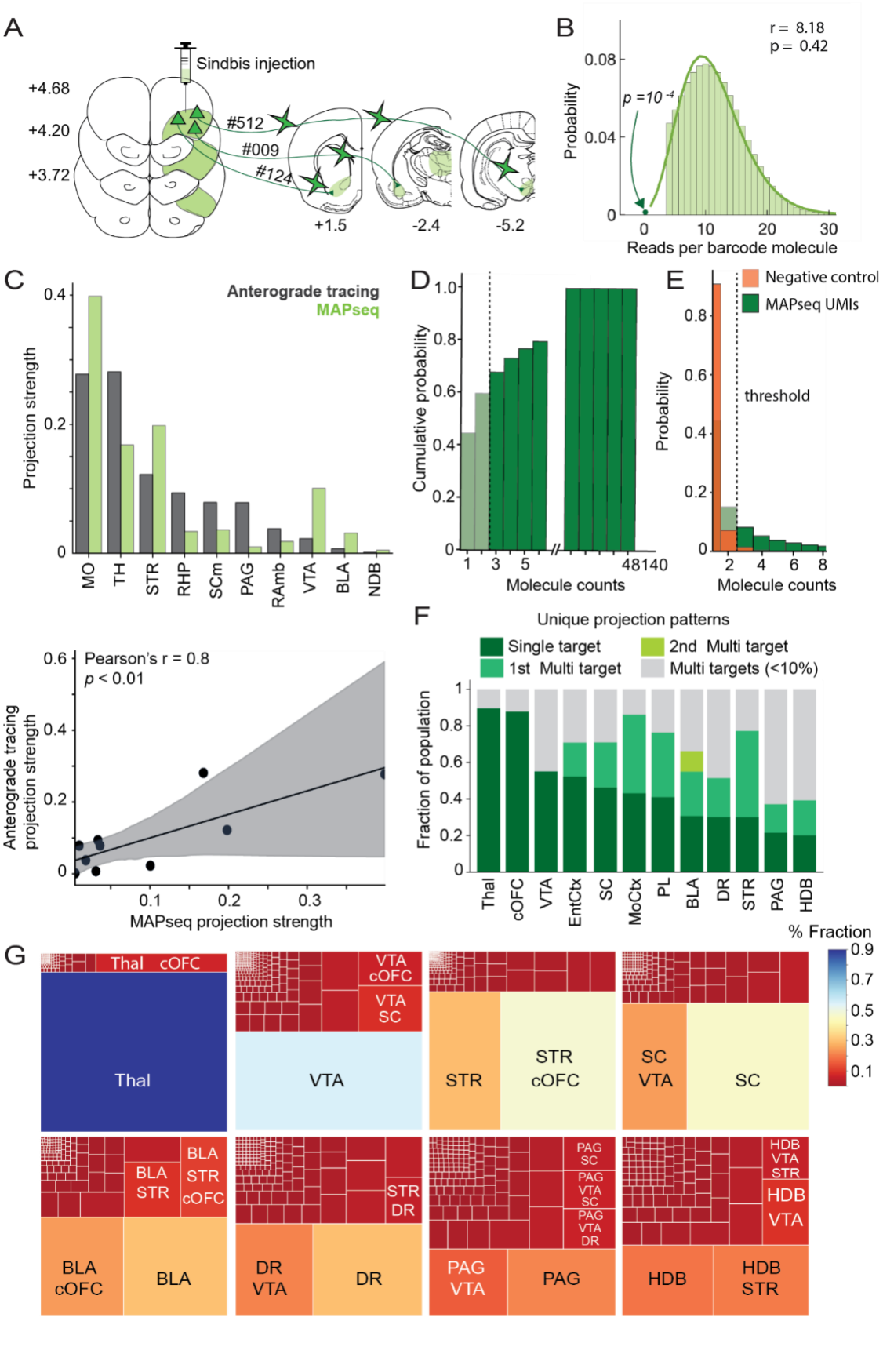
Single-neuron resolution mapping reveals projection defined cell types and non-random motifs

We next sought to determine whether individual projection neurons in OFC selectively target a single subcortical region or if they broadcast broadly to multiple subcortical regions. Based on the anterograde mapping, we selected OFC's major subcortical targets as well as additional cortical and thalamic target areas for simultaneous single neuron tracing. We adapted a high-throughput, single neuron tracing technology - multiplexed analysis of projections (MAPseq) for use in rat OFC (Figure 2A) (Kebschull et al., 2016a). We uniquely tagged neurons in OFC with a random RNA sequence (30-nucleotide barcodes) linked to a GFP fluorophore by injecting a library of viral vectors (barcoded Sindbis virus) across OFC. Barcode RNAs were then robustly expressed and actively transported to pre-synaptic terminals. We found stable expression of barcode mRNA in distal projection targets across a range of barcode trafficking times (36 to 68 hours), ensuring that barcode detection was unaffected by axonal length (see Methods, Figure 10). We focused on 12 target areas that included a handful of cortical as well as the major subcortical projection targets. From anatomic landmarks, we manually identified and hand dissected the OFC and our 12 target areas from frozen coronal sections (see Figure 11 for dissection maps), isolated and purified barcode RNAs and sequenced these pooled samples. Notably, because OFC output is targeted and areas were non-adjacent, we determined areal boundaries more liberally in order to maximize the number of captured barcodes. We identified unique barcodes across all dissected target areas to

generate a 'barcode matrix' of single neuron projection patterns, in which the elements represent whether a barcode was present in each area. Thus, a barcode that is only detected in one target area represents one neuron that specializes in communicating with one region, whereas a barcode that is detected in multiple target areas represents a neuron that broadcasts information to multiple regions.

Figure 2 Most OFC projection neurons target a single area.

(A) MAPseq experiment. Sindbis barcoding virus was injected into OFC and to label neurons with unique barcodes. Barcodes are then actively transported down to the presynaptic terminals, target areas dissected and sequenced. Green overlays illustrate the dissected injection and downstream regions. **(B)** The distribution of sequencing reads per barcode in target regions and negative binomial fit. **(C)** For matched projection targets regions, projection strength estimated from whole-brain anterograde tracing and pixel quantification versus MAPseq were significantly correlated (Pearson's r = 0.8, $p<0.01$). **(D)** The distribution of barcode molecule counts in target regions (green) and in negative control samples (red). To determine whether a neuron projects to a given target area, we chose a low threshold (dotted line) of at least 3 barcode molecules, filtering barcodes due to noise (what's the left panel). **(E)** The fraction of single target projection patterns compared to the next most common multi-target projection pattern(s) above 10% for each target. Projection patterns below 10% are pooled and visualized in gray. **(F)** The fraction of single-target and multi-target projection patterns above 5% for each target region. Projection patterns below 5% are not labelled for visualization. *Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA, basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band.*


A
Sindbis injection
+4.68
+4.20
+3.72
#512
#009
#124
+1.5
-2.4
-5.2
B
r = 8.18
p = 0.42
p =10 -4
Probability
0.08
0.04
0
0
10
20
30
Reads per barcode molecule
C
Anterograde tracing
MAPseq
Projection strength
0.4
0.2
0
MO
TH
STR
RHP
SCm
PAG
RAmb
VTA
BLA
NDB
D
Cumulative probability
1.0
0.8
0.6
0.4
0.2
0
1
3
5
48140
Molecule counts
E
Negative control
MAPseq UMIs
Probability
0.8
0.6
0.4
0.2
0
threshold
2
4
6
8
Molecule counts
Pearson's r = 0.8
p < 0.01
Anterograde tracing
projection strength
0.6
0.4
0.2
0
0.1
0.2
0.3
MAPseq projection strength
F
Unique projection patterns
Single target
1st Multi target
2nd Multi target
Multi targets (<10%)
Fraction of population
1
0.8
0.6
0.4
0.2
0
Thal
cOFC
VTA
EntCtx
SC
MoCtx
PL
BLA
DR
STR
PAG
HDB
G
% Fraction
0.9
0.7
0.5
0.3
0.1
Thal
Thal cOFC
VTA
VTA cOFC
VTA SC
STR
STR cOFC
SC
SC VTA
BLA
BLA STR
BLA STR cOFC
BLA cOFC
DR
STR DR
DR VTA
PAG
PAG SC
PAG VTA SC
PAG VTA DR
PAG VTA
HDB
HDB VTA STR
HDB VTA
HDB STR

MAPseq is subject to different sources of error that lead to either mistaking noise for true barcodes (i.e., false positives) or missing true barcodes (i.e., false negatives), as also seen in other conventional tracing techniques. The major sources of error include inefficient expression and trafficking of barcode mRNAs, low sequencing depth, and sequencing noise (see Methods for a full discussion on these and other sources of error). We determined that barcode expression and the number of reads per barcode molecule were sufficient to rarely miss true barcodes (Figure 2B, Figure 12), nevertheless errors introduced from sequencing are unavoidable. Within the barcode matrix, the uncertainty of a detected barcode is inferred from its molecule counts, that is, the number of barcode molecules detected in the area. Thus, we used a barcode molecule count threshold to filter out barcodes that could stem from noise. We obtained a null distribution of barcode molecule counts by sequencing negative control areas with no OFC projections from the same brains (e.g., contralateral olfactory bulb), which showed generally low molecule counts (below 22, peak at 1, Figure 2C). Because we first asked whether projection neurons have dedicated projection targets, we chose a low threshold for accepting a barcode (> 2 barcode molecules) to minimize missing barcodes (i.e., low false negatives). A higher barcode count threshold might bias the analysis toward identifying more single target projection neurons. This choice of threshold increased the number of false multi-target projection patterns (i.e., false positives), therefore it provides a lower bound for determining the fraction of single target projection neurons. From this analysis, we identified the projection patterns of ~450,000 OFC neurons (i.e., unique barcodes) to 12 target regions across three individual rats. Note that we observed generally similar results for higher thresholds (e.g., >10 barcode molecules, Suppl. 3) that minimized the chance of mistaking noise for a true barcode (i.e., low false alarms), but could

miss more barcodes (i.e., false negatives) since a larger fraction of the data falls below the threshold (Figure 2C).

We next compared these MAPseq data to anatomical estimates of projection strength based on whole-brain anterograde tracing (Figure 1). We visually matched the dissected regions in MAPseq experiments to their counterparts in the rat Common Coordinate Framework, and selected subregions such as the ventral region of striatum and motor related region of superior colliculus. We performed a linear regression on projection strengths determined from hand-dissected regions in MAPseq and matched regions from whole-brain anterograde tracing. Despite the vast methodological differences, we found that the results were largely in line (Figure 2C, Pearson's $r = 0.8$, $p < 0.01$). The broad correspondence of these two methodologically different measures of OFC outputs provides confidence in the quantitative map of projections strengths that we analyze in detail.

We found that neurons with a single target were the most common pattern (~60% of subcortical; ~90% of cortical) of output connectivity from OFC. We observed that single-target projections to the contralateral OFC were the most abundant, making up ~74% of all projection patterns. There were also substantial populations of single-target neurons to cortical regions including in the prelimbic area (~41%), entorhinal cortex (~52%), and motor cortex (~43%). Regarding subcortical targets, as expected, single-target projection neurons make up ~90% of input to the thalamus (Bourassa and Deschênes, 1995; Bourassa et al., 1995; Gabbott et al., 2005a; Lévesque et al., 1996b). Single-target projection neurons also make up ~50% of input to the ventral tegmental area and superior colliculus (Figure 2D). In contrast, projection neurons to the basolateral amygdala,

horizontal diagonal band, striatum, and dorsal raphe were largely dominated by two distinct patterns of connectivity: a single-target subpopulation as well as a second multi-target subpopulation (Figure 2D, E). Since we only considered 12 major projection targets, it remains possible that some of these neurons had undetected axonal collaterals to other regions. Notably, we used a conservative, low barcode threshold for this analysis that underestimated the contribution of single target subpopulations (see Figure 12B,C for the fraction of single target neurons at different molecule count thresholds). Choosing a higher barcode threshold (e.g., >10 barcode molecules, see Han et al., 2018 and Kebschull et al., 2016) is expected to increase the fraction of single target neurons and decrease the fraction of rare, multi-target neurons. Indeed, using a higher barcode threshold that minimized sequencing noise increased the number of single-target projections such that it became the dominant projection pattern across all regions (~70% ventral tegmental area; ~70% superior colliculus; ~50% periaqueductal gray; ~50% dorsal raphe; ~50% basolateral amygdala, ~40% horizontal diagonal band), apart from the striatum (Figure 12C).We next asked whether the remaining non-dedicated projection neurons that target multiple areas showed non-random connectivity patterns.

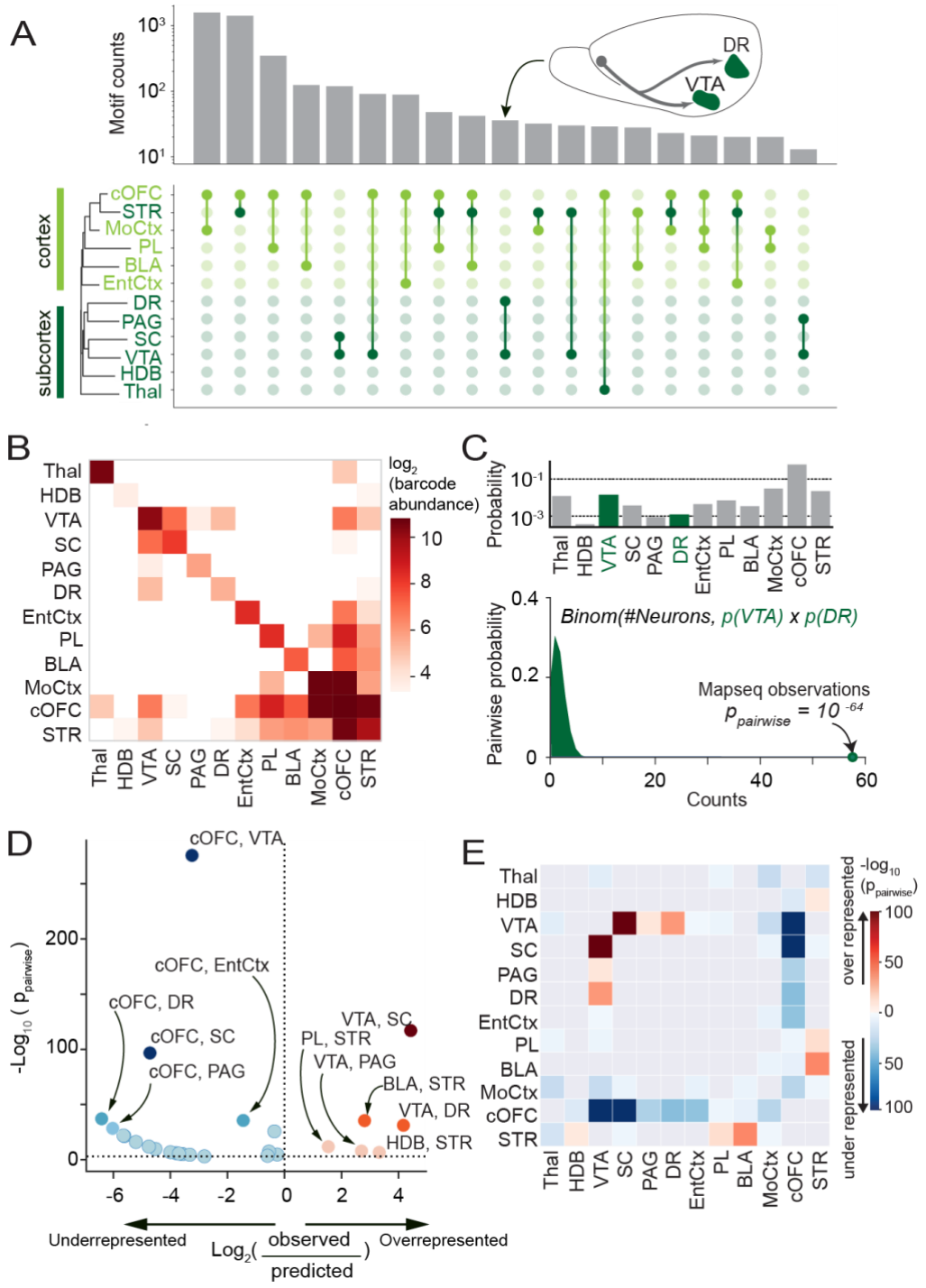


Figure 3 Multi-target projection patterns are non-random "motifs".

(A) Multi-target projection patterns in descending order by their occurrence, filtered to only include projection patterns more than 10 counts (top). Dendrogram of unsupervised hierarchical clustering of OFC projection neurons and the regions that make up the most common multi-projection patterns (bottom). Cortical regions are colored light green and subcortical regions are colored dark green. **(B)** Abundance matrix that holds the number of co-occuring barcodes per pairwise combination of regions, with the number of single-target barcodes along

the diagonal. **(C)** Probability of selecting a barcode belonging to a target region (top). Example to calculate the likelihood of observing a bifurcation motif under the null distribution (binomial probability distribution, purple) assuming independence(bottom). **(D)** Likelihood of bifurcation motifs ($p_{pairwise}$ as in E) versus fold change of observed over the predicted number of co-occuring barcodes from the null model. Horizontal dotted line indicates significance level ($P_{adj} < 0.01$, Bonferroni corrected). Underrepresented motifs are to the left of the vertical dotted line, whereas overrepresented motifs are to the right. **(E)** Likelihood of observing bifurcation motifs for all area combinations. Bifurcation motifs that are overrepresented (observed more often than predicted by the null model) are shown in red, underrepresented motifs in blue. *Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA, basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band.*

Subcortical multi-target projection patterns are non-random

Although OFC neurons projecting to multiple subcortical regions were rare, we identified several projection motifs, that is, repeating patterns of neurons targeting multiple subcortical regions. As the identification of multi-target projection patterns is sensitive to false positives, mistaking noise for true barcodes, we chose a stricter threshold based on the maximum barcode molecule count in the negative controls (>10 barcode molecule count, see Methods). We then sorted these high-confidence, multi-target projection patterns based on their frequency of occurrence, with a focus on projection patterns that were observed at least 10 times (on average, ~3 times per brain) (Figure 3A). We found that most multi-target projection neurons target cortical, rather than subcortical structures, apart from the striatum (Figure 3A). In contrast, subcortical ‘n-furcations’ (neurons projecting to n target areas, n > 1) were less abundant by an order of magnitude. The most common bifurcations (n=2) included the ventral tegmental area and superior colliculus (~7% of projections to the ventral tegmental area; 103 counts), and the ventral tegmental area and dorsal raphe (~2% of projections to the ventral tegmental area; 33 counts). This is further reflected in an abundance matrix, in which each element is the number of above-threshold (>10 barcode molecule count), co-occurring barcodes for a region pair, with single target (i.e., dedicated) barcodes represented along the diagonal (Figure 3B). Given their rarity in these MAPseq data, we independently

validated multi-target subcortical projections using a dual viral strategy, with a retrograde virus (Cav2-Cre) that labelled a specific projection population and activated a cre-dependent conditional virus in OFC that labelled synaptic terminals with a fluorescent marker (AAV8-FLEX-Synaptophysin-EGFP). We labelled outputs from OFC to the ventral tegmental area, and identified their collateral branches in the superior colliculus and dorsal raphe, as predicted from MAPseq (Figure 13).

Next, we determined that these projection neurons chose their targets in specific combinations, representing motifs. Projection motifs are connectivity patterns that occur with a probability greater than expected by chance in a randomized network, with the assumption that the probability of a neuron projecting to any one region is an independent event. This model represents the null hypothesis that projections targets are selected independently and ensures that repeated motifs are not considered significant due to the different numbers of barcodes between regions (Figure 3C). The expected probability of co-occurring barcodes in our null hypothesis is then reduced to a statistical problem akin to simultaneously flipping a set of unfair coins. We calculated the probability of observed bifurcating neurons under our null distribution, yielding a p-value per bifurcation motif, and the fold change, that is, the number of observed over the number of predicted co-occurring barcodes (Figure 2D and E, $p_{pairwise,}$ Bonferroni-corrected). This analysis identified rare, but repeated subcortical projection motifs in OFC that included the ventral tegmental area and superior colliculus, the ventral tegmental area and dorsal raphe, and the basolateral amygdala and striatum (Figure 3D and E).

Orbitofrontal to striatal projection patterns

Unlike most other subcortical projection neurons, striatal projection neurons were largely non-dedicated and innervated multiple regions across both the cortex and subcortex. Here, we did not explicitly differentiate between the dorsal and ventral striatum, although the hand dissection focused on medial ventral striatal regions around the nucleus accumbens (Suppl 3 for dissection map). We found that striatal projection neurons from OFC were overrepresented in nearly all bifurcations (n=2 target areas) in our MAPseq dataset, with the striatum and basolateral amygdala identified as the most statistically over-represented motif, and the striatum and contralateral OFC as the most abundant bifurcation (Figure 3E and 4A).

One potential challenge of MAPseq arises from the dissection of passing fibers without local synaptic terminations that could result in false multi-target projection patterns (i.e., false positives). Passing fibers are a particular concern for the striatum because white matter bundles are embedded within it dorsally and centrally, and it is directly adjacent to the internal capsule that contains all descending fibers to midbrain, brainstem, and spinal cord (Berendse et al., 1992; Coizet et al., 2017). First, we avoided the internal capsule and dorso-lateral regions containing the majority of passing fibers during dissection. While past MAPseq experiments have demonstrated that barcodes in passing fibers make a negligible contribution to overall barcode counts (see Discussion or Chen et al., 2019), we next performed a series of validation experiments. We visualized multi-target projection neurons using classical tracing, selecting the contralateral OFC and striatum (n=2) as well as the basolateral amygdala and striatum (n=2), and the ventral tegmental area and striatum (n=3). We identified rare co-labelled cell bodies in a series of dual-color retrograde tracing

experiments that roughly matched the fraction of striatal collaterals in MAPseq (Figure 4B-D). However, tracing experiments under-estimated the overlap fraction between OFC and striatum at ~11% compared to MAPseq at ~70%. Several factors could lead to under-estimating cell counts in classical tracing such as limited field of view for cell counting or inefficient retrograde labelling, either due to viral tropism or the limited injection coordinates, while biases in MAPseq experiments (e.g., multiple barcodes per neuron, see Discussion) could over-estimate cell counts. Nevertheless, dual retrograde tracing validated the existence of striatal bifurcation motifs we observed with MAPseq (Figure 4E).

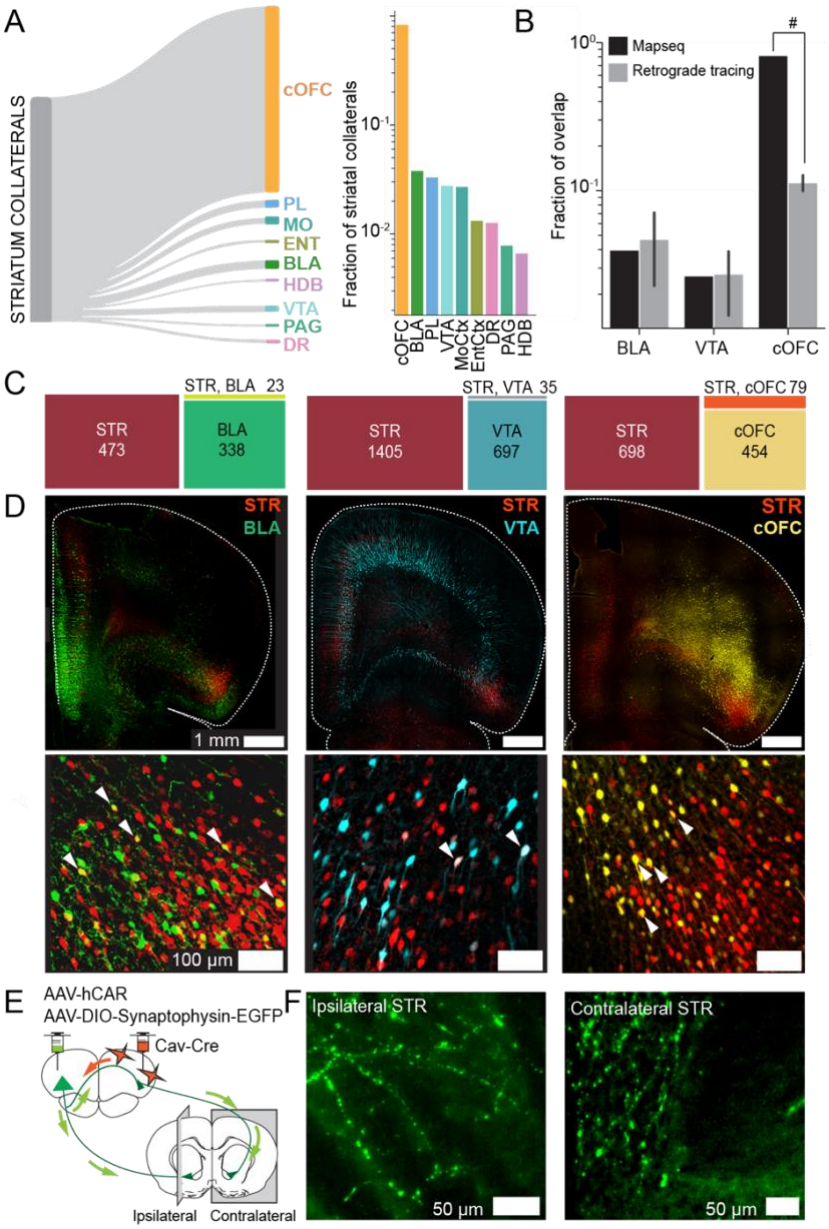


Figure 4 Striatal motifs are the most common multi-target patterns.

(A) Collaterals of multi-target striatal projection neurons from MAPseq dataset. **(B)** Overlap between projection populations in tracing experiments versus MAPseq; *pval* < 0.05 (#), one-sample t-test. **(C)** Overlap between projection populations at the individual cell level (neuron overlap: STR-BLA, 4.86% or 23/473 neurons from 2 animals; STR-VTA, 2.49% or 35/1405 from 3 animals; and STR-cOFC, 11.3% or 79/454 from 2 animals). **(D)** Dual retrograde tracing of bifurcating neurons targeting the striatum and basolateral amygdala (left), striatum and ventral tegmental area (middle), and striatum and cOFC (right). **(E)** Visualization of synaptic terminals for bifurcating neurons targeting the cOFC and striatum through injections of conditional synaptophysin-EGFP in the OFC and retrograde Cav2-Cre in the cOFC. **(F)** Axons and synaptic boutons in the ipsilateral (left) and contralateral striatum (right). *Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA,*

basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band.

We next identified synaptic terminals in striatum from the neuronal population that targeted the contralateral OFC. We used a dual viral strategy to identify synaptic connections in secondary targets of OFC neurons projecting to the contralateral OFC. To avoid fibers of passage in the striatum, we injected a conditionally expressing AAV8-Flex-Synaptophysin-EGFP virus in the left OFC and a retrogradely transported Cav2-Cre in the contralateral (i.e., right) OFC (Figure 4E). We identified synapses in the ipsilateral striatum and, unexpectedly, in the contralateral striatum, revealing that a subpopulation of cortically projecting neurons also targets the striatum bilaterally (Figure 4F). We further validated these synaptic connections by co-staining for the pre-synaptic protein Bassoon (Figure 10B). These observations reveal diverse projection patterns within the OFC to striatum, innervating both cortical and subcortical structures, in some cases bilaterally.

Single target subcortical projection neurons in OFC are organized into distance ordered sublayers

As subcortical projection populations represent distinct, connectivity-defined neuronal populations, we next asked if these might also be spatially organized within the laminar structure of cortex. The major projection neuron classes are known to be mostly localized within specific cortical layers. Intra-telencephalic neurons that target the striatum reside in superficial layer 5a and extra-telencephalic neurons that target the thalamus reside in layer 6 (Harris and Shepherd, 2015; Shepherd and Rowe, 2017). However, it is not known whether there is further spatial structure within layer 5 based on projection target. As most subcortical projection populations in OFC were defined by their connections to a single target, we hypothesized that these might also have distinct

laminar positions within layer 5b, the sublayer in which most extra-telencephalic projection neurons reside. We selected subcortical targets that included the striatum (n=4), ventral tegmental area (n=3), superior colliculus (n=4), dorsal raphe (n=4), and thalamus (n=2) (Figure 5 D-G). As previous anterograde experiments showed that OFC outputs to subcortical targets were highly targeted and easily separable, we first optimized single region injections and injected over a wide area to maximize the extent of labeling and minimize variability (see Methods for coordinates, Figure 14). We reasoned that we were less likely to label cells non-specifically (e.g., diffusion from the injection site), and more likely to miss cells altogether. We performed dual and triple retrograde injections in adult rats with different colored tracers (i.e., CTB 488, 594, 647 and retroAAV-CAG-GFP, retroAAV-CAG-tdTomato) to determine the spatial organization of soma positions of subcortical projection neurons in layer 5b. As expected, striatum projection neurons were the most superficial and thalamic projection neurons were the deepest. Projection neurons to the ventral tegmental area, superior colliculus, and dorsal raphe were localized within layer 5b in overlapping sublayers that were bounded in between striatal- and thalamic-enriched layers (Figure 5A, right).

Subcortical projection neurons were organized into intermediate sublayers within layer 5b that were ordered according to their distance of their projection target area. Because projection neurons to the superior colliculus are localized to more anterior portions of OFC, we focused on a single representative coronal plane 4.0 mm anterior of bregma for quantitative analysis, which also minimized curvature-induced distortions of the deep cortical layers. The superficial layer 5 border was approximated by staining against Cux-1, a classic layer 2/3 marker (Figure 5A, left). Soma positions were then calculated using the Euclidean distance from this superficial layer 5 border

across the different OFC subdivisions (Figure 5B). We then normalized these distances, with 0 indicating that a soma was localized at the layer 5 border, and 1 indicating the most distant soma observed. The median soma positions were ordered according to the distance of the target (Figure 5C); that is, projection neurons that target the ventral tegmental area tend to be more superficial than the superior colliculus, which in turn, is superficial to the dorsal raphe. We next calculated the approximate Cartesian distance from OFC to its subcortical targets using the stereotaxic injection coordinates (ignoring fiber trajectories). Descending projection neurons in OFC must traverse ~3 mm to the striatum, ~9 mm to the ventral tegmental area, ~11 mm to the superior colliculus, and ~14 mm to the dorsal raphe. The soma position of different subcortical projection populations correlated with the distance of their subcortical targets (Figure 6A). Thus, our results demonstrate a superficial-to-deep laminar sublayer organization that reflects a coarse anterior-to-posterior distance rule.

Figure 5 OFC projection neurons to subcortex are ordered in overlapping sublayers of L5b according to target distance.

(A) (left) Visualization of Cux staining delineating L2/3 and L5 highlighting the OFC subregion used for quantification of soma locations. Retrograde labelled projection neurons to superior colliculus are pseudo-colored in yellow. (right) Cartoon-ized representation of the left image highlighting L23 (grey), L5 (white), and L6 (black) with digitized superior colliculus projection neurons. **(B)** Soma locations of L5/6 projection neurons. **(C)** L5 depth of individual retrogradely-labelled projection neurons across OFC subdivisions. **(D)** Cumulative probability distribution of pooled, normalized distances (top) and the matching kernel density estimate (bottom) **(E-H)** Dual retrograde labelling of subcortical targets: **(E)** striatum and ventral tegmental area, **(F)** ventral tegmental area and superior colliculus, **(G)** superior colliculus and dorsal raphe, and **(H)** dorsal raphe and thalamus. *DLO dorsolateral orbitofrontal cortex; LO, lateral orbitofrontal cortex; VLO ventrolateral orbitofrontal cortex; Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA, basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band*

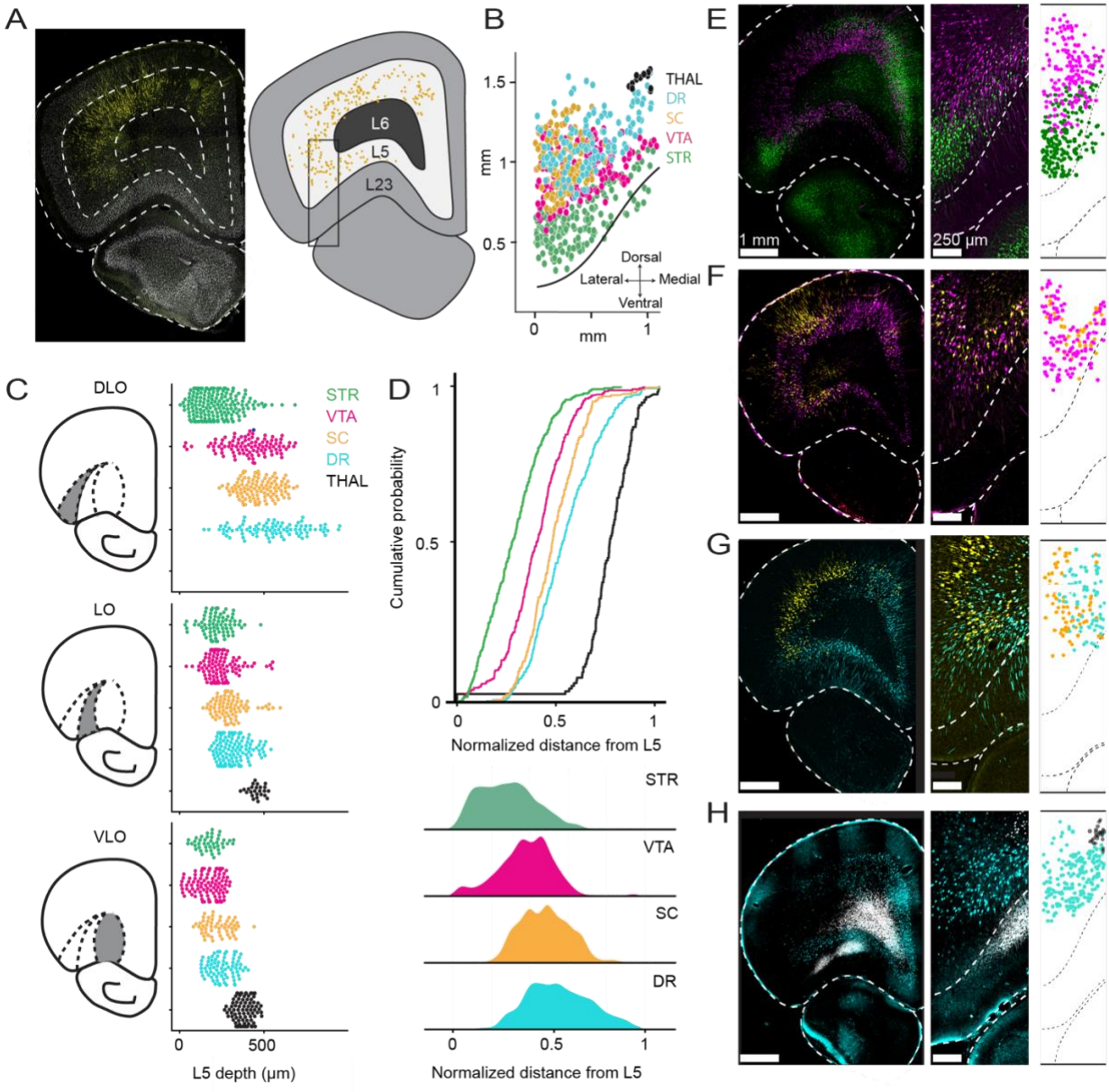
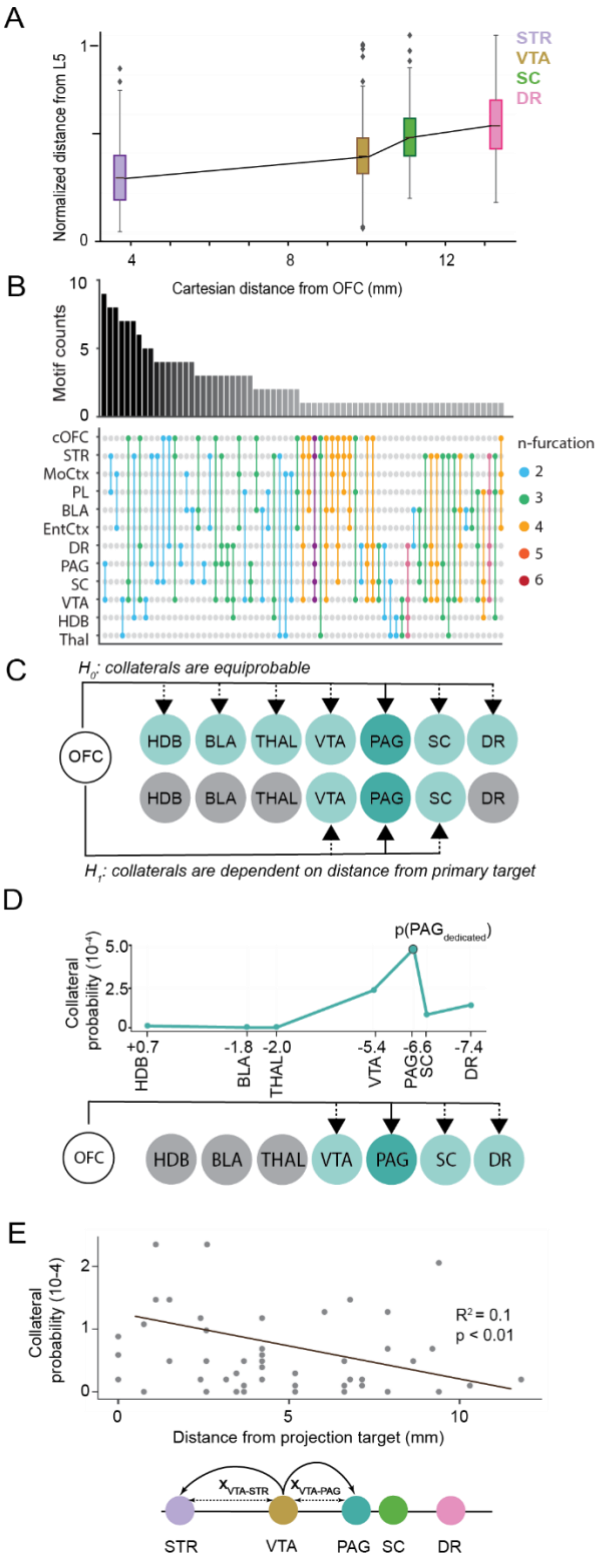

A
L6
L5
L23
B
THAL
DR
SC
VTA
STR
mm
1.5
1
0.5
0
1
Dorsal
Lateral
Medial
Ventral
C
DLO
LO
VLO
STR
VTA
SC
DR
THAL
0
500
L5 depth (μm)
D
Cumulative probability
1
0.5
0
0
0.5
1
Normalized distance from L5
STR
VTA
SC
DR
0
0.5
1
Normalized distance from L5
E
F
G
H
1 mm
250 μm

Figure 6 Rare multi-target projection neurons preferentially send collaterals to neighboring structures.

(A) The location of neuronal cell bodies in L5 were correlated with the distance of the projection target. **(B)** Rare, multi-target projection patterns that were observed less than 10 times. Discrete projection patterns are ordered in descending frequency (top), with the target combinations shown below each bar (bottom). N-furcation (n>1) is color coded according to the number of projection targets. **(C)** Null hypothesis that collaterals of rare, multi-target projection patterns are distributed randomly versus in a distance-dependent manner. **(D)** Probability of observing neurons that project to periaqueductal gray (PAG) and additional secondary targets. **(E)** Probability of observing co-innervated targets as a function of co-target distance (each point corresponds to a bifurcation motif). *Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA, basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band*



Rare multi-target projection neurons preferentially target neighboring structures

Our previous result that projection neurons were organized into ordered sublayers according to the distance of their target is suggestive of an overarching rule that guides global target selection. We therefore asked whether the connectivity patterns of rare, multi-target projection neurons might also be distance dependent. Although most multi-target projection patterns were observed less than 10 times across ~60,000 neurons, these account for much of the projection diversity (Figure 6B). Since we used a barcode detection threshold that minimizes false alarms (i.e., with a strict barcode molecule count threshold >10, see Methods), we believe these rare motifs represent genuine, but rare branching neurons. We then asked whether the individual targets of these rare multi-target projection neurons were distributed randomly across all targets, or if they show non-random structure (Figure 6B). Specifically, we hypothesized that neurons projecting to a given target area would tend to project to spatially close targets (Figure 6A). For example, we found that projection neurons to the periaqueductal gray preferentially branched to nearby areas such as the ventral tegmental area and the dorsal raphe (Figure 6C). To extend this analysis, we calculated the relative distance between a projection target and its collaterals. We then performed a linear regression on these distances and the probabilities of observed bifurcating neurons ($p_{airwise}$). We found that the likelihood of a collateral was inversely related to the distance from the reference projection target (Figure 6D, $R^2 = 0.1$, $p < 0.01$, regression analysis). Note that even if target areas were spatially close, they were nevertheless anatomically segregated, and overlap is unlikely due to dissection errors (see Methods). Thus, rare multi-target projection neurons tend to innervate targets in a distance-dependent manner that favors neighboring subcortical structures.

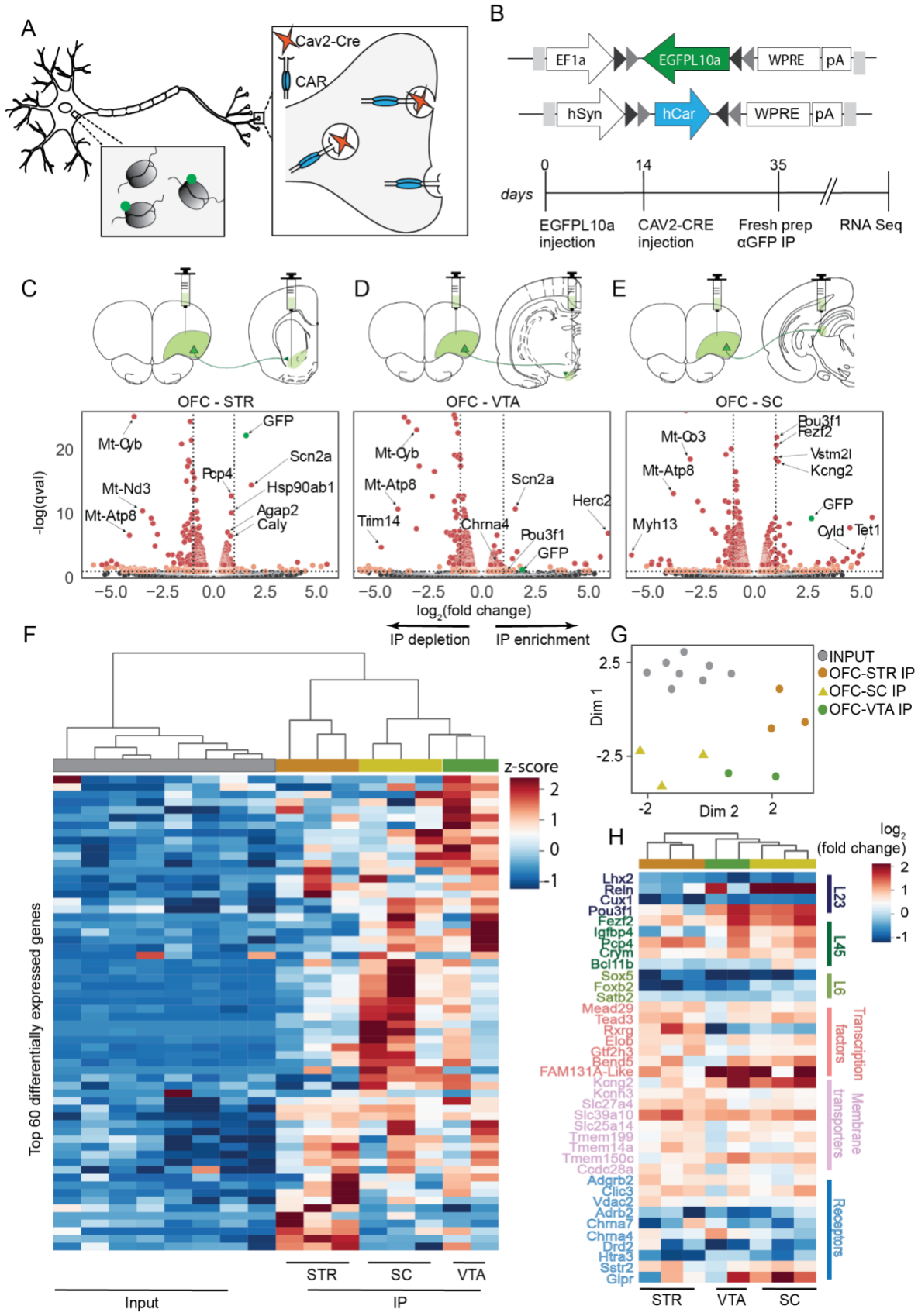
OFC projection neurons to Striatum, VTA, and SC represent molecularly distinct populations

We found that three types of projection neurons defined by their subcortical targets also represent molecularly distinct neuron types. Based on our results that subcortical projection neurons can be divided into anatomically distinct types based on projections and spatial location of their somata, we wanted to determine how these properties are reflected in their molecular landscape. We determined the gene expression patterns of three projection-defined OFC neural populations to the striatum, the ventral tegmental area and the superior colliculus. We combined a retrograde viral targeting strategy with viral TRAP (vTRAP, viral Translating Ribosome Affinity Purification) to tag translating ribosomes to capture cell-type-specifc mRNA followed by RNA sequencing (Figure 7A,B). Prior to being translated into a protein, mRNAs are attached to polysomes (Figure 7A). The conditional vTRAP construct (AAV5-DIO-EGFPL10a) delivers a ribosomal subunit protein N-terminally fused to a green fluorophore (EGFPL10a) that enables the specific pull-down, that is immunoprecipitation (IP), of ribosomes and their bound, translating mRNAs. This approach thus provides a snapshot of actively translated transcripts, i.e., the translatome. To ensure the robust expression of EGFPL10a that is not subject to preferential viral uptake across distinct projection populations (i.e. viral tropism), we primed neurons in OFC to uptake retrograde Cav2-Cre by over-expressing the native canine adenovirus receptor (i.e., CAR). We first injected a mixture of conditional AAV5-DIO-EGFPL10a and AAVdj-hSyn-DIO-{hCAR-Myc}$_{off}$-WPRE-pA in the OFC and then injected retrograde Cav2-Cre in subcortical projection targets to activate EGFPL10a in a cell type specific manner (Figure 7B). We then performed RNA sequencing on three anatomically distinct projection populations to the striatum (n=3), ventral tegmental area (n=2),

and superior colliculus (n=3) to identify differentially enriched transcripts across these anatomically distinct projection populations.

Figure 7 RNA-sequencing of subcortical projection populations demonstrates that anatomically-distinct cell types are also molecularly distinct.

(A) Schematic of dual viral strategy to selectively express EGPL10a in projection specific populations in an unbiased manner. Viral translating ribosomal affinity purification (vTRAP) is a technology to purify translating mRNAs from projection defined cell types. Retrograde delivery of Cre gates ribosomal integration of the GFP tagged ribosomal subunit EGFPL10a to allow for cell type specific immunoprecipitation and RNA sequencing. **(B)** Viral constructs (top) and experimental timeline (bottom). **(C-E)** Statistical significance versus fold change of transcripts in **(C)** striatum, **(D)** ventral tegmental area, and **(E)** superior colliculus projection populations. **(F)** Unsupervised hierarchical clustering of the top 60 differentially enriched genes across cell type specific pull-downs (IP) and transcriptomic background (input). **(G)** Multi-dimensional scaling of the top 60 differentially enriched genes. **(H)**Unsupervised hierarchical clustering of genes identified as layer markers from the literature or transcription factors, membrane transporters, and receptors from Panther gene ontology analysis. Included genes are either significantly (false discovery rate < 0.01) enriched or depleted in at least one projection population. *vTRAP, viral translating ribosomal affinity purification; IP, immunoprecipitated; STR, striatum; VTA, ventral tegmental area; SC, superior colliculus;*


A
Cav2-Cre
CAR
B
EF1a
EGFPL10a
WPRE
pA
hSyn
hCar
WPRE
pA
0
14
35
days
EGFPL10a injection
CAV2-CRE injection
Fresh prep αGFP IP
RNA Seq
C
D
E
OFC - STR
OFC - VTA
OFC - SC
-log(qval)
log2(fold change)
IP depletion
IP enrichment
Mt-Cyb
GFP
Scn2a
Pcp4
Mt-Nd3
Hsp90ab1
Mt-Atp8
Agap2
Caly
Trim14
Chrna4
Pou3f1
Herc2
Mt-Co3
Fezf2
Vstm2l
Kcng2
Myh13
Cyld
Tet1
F
G
H
z-score
Top 60 differentially expressed genes
Input
IP
STR
SC
VTA
INPUT
OFC-STR IP
OFC-SC IP
OFC-VTA IP
Dim 1
Dim 2
log2 (fold change)
L23
L45
L6
Transcription factors
Membrane transporters
Receptors
Lhx2
Reln
Cux1
Pou3f1
Fezf2
Igfbp4
Pcp4
Crym
Bcl11b
Sox5
Foxb2
Satb2
Mead29
Tead3
Rxrg
Elob
Gtf2h3
Bend5
FAM131A-Like
Kcng2
Kcnh3
Slc27a4
Slc39a10
Slc25a14
Tmem199
Tmem14a
Tmem150c
Ccdc28a
Adgrb2
Clic3
Vdac2
Adrb2
Chrna7
Chrna4
Drd2
Htra3
Sstr2
Gipr

Subcortical projection populations demonstrated distinct molecular phenotypes, mirroring their partitioned target selection and laminar distribution. Differential analysis identified 147 OFC-to-striatum genes, 94 OFC-to-ventral tegmental area genes, and 114 OFC-to-superior colliculus genes enriched in the cell type specific IPs over OFC input (Wald test, false discovery rate<0.05, Benjamini-Hochberg corrected; fold change (log_2) > 0.5) (Figure 7C-E). These include positive control genes such as GFP and the neuronal marker Scn2a, as well as genes related to calcium kinetics and synaptic plasticity (e.g., Pcp4), transcription factors (e.g., Pou3f1), neuromodulatory receptors and ion channels (e.g., Chrna4, Kcng2). We next performed unsupervised hierarchical clustering on the top 60 differentially enriched genes (Wald test, false discovery rate<0.01, fold change (log_2) > 1) across the cell type specific IPs and OFC inputs, and demonstrated close relationships across IPs from the same brain region (Figure 7F and G, see Supplemental figure 14 showing similar relationships in an analysis across all genes with significant differences in enrichment). Across the cell type specific IPs, the first major branch separated target-defined projection populations by their developmental origins. The intra-telencephalic striatal projection population was more distinct from the other 2 extra-telencephalic populations. In turn, 2 ventral tegmental area samples were most similar to one another, and 2 out of 3 superior colliculus samples were also most similar to one another. Surprisingly, 1 superior colliculus sample did not cluster with its target population and was more similar to the ventral tegmental area samples. While this might reflect sample quality (e.g., variability across rats, injections, or in sample processing), we previously demonstrated the diversity of connectivity patterns at a single neuron resolution. From MAPseq, we estimated that ~20% of projection neurons to the superior colliculus also sent collateral branches to the ventral tegmental area, making up the second most common

subpopulation (Figure 2E). As such, this outlier superior colliculus sample might be more similar to the ventral tegmental area because it is enriched in a subpopulation that also sends collaterals to the ventral tegmental area. Taken together, these results reveal that the pattern of connectivity is strongly reflected in their genetic composition.

We next refined this analysis to evaluate genes that might specifically play a role in neuronal development. We selected for genes that were implicated in mediating laminar position and cell signaling. Target-defined populations showed distinct developmental landscapes. From a gene ontology analysis and literature search, we curated a gene list that included transcription factors, membrane transporters, receptors, and classic layers markers in mice (Molyneaux et al., 2007; Zeng et al., 2012). Hierarchical clustering revealed that the molecular landscapes of these projection populations were largely distinct (Figure 7G). Again, we found that the striatum replicates were more distinct than the ventral tegmental area and superior colliculus replicates. Although the ventral tegmental replicates were separated, these did not form a tight cluster, unlike the superior colliculus replicates that emerged from a distinct branch. The overall expression pattern of known cortical layer markers was largely concordant with findings in mice, but we also observed a number of notable differences. For instance, Pou3f1 and Reln are transcription factors expressed in layer 2/3 neurons in mice were enriched in layer 5 neurons in rats. On the other hand CTIP2 (alias. Bcl11b), a prototypical deep layer marker in mice, was not enriched in our layer 5 projection populations.

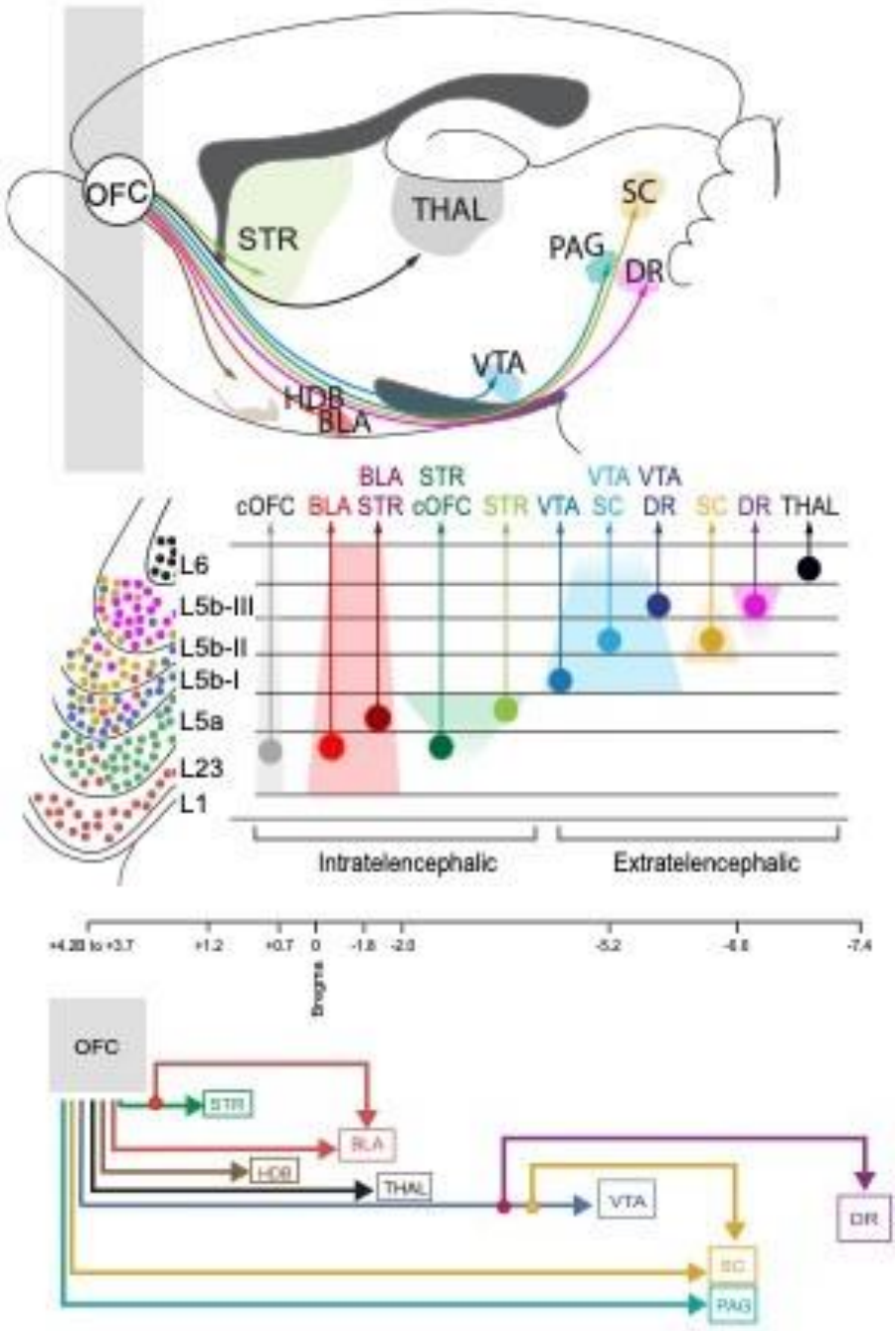


Figure 8 Projection cell types in rat OFC.

(A) OFC and its major downstream subcortical targets. **(B)** Subcortical projection populations can be defined by their connectivity to a single target and are spatially distributed in different sublayers of L5b. **(C)** This anatomical division-of-labor identifies distinct output channels for future functional studies.

Discussion

Here, we used a multi-modal approach to uncover the architectural logic of subcortical projection neuron types in the rat OFC. We established a comprehensive map of OFC outputs to subcortex at a single neuron resolution that has the statistical power to differentiate a one-off pattern from a conserved one. We found that subcortical projection neurons preferentially target a single structure and complemented this high-throughput single neuron resolution projection mapping with dual fluorescent retrograde tracing. Within OFC, we identified a high degree of laminar structure. These target-defined projection populations were organized into sublayers within L5, beyond the classic L5a and L5b distinction. These sublayers were ordered in a manner that matched the distance of the projection target, such that neurons in the deepest sublayer targeted the most caudal subcortical structure (i.e., dorsal raphe) and those in the most superficial sublayer targeted the most rostral subcortical structure (i.e., striatum). We next demonstrated that these anatomically distinct subcortical projection populations were also genetically distinct, with different patterns of transcription factors and membrane proteins. As such, these subcortical projection neurons represent distinct cell types that differ in their output connectivites, spatial distributions, and molecular phenotypes. We expect that the distinct connectivity and transcriptomic variation in subcortical neuron types provide the foundation for finding a cellular basis for how OFC supports complex behavioral functions.

Novel methods, such as whole-brain imaging and next-generation sequencing revolutionized anatomical studies in mice, whereas studies in rats have been typically confined to low-throughput, labor-intensive methods (Chen et al., 2019; Gabbott et al., 2005a; Han et al., 2018a; Muñoz-Castañeda et al., 2020; Murphy and Deutch, 2018; Oh et al., 2014; Winnubst et al., 2019). At the

same time rats remain one of the preferred model organisms to study the neural basis of complex behavior, and OFC has been identified as a core region for economic decisions resulting in a disconnect between our rich understanding of OFC neural activity and the fundamental lack of knowledge about its underlying circuit architecture (Constantinople et al., 2019b; Gardner et al., 2018; Groman et al., 2019; Masset et al., 2020; Singer et al., 2018; Steiner and Redish, 2014; Vandaele et al., 2016). Our results establish the output circuit architecture and major subcortical projection types in rat OFC. These subcortical projection types largely adhered to a one-neuron-one-target logic, a finding that suggests that these act as highly specialized information channels. Whether these outputs route different information to their targets, and whether these representations are uniform or mixed, remains to be seen. However, this work serves as a foundation for future circuit-driven functional studies that can link these dissociable circuits to neural representations to psychiatric concepts like compulsive reward seeking (Pascoli et al., 2018). Further, as cortical neurons are born in an inside-out manner, the laminar segregation of these subcortical projection types raises an outstanding question in development: target selection. From these local observations, we propose that these anatomically segregated output channels reflect a fundamental 'deeper farther' organizational principle that contrains the flow of information from frontal cortex to phylogenetically older, specialized structures in subcortex.

Technical considerations and limitations

We found evidence that a large subpopulation of subcortical projection neurons in OFC target a single area. Although MAPseq represents a powerful technology to rapidly map the projectome in mice and rats, there are both practical and inherent limitations for this finding. First, even though hand-dissection maximized RNA yield, hand-dissection limited the dissection precision and the

number of areas we could include in our study. Second, like all viruses, Sindbis barcoding virus likely suffers from biological tropism, that is, variable uptake and expression. Finally, limitations in detecting barcodes could lead to missed targets. Previous MAPseq studies in mice do not identify biases in detecting single versus multi target projection patterns and demonstrate comparable sensitivity of MAPseq to Lumafluor retrograde beads and recently, fluorescence-based single cell tracing (Chen et al., 2019; Han et al., 2018a; Huang et al., 2020; Kebschull et al., 2016a). Further, in this study, we ensured sufficient sequencing depth and minimized false negatives by setting a low detection threshold in our analysis. Therefore, we believe that ~60% dedicated input to subcortical areas represents the lower bound estimate of the contribution of these specialized information channels. At a high detection threshold that minimizes sequencing noise, the dedicated input to subcortical areas is ~90% (Figure 12D). Though subcortical projection neurons are better well-known for their axonal exuberance and wide-spread collateral branching, this is only the second large-scale study of subcortical projection neurons at a cellular resolution, and the first in rat frontal cortex. Even within the much more diverse intra-telencephalic projection class, however, dedicated outputs in represented a substantial form of inter-areal communication (~23%) in visual cortex (Han et al., 2018a). Additional theoretical concerns of MAPseq include degenerate barcode labelling (i.e., multiple barcodes per neuron), insufficient library diversity, and the detection of barcodes in passing fibers (recent studies indicate this is a minor source of contamination due to the active transport of barcodes and small amounts of axonal RNA, see (Chen et al., 2019) for further discussion). These concerns are described in detail elsewhere (Chen et al., 2019; Han et al., 2018a; Kebschull et al., 2016a) and would lead us to further underestimate the number of single-target neurons.

We complemented single-cell tracing with classical bulk tracing, and further demonstrated that connectivity-defined projection populations reside in overlapping sublayers in layer 5b. This, however, is also subject to several limitations. First, injection specificity is inherently limited by the selection of injection coordinates as well as the diffusion and neuronal uptake (e.g., tropism) of the tracer. Although the broad and patchy nature of subcortical projections from OFC make barcode contamination across brain regions unlikely (as in hand-dissection), the selection of these regions was, itself, based on viral anterograde tracing. In anterograde tracing experiments, we focused on the anterior segment of OFC, from approximately +4.5 mm to +3.5 mm anterior of bregma. Second, high inter-injection and inter-animal variability renders the high-level identification of projection patterns and soma distribution difficult. Even if viral tropism might be circumvented via either a receptor complementation strategy (Li et al., 2018), or the use of non-viral tracers, studying many projections is unfeasible given the limitation to two to three areas per animal. Therefore, our findings that projection neurons are localized in laminar sub layers rests on pooling results from multiple brains, making quantitative comparisons of soma localization challenging. Although we were able to replicate our main findings using double- and triple-retrograde-tracing, further studies are needed to better refine the laminar distribution of connectivity-defined cell types in a high-throughput manner.

We found that three major connectivity-defined populations - the striatum, ventral tegmental area, and superior colliculus - also represent molecularly distinct populations. Although vTRAP is an extremely versatile technique, especially in concert with an unbiased receptor complementation strategy (Li et al., 2018), it is limited in its resolution (Nectow et al., 2017). First, as a pull-down technique, it suffers from a higher degree of background and is therefore, less sensitive to

depletion. This might be especially the case in rats where tissue sizes are much larger and there is more non-neuronal debris (e.g., myelin, fibroblasts). More relevant to this study, however, is that as a bulk sequencing technique, we were unable to resolve the within-population heterogeneity. We made this decision to trade-off single cell resolution for high sensitivity to detect low copy number transcripts and capture population characteristics that might reveal more generalized developmental or functional similarities. Despite this population-driven approach, however, the outlier superior colliculus sample suggests that it includes the second most common subpopulation of superior colliculus projection neurons, that is, those that also send collaterals to the ventral tegmental area. While this raises the intriguing possibility that the connectivity patterns of subcortical projection neurons predict molecular phenotypes, it also represents a major challenge for future single-cell studies considering the diverse (and perhaps stochastic) patterns at a single neuron resolution (Lui et al., 2021). Although these three techniques each have limitations, taken together, they provide three independent lines of evidence that descending projection neurons in OFC represent distinct cell types.

Intra-telencephalic projection neurons: an exception to the typical one-neuron-one-target logic in OFC

Across both frontal and sensory cortex, the diversity of intra-telencephalic projection patterns is unique, and at times, seemingly random (Oh et al., 2014; Wilson, 1987; Winnubst et al., 2019). Here, we focus on cortico-striatal projection neurons, a neuron type within the intra-telencephalic class. OFC-to-striatum projecting neurons demonstrated the most diverse output patterns at a single neuron resolution, sending collaterals to almost every structure, both cortical and subcortical. While cortico-striatal neurons are often treated as a definitive cell type, these findings

are in line with both seminal and modern single cell tracing studies across cortex that also report incredible projection diversity across intra-telencephalic projection neurons (Muñoz-Castañeda et al., 2020; Wilson, 1987; Winnubst et al., 2019). Nevertheless, we identified statistically non-random collateral branches in the basolateral amygdala, and a putative tri-furcation that targeted the striatum bilaterally as well as the contralateral OFC.

OFC outputs to striatum are thought to have distinct roles in learning and choice behavior (Gremel and Costa, 2013; Hirokawa et al., 2019; Malvaez et al., 2019). We previously recorded from this projection population and found that their single neuron responses were highly uniform (Hirokawa et al., 2019). Indeed, in rats, these OFC-to-striatal projection neurons are reported to encode outcome information in varied behavioral tasks, from a two-alternative forced choice to probability reversal (Groman et al., 2019; Hirokawa et al., 2019; Hocker et al., 2021). As such, we did not expect to identify such diverse striatal projection patterns. We speculate that this might reflect the striatum's more recent evolutionary role in valuation. Current evolutionary hypotheses propose that reptilian cortical glutamatergic neurons are homologous to early born mammalian deep cortical layers (L5-6), and that the more superficial neurons, like the striatal projection neurons in layer 5a, were subsequently added in the mammalian lineage (Nieuwenhuys, 1994; Shepherd and Rowe, 2017). Thus, the diversity of striatal projection patterns might reflect the need for outcome information to be widely distributed across the brain.

Recent large-scale transcriptomic mapping in motor cortex demonstrated that intra-telencephalic projection neurons do not constitute discrete cell types, but rather are largely continuous in their gene expression (Zhang et al., 2021). This lack of discrete genetic subtypes makes it challenging

to understand how single neuron patterns of connectivity might arise in development, but also somewhat accounts for their inscrutable choice of targets. In contrast, extra-telencephalic neurons in motor cortex are consistently identified as discrete transcriptomic and anatomic types (Economo et al., 2018a; Sahni et al., 2021a; Yao et al., 2021; Zhang et al., 2021). This likely reflects a fundamental difference in the mode of inter-areal communication across these projection classes. However, there is a hint at a higher-level of organization in motor cortex that comes from the matched gradient of gene expression and cortical depth for intra-telencephalic neurons. Whether this laminar organization might also resolve continuous neuron types within the OFC-to-striatal population remains to be seen. However, the further refinement of cortical architecture at a cellular resolution is important not only for a better understanding of cell types, but also because it has an intrinsic relationship to cortical development.

"Deeper farther" laminar structure and its relationship to "inside out" cortical development

Our observation that there is a spatial gradient of projection neurons within layer 5b has intriguing developmental implications. Cortical layers develop from the "inside-out" generation of distinct neuronal classes, with the earliest born neurons in layer 6 (and layer 1) and the late born neurons in layer 2 (Angevine and Sidman, 1961; Caviness and Rakic, 1978; Frantz and McConnell, 1996; Gilbert and Kelly, 1975; Harris and Shepherd, 2015; McConnell and Kaznowski, 1991a; Rakic, 1974; Shepherd and Rowe, 2017). We hypothesize that this 'deeper further' laminar structure might result from the staggered birth of subcortical projection neurons, such that neurons that project to the furthest, and incidentally the most primitive structures, are born first, with the rest following in descending axon length order.

The organization of projection neuron types into distance-ordered sublayers implies a more directed mechanism for target selection in development. Pioneering tracing studies in rats suggest that all subcortical projection neurons adhere to a stereotyped branching pattern and extend their parent axon to the spinal cord, prior to forming pre-specified collateral branches (O'Leary and Koester, 1993; O'Leary and Terashima, 1988). Indeed, these mature branching patterns in rats do not develop until the first or second post-natal week of life (Bates and Killackey, 1984; Gribnau et al., 1986). Though past heterotopic cortical explant experiments demonstrated that regional-specific factors shaped connectivity patterns, it is unknown how neurons choose which branches to eliminate (Fishell, 1995; Schlaggar and O'Leary, 1991; Stanfield and O'Leary, 1985; Stanfield et al., 1982). ET projection neurons in mouse motor cortex were incredibly diverse, a finding that suggests either a semi-random or at least, complex mechanism for target selection (Callaway et al., 2021; Muñoz-Castañeda et al., 2020; Winnubst et al., 2019). However, the ordered structure in rat frontal cortex raises the possibility that there is an overarching organization that might be noisier in mouse motor cortex. We were able to identify this "deeper farther" principle because subcortical projection neurons in OFC were defined by their connections to a single target, this claim is separate from that of dedicated lines. First, we do not identify a strong degree of overlap across these major subcortical projection targets. Indeed, future experiments at a higher spatial resolution might reveal that these target-defined projection neurons are better described as regionally-defined, these limited targets still resolve this population into molecularly- and spatially- distinct cell types. Second, as briefly mentioned, single cell transcriptomic sequencing in mouse motor cortex also consistently identifies distinct groups of subcortical projection neurons that have distinct spatial distributions and connectivity patterns. An analogous "deeper farther"

principle is difficult to assess because mouse motor cortex subcortical projection neurons are mostly multi-target. However, a consistent finding seems to be that cortico-medullary neurons are in the deepest sublayer of L5b, whereas cortico-pontine and cortico-tectal neurons tend to be more superficial (Economo et al., 2018a; Muñoz-Castañeda et al., 2020; Yao et al., 2021; Zhang et al., 2021). This unclear relationship between laminar position and target distance in mouse motor cortex might suggest that this "deeper farther" principle does not reflect target distance per se, but rather a correlated factor such as axon length or the roughly caudal-to-rostral maturation of subcortical targets (Altman and Bayer, 1978, 1980, 1981; Bayer, 1980; Fentress et al., 1981; Finlay and Darlington, 1995).

An anatomical framework for OFC with specialized subcortical output channels

Understanding the cell type-specific architectural logic of frontal cortex is critical to gain mechanistic insights into the neural substrates of cognitive functions. The micro-circuitry of orbitofrontal cortex (OFC) is thought to underlie a wide variety of cognitive functions including economic decisions, valuation and learning (Gallagher et al., 1999; Gottfried, 2003; Gremel and Costa, 2013; Lak et al., 2014b; Miller et al., 2020b; O'Doherty et al., 2001; Padoa-Schioppa and Assad, 2006a; Tremblay and Schultz, 1999). It is unknown, however, how neurons in OFC accomplish these functions and whether these can be mapped to separate circuits. Within OFC, recent studies suggest that such projection populations, defined by their connectivity to a single subcortical target, represent dedicated information and assume specialized functions for learning and decision-making (Bariselli et al., 2020; Groman et al., 2019; Hirokawa et al., 2019; Izquierdo, 2017; Lui et al., 2021; Malvaez et al., 2019; Namboodiri et al., 2019; Pascoli et al., 2018). These

findings point towards a modular output architecture with segregated streams of information to subcortical areas.

We obtained three independent lines of evidence to support the hypothesis that these connectivity-defined descending projection populations might represent functionally distinct cell types. First, descending projection populations in OFC are highly target-selective, being either dedicated or repeated multi-target motifs, suggesting that target selection is tightly regulated. Second, these descending projection populations reside in different laminar sublayers within layer 5b, adhering to a 'deeper further' organization. Third, descending projection populations are molecularly distinct, and express different combinations of transcription factors and surface receptors. These findings support an output architecture wherein subcortical projection types make-up anatomically segregated circuits. This architecture is a pre-requisite for functionally segregated circuits and raises the possibility subcortical projection types route tailored information to downstream targets, and thereby assume functionally specialized roles in decision-making.

Methods

Adult male Long Evans rats (~300–500 g) were used for the study (Envigo). Rats were group-housed and maintained on a reverse 12 h light/dark cycle. All procedures were carried out in accordance with National Institutes of Health standards and were approved by the Cold Spring Harbor Laboratory Institutional Animal Care and Use Committee and Washington University in St Louis.

Stereotactic Surgeries

Rats were anaesthetized with 1-3% isoflurane, injected with pre-operative buprenorphine (0.03 mg/kg) and placed in a Kopf stereotactic setup. For injections, we used pulled 5 uL calibrated glass micropipettes, cut to an opening diameter of ~10-20 um. Craniotomies were made with a dental drill and dura cut with a 27 gauge needle. In post-operative recovery, rats were provided with oral carprofen tablets and observed either daily for a minimum of 6 days, or to euthanasia.

For anterograde tracing experiments, rats (n=3) were injected with 1,800 nL of AAV1-CAG-TdTomato in OFC, evenly distributed across two injection sites (4.2 mm anterior of bregma (+4.2AP), 2.8 mm lateral of bregma (+2.8ML) and 3.7AP/3.2ML and across 3 or 4 depths (2.4–3.3 mm ventral from brain surface (2.4–3.3DV) in increments of 0.3 mm and 2.4–3.0DV, respectively). N=1 rat was injected bilaterally across two injection sites (+3.7AP $\pm$ 2.6) and across 3 depths (2.7 – 3.3DV in increments of 0.3 mm).

For MAPseq experiments, rats were injected unilaterally with 2,200 nl Sindbis virus (1 x 10^{10} genome copies/ml (GC/ml), diversity of > 10^6 different barcode sequences; CSHL MAPseq Core) in OFC at 5 target locations to attain broad coverage, to (i) cover the extend of rat OFC, (ii) reduce missing projections targets. We used two anterior locations (4.2AP; 2.4ML and 3.2ML; 2.4–3.3DV each (0.3 mm steps) and three posterior locations (3.6AP; 2.2ML, 2.8ML, and 3.4ML; 2.4–3.3DV each (0.3 mm steps)).

For synaptophysin experiments, rats were injected with a 1:1 mixture of AAV8-CAG double flox-synaptophysin-EGFP (Universite Laval – Neurophotonics, 5.8x10^{12} GC/ml) and AAV_{dj}-hSyn-DIO-{hCAR-Myc}$_{off}$-WPRE-pA (5.9x10^{12} GC/ml) in the OFC at two locations (4.2AP; 2.8ML;3.3–2.4ML (0.3 mm steps), 3.7AP;3.2ML;3.0–2.4DV (0.3mm steps)) and Cav-Cre

(IGMM Vector Core France, 1.7–2.5x10^{12} viral particles per ml) in the contralateral OFC (n=2) and ventral tegmental area (n=2).

For retrograde tracing experiments, either 150 nL of cholera toxin B subunit conjugated to fluorophores (CTB-488, CTB-555, CTB-647, Life Technologies Corporation) or retrograde AAVs ($AAV_{retrograde}$-CAG-GFP, 37825-AAVrg; $AAV_{retrograde}$-CAG-tdTomato, 59462-AAVrg) were injected at each coordinate into: (1) basolateral amygdala (2.0 – 2.4AP each (0.2 mm steps); 0.75ML; 7.2 and 7.4DV) (2) thalamus (2.0AP; 1.8ML; 4.0DV) (3) superior colliculus (6.6 and 6.9AP; 2.45 ML; 4.2 and 4.5DV) (4) dorsal raphe (7.4AP; 2.0ML; 6.6DV at a 20° angle).

For molecular profiling experiments, rats were injected with 4,400 nl of AAV5-IV-GFPL10 (final concentration: 5 x 10^{12} GC/ml) and AAV9-hSyn-DIO-{mCAR-Myc}$_{off}${ChETA-HA}$_{on}$-WPRE-pA (final concentration: 2.5 x 10^{12} GC/ml) at a 1:1 ratio in the OFC bilaterally at the same coordinates used for the MAPseq experiments. Following a 2 week incubation, 300 nL of Cav-Cre (1.7–2.5x10^{12} viral particles per ml) was then injected bilaterally in the either the striatum, ventral tegmental area, or superior colliculus at the coordinates used for retrograde tracing experiments.

Immunohistochemistry and standard imaging

For anterograde tracing experiments, rats were sacrificed ~10-14 days after injections. Rats were first anesthetized with 2-3% isoflurane and injected i.p. with pentobarbital (Fataldose, 150 mg/kg). Rats were then transcardially perfused with heparinized saline, followed by fixation with 4% paraformaldehyde (PFA) in phosphate buffered saline solution (PBS). Brains were dissected and post-fixed in 4% PFA overnight. We cut 80–100 μm thick coronal sections on a Leica VT 1200 vibratome.

For retrograde tracing experiments, rats were sacrificed 5-7 days after injections and processed as for the anterograde tracing experiments. We additionally performed free floating immunohistochemistry. Sections were first washed in PBS and background fluorescence quenched in a 30 minute 3% H_2O_2 solution, and then washed again in PBS. Sections were blocked for 1 hr at room temperature in PBS containing 0.3% Triton X-100 and 5% normal donkey serum, followed by overnight incubation at 4°C with primary antibodies: Rabbit anti-Cux (11733-1-AP, 1:1000), Mouse anti-GFP (Santa Cruz Biotechnology Cat# sc-9996, RRID: AB_627695, 1:500). The next day, after serial washes in PBS, sections were stained with Alexa Fluor conjugated secondary antibodies (Invitrogen, 1:500) at room temperature for 2 hours. After washing, the sections were counterstained with DAPI and mounted. We verified that injections were localized to different subcortical projection targets. Coronal sections were then imaged (Leica THUNDER Imager) to visualize retrograde labeling in OFC. Representative sections at ~4.2 mm anterior of bregma to visualize the full breadth of cortical layers 2/3 (L2/3) to 6 (L6).

For synaptophysin experiments, rats were sacrificed 5 weeks after injections and brains dissected as described above. Coronal (60 μm thickness) and sagittal (80 μm thickness) sections were cut on a Leica VT 1200 vibratome. In addition, selected slices were stained against the Mouse anti-Bassoon (ab82958), as described above. These sections were imaged at 63x oil immersion (Leica THUNDER Imager).

Likewise, we also imaged the MAPseq (Figure 11A,B) and AAV5-IV-EGFPL10a (data not shown) injection sites in initial pilot experiments to confirm expression and Cre-dependence

respectively. Regarding the latter, we confirmed the Cre-dependence of AAV5-IV-EGFPL10a in a previously validated PV-Cre^{+} rat line.

Confocal image registration

From n=1 bilaterally injected rat, selected slices were used in registration analysis to quantify pixel number and density, and averaged across the hemispheres. Slices ranged from +2 mm anterior to -8 mm posterior of bregma and were imaged on a Zeiss LSM710 confocal microscope at 300 um intervals.

The mouse Common Coordinate Framework (CCF) reference brain was first registered onto a serial photon tomography (STP) imaged fluorescent Nissl (NeuroTrace) rat brain. After registration, transformation parameters obtained were used to move previously modified Allen Reference Atlas (ARA) for mouse onto rat STP brain (Kim et al., 2017; Muñoz-Castañeda et al., 2020). NeuroTrace staining was used to visually validate registration accuracy. After manual validation of the new rat registered CCF, a subset of slices corresponding to the starting and end points of the confocal images were selected for posterior registration. A rat anatomically enhanced reference brain was created by correcting intensity and detecting intrinsic anatomical features using Sobel operator with custom python scripts. Confocal images were denoised by removing outer-brain artifacts and 2D aligned based on the rat registered CCF images subset. Images intensity was corrected and intrinsic brain anatomical features detected using Sobel operator with custom python scripts. Registration of the rat registered CCF onto rat confocal images was performed using 3D affine transformation. Similarity was computed using Advanced Mattes mutual Information metric by Elastix registration toolbox (Klein et al., 2010).

Image quantification

All images were quantified using CellProfiler and Fiji. The L2/3 border was manually traced in Fiji, and pixel coordinates imported into Python for analysis. OFC subregion borders were visually determined and manually traced based on Paxinos et al., and images cropped to separate the dorsal lateral, lateral, and ventral lateral OFC subregions. Cell bodies were detected via a built-in object identification module (i.e., Identify primary objects) in CellProfiler. The module settings were manually adjusted on-line per channel. To exclude axons, primary objects were then filtered based on their median eccentricity values (i.e., MeasureObjectSizeshape, FilterObjects). The pixel coordinates of these filtered objects were then imported into Python for further analyses.

Layer 5 depth was calculated as the perpendicular distance (i.e., minimum Euclidean) from the interpolated L2/3 border for the striatum (n=303 from 2 animals) , ventral tegmental area (n=307 from 3 animals), superior colliculus (n=247 from 4 animals), dorsal raphe (n=3, neurons=293), and thalamus (n=113 from 1 animal). To pool this data across the different OFC subregions, we normalized the L5 depth to the maximum subregion distance. That is, in each subregion, neurons hold a normalized distance value ranging from 0 to 1, with 1 being the most distant neuron from the L2/3 border. We then plotted the cumulative probability function and corresponding distributions using custom scripts in Python.

Due to the low frequency of co-labeled neurons in multi-retrograde-tracing experiments, colocalization analyses was restricted to qualitative, visual assessment.

MAPseq experiments and analyses

After a 44-hr incubation period, Sindbis-injected rats were rapidly decapitated, and their brains were immediately frozen on dry ice. Though we initially sought to preserve the AP spatial distribution by, for instance, taking a series of 100 um slices as independent samples, we found that the number of detected subcerebral barcodes was extremely low, potentially due to inefficient RNA recovery in small tissue samples. In addition, we explored more refined methods of tissue isolation using laser capture microscopy, but found the RNA quality of ethanol dehydrated tissue to be extremely variable, with RNA integrity numbers less than 7.0.

We sectioned 300 um coronal slices for hand dissection using RNA safe methods in collaboration with the MAPseq core at Cold Spring Harbor Laboratory (CSHL). The injection site and projection targets were kept frozen on a metal block and dissected using cold scalpels. We avoided fiber bundles and additionally dissected a peripheral border around projection targets that were not included in the later analysis. In total, we dissected all major subcortical projection targets identified in our anterograde tracing experiments and identified high-level projection patterns that would otherwise necessitate dual- or triple- color injections across tens to hundreds of rats. Brain regions were pooled across the antero-posterior axis. After dissection, the RNA for each target area was extracted in TRIzol, purified, and reverse transcribed to cDNA by the CSHL MAPseq Core in a protocol that maximizes barcode recovery [ref]. Dissected tissue samples from the orbitofrontal cortex, a strong projection target (e.g., the striatum), and a weak projection target (e.g., the periaqueductal gray) were then selected for quality control testing (i.e., DNA bioanalyzer,

RT-qPCR) prior to sequencing to ensure sufficient RNA quality (RNA integrity number > 7). Samples were pooled and sequenced by the Next Generation Sequencing Core.

In total, we sequenced 3 brains, with the first brain in one run (i.e., Mseq106) and the remaining brains in a second run (i.e., Mseq130). Pooled multiplexed libraries were sequenced via 36-bp, paired-end reads on an Illumina NextSeq500 High Output Flow Cell. The CSHL MAPseq Core processed the raw sequencing data as previously described (ref). Briefly, the barcode sequences were de-multiplexed across samples and a read threshold set to filter low read molecules that might be due to PCR and/or sequencing errors (minimum 2 reads per barcode in the OFC and minimum 3 reads per barcode in projection targets). After read thresholding, all barcodes with a hamming distance of less than 3 were collapsed to correct for errors in sequencing. From this, a barcode matrix was generated, wherein each row represented a single barcode, and each column represented a single dissected brain area. Each element of the matrix then corresponded to the molecule number of the barcode (UMI). From the negative control UMI distribution, a threshold UMI (either >2 UMI in dedicated analysis or >10 UMI in motif analysis) was then set to further filter barcodes that might be due to noise. These data were then binarized and concatenated across brains. Subsequent analyses were performed using custom Python scripts.

Sindbis trafficking in rats. Straddling this 44 hour time point, we performed a time series experiment ranging from 36 to 68 hours of incubation (Figure 10A). In 4 adult rats, we injected Sindbis virus in the OFC and rapidly extracted the brain at 8-hour intervals and immediately froze the brains on dry ice. As above, the CSHL MAPseq Core sectioned 300 um coronal slices for hand dissection of the OFC and distal subcortical targets including the periaqueductal gray, superior colliculus, dorsal raphe, and a negative control (i.e., the tip of the contralateral olfactory bulb). The

CSHL MAPseq Core then extracted and purified the RNA as previously described and ran a qPCR for barcode mRNA to determine if expression was significantly affected by Sindbis incubation time and/or axon length.

False positives that might result in false projection patterns. Sample cross-contamination in the RNA processing steps and sequencing noise might result in false positives and, in turn, false projection patterns. We therefore sequenced negative control samples that did not receive OFC input in parallel. Negative control samples were taken from random tissue punched, un-injected brains as well as from the olfactory bulb and contralateral medial colliculus in injected brains. Based on the distribution of unique molecular identifiers (UMI) for each barcode detected in the negative control samples, we set the threshold at the distribution tail (>10 UMI) to binarize the data. Though this differs from previous studies that subdivided cell types by their projection strength, it is unclear to us that this has a straight-forward functional interpretation. Rather, we classified cell types purely based on their unique projection patterns. That is, we did not differentiate a neuron that projects strongly to region A and weakly to B, versus one that projects strongly to region B and weakly to region A.

False negatives that might contribute to false singletons. Missing barcodes from multi-target projections could bias the detection of single-target barcodes. There are three main potential sources of error that could lead to false negatives: (1) inefficient barcode trafficking (2) low RNA quantity/quality and (3) insufficient sequencing depth. Previous MAPseq experiments that have either probed locus coeruleus projections that uniquely encircle the brain and innervate in a caudal to rostral direction, or directly compared MAPseq to single neuron tracing indicate that the trafficking efficiency is not dependent on distance. In addition, in the previously described time

series experiment, we did not find a correlation between incubation time and Sindbis expression in distal projection targets. As such, although the rat brain is much larger than the mouse brain, it is unlikely that false negatives are due to inefficient barcode trafficking to distal axons. Second, low RNA yield might result in lost barcode molecules and false negatives, however, the MAPseq RNA extraction protocol is optimized as previously described to produce the highest yield and is quality tested prior to submitting the samples for sequencing. Last, barcode molecules might be lost due to insufficient sequencing depth, resulting in a substantial proportion of undetected barcodes. To estimate the probability of not detecting a barcode, we fit a negative binomial distribution to the number of sequencing reads per unique barcode. The distribution of target reads is bimodal, with a distinctive peak at 0 that corresponds to PCR and sequencing noise. We left truncated this distribution at ~5 reads and find that the number of reads per barcode is described by a negative binomial distribution NB(8.18, 0.42). From this, we determined that the probability of not detecting a barcode,P(0 reads | barcode), is ~ 10^{-4}.

Further, we ran a separate analysis to determine the relationship between the theoretical drop-out rate in sequencing and the detection of false singletons. As the UMI distributions of the dedicated and non-dedicated populations are indistinguishable and we do not believe that there are significant biological biases in Sindbis infectivity or transport. We then took the population of non-dedicated neurons and progressively subsampled that dataset to determine the impact of dropouts on the false singleton detection frequency, that is, true motifs that are mistaken as dedicated neurons.

Motif over- and under representation analyses. We next sought to determine which, if any, multi-target projection patterns might represent non-random motifs. As an overview, we first performed

unsupervised hierarchical clustering on the binarized barcode matrix (agglomerative clustering using a farthest point algorithm with cosine distance metric). To estimate if any two target areas are strongly associated, we constructed a null model in which each neuron projected to each area independently. We first estimated the probability of a single neuron projecting to each area. We define the probability that a given neuron projects to area A_i as $P(A_i) = N_{Ai}/ N_{total}$, in which N_{Ai} is the number of neurons in the dataset that project to area A_i, i=1….k, for k downstream targets, and N_{total} is the total number of neurons in the sample. N_{total} is not readily observable as we do not know the absolute number of neurons in OFC infected by Sindbis virus and there are infected neurons project to none of the dissected areas. A UMI threshold was set to select for "high-confidence" cell bodies inside the OFC, that is, unique barcodes in OFC that were detected at 10x the minimum UMI count (e.g., 100 UMIs) that crossed the threshold determined from the negative control UMI distribution (e.g., 10 UMIs). We therefore inferred N_{total} by fitting an exponential curve to the left-truncated (i.e., >100 UMI) UMI distribution in OFC.

Next, to calculate statistically significant motifs, we assumed a simple model in which each neuron projected to each area independently. We first estimated the probability of a single neuron projecting to each area. We define the probability $P(A_i)$ that a given projects to region A_i as $P(A_i) = N_{Ai}/ N_{total}$, in which N_{Ai} is the number of neurons in the dataset that project to area A_i, i=1….k, for k downstream targets, and N_{total} is the total number of neurons in the sample. As N_{total} is not readily observable due to noise, we infer N_{total} by fitting an exponential curve to a left truncated UMI distribution (>50 UMI) of the "high-confidence" neurons orbitofrontal cortex. We then sum the number of barcodes from the left truncated UMI distribution and the predicted number of barcodes from the exponential fit to estimate N_{total}.

We then calculated the pairwise frequency across target areas. That is, the number of instances that any two regions co-occur, regardless of n-furcation. This results in a symmetric matrix, PairDf in which each row and column corresponds to an area and each entry corresponds to the number of co-occurring barcodes for that area pair. We then calculated a theoretical symmetric matrix that predicts the expected pairwise frequency across target areas from a binomial distribution, with the probability of a neuron projecting to A_i and $A_{ii:}$ $p(A_i)$ x $p(A_{ii})$ and the number of observations: N_{total}. From this distribution, we determined the statistical probability of observing the co-occurring barcodes PairDf ($p_{pairwise}$). For visualization, we then masked values with a $p_{pairwise}$ <0.01, -log transformed the matrix, and pseudo-colored positive z-scores (i.e., over representation) red and negative z-scores (i.e., under representation) blue. As an independent metric of over representation, we calculated the odds ratio using the pairwise frequency. We constructed 2x2 frequency tables and calculated the odds ratio:

$$\frac{(\mathrm{Ai}+\ \text{and}\ \ \mathrm{Aj}+)\ \mathrm{x}\ (\mathrm{Ai}-\text{and}\ \mathrm{Aj}-)}{(\mathrm{Ai}+)(\mathrm{Aj}+)}$$

As with all odds ratio analyses, this is biased to overestimate the impact of rare projection populations and as such, HDB was excluded from this analysis. Significance values were determined using a two-tailed Fisher's exact test and p-val < 0.01.

Distance-dependent collaterals

The Cartesian distances for different OFC targets was estimated to be the averaged anterior-posterior injection coordinates that were used in retrograde tracing experiments. The normalized distance from L5 was the Euclidean distance from the L2/3 border, as described above. The collateral probability was the pairwise frequency taken from the symmetric matrix, PairDf divided by the N_{total} estimate (i.e., non-log transformed data in Figure 3). We then performed a linear regression analysis in Python for these collateral probabilities and the distance between projection targets. For instance, the relative distance between the ventral tegmental area (AP: -5.0B to -5.7B, mean: -5.35B) and the superior colliculus (AP: -6.0B to -6.9B, mean: -6.45B) is 1.1 mm.

Molecular profiling experiments

In order to selectively express EGFPL10a in a descending projection population, we used a dual virus approach and injected conditional AAV5-DIO-EGFPL10a in the OFC and retrograde CAV2-Cre in subcortical projection targets. Variable viral update or tropism limits the reliable and unbiased targeting of projection neurons. Thus, to mitigate this biological tropism, we used a viral receptor complementation strategy to express coxsackievirus and adenovirus receptor (hCAR), the endogenous receptor for retrograde CAV2 internalization and transport in projection neurons. Briefly, rats were injected with a 1:1 mixture of AAV5-DIO-EGFPL10a:AAVdj-hSyn-DIO-{hCAR-Myc}$_{off}$-WPRE-pA. After a 2 week incubation period, we injected Cav-Cre in the striatum, superior colliculus, and ventral tegmental area.

Approximately 5 weeks post-injection, rats were rapidly decapitated and their brains hand-dissected on ice to isolate the OFC. The dissected brain tissue from a one rat constituted a single

sample (striatum: n=3, ventral tegmental area: n=4, superior colliculus: n=3). The dissected brain tissue was pooled across hemispheres and mechanically homogenized using a buffer containing 10 mM HEPES-KOH (pH 7.4), 150 mM KCl, 5 mM MgCl2, 0.5 mM DTT, 100 μg/ml cycloheximide, RNasin (Promega, Madison, WI) and SUPERas-InTM (Life Technologies) RNase inhibitors, and Complete-EDTA-free protease inhibitors (Roche). To correct for the increased tissue sizes in rats, we increased the volume of homogenization buffer to 3 ml and increased the volumes of additional reagents in concert in order to maintain the final concentrations. For instance, after tissue homogenization in 3 ml, we added a tenth of a volume of 300 mM DHPC (300 μl) and another tenth of a volume of 10 % NP-40 (333 μl) to clarify the supernatant. GFPL10-positive ribosomes were then immunoprecipitated via magnetic Dynabeads coupled to anti-EGFP (HtzGFP-19C8, HtzGFP-19F7) in a 1.5-hour incubation at room temperature. Prior to the elution of the cell type specific, GFP immunoprecipitated (IP) RNA, 50 μl of supernatant was removed to be used as the transcriptomic background (input). RNA from the IP and input were then processed in tandem using an adapted protocol from the Stratagene Absolutely RNA Nanoprep Kit, as previously described (refs). IP and input RNA samples were kept at –80°C and sequenced (CSHL Sequencing Core).

Sequencing and bioinformatics

The CSHL Sequencing Core prepared libraries from IP (cell type specific GFP immunoprecipitated) and input samples using an Ultralow Input NuGen Ovation RNA-seq Kit. RNA quality was assayed on an Agilent 2100 Bioanalyzer. Only samples with RNA integrity values > 7.0 were used for RNA-seq. In total, twenty samples were sequenced (10 IPs paired with 10 inputs – VS (3), VTA (4), and SC (3)). Pooled multiplexed libraries were then sequenced via

76-base pair, single-end runs on the Illumina HiSeq platform with ~40 million reads/sample. Reads were pseudo-aligned to rat reference genome (Rnor_6.0) through Kallisto with default single-end parameters. Estimated counts were imported into the Sleuth package in R/Bioconductor, and differential expression and fold change (IP/Input) was assessed using a Wald test adjusted for multiple comparisons to achieve a false discovery rate (FDR < 0.05, Benjamini-Hochberg).

The resulting data were analyzed and plotted in Python. Prior to further analysis, we assessed enrichment in control genes including GFP, CRE, and Rpl10a (Figure 15A). For downstream analysis, we required that samples had at least a 2x significant enrichment in GFP. As such, two ventral tegmental area samples (VTA1 and VTA3, Figure 15A) were excluded from analysis. We performed unsupervised hierarchical clustering with a cosine distance metric and multi-dimensional scaling on the top 20 enriched and depleted (absolute value of fold change) genes in each condition. We next performed a gene ontology analysis of these top 60 genes via the online Panther Gene List Analysis and mined for genes that were classified as membrane transporters, transcription factors, and receptors. From the literature, we then added genes that were identified as cortical layer markers to create a curated gene list and again, performed unsupervised hierarchical clustering with a cosine distance metric.

Supplemental figures

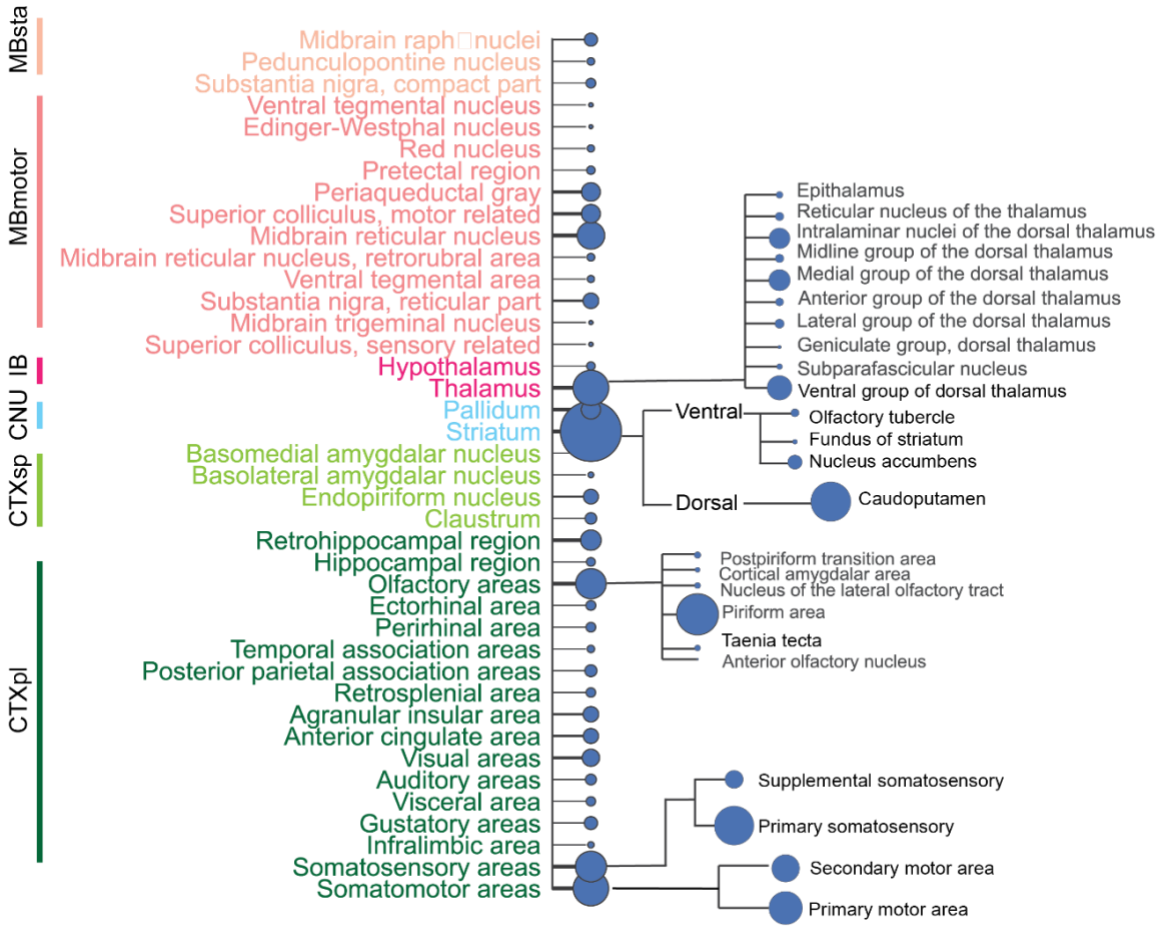


Figure 9 Extended projection strength quantification from whole-brain anterograde tracing.

A more detailed plot as in Figure 1, wherein the marker size reflects the relative projection strength. Leaves off of branches are ordered according the Common Coordinate Framework and reflect the relative projection strength (summed to 100%) for structure subregions.

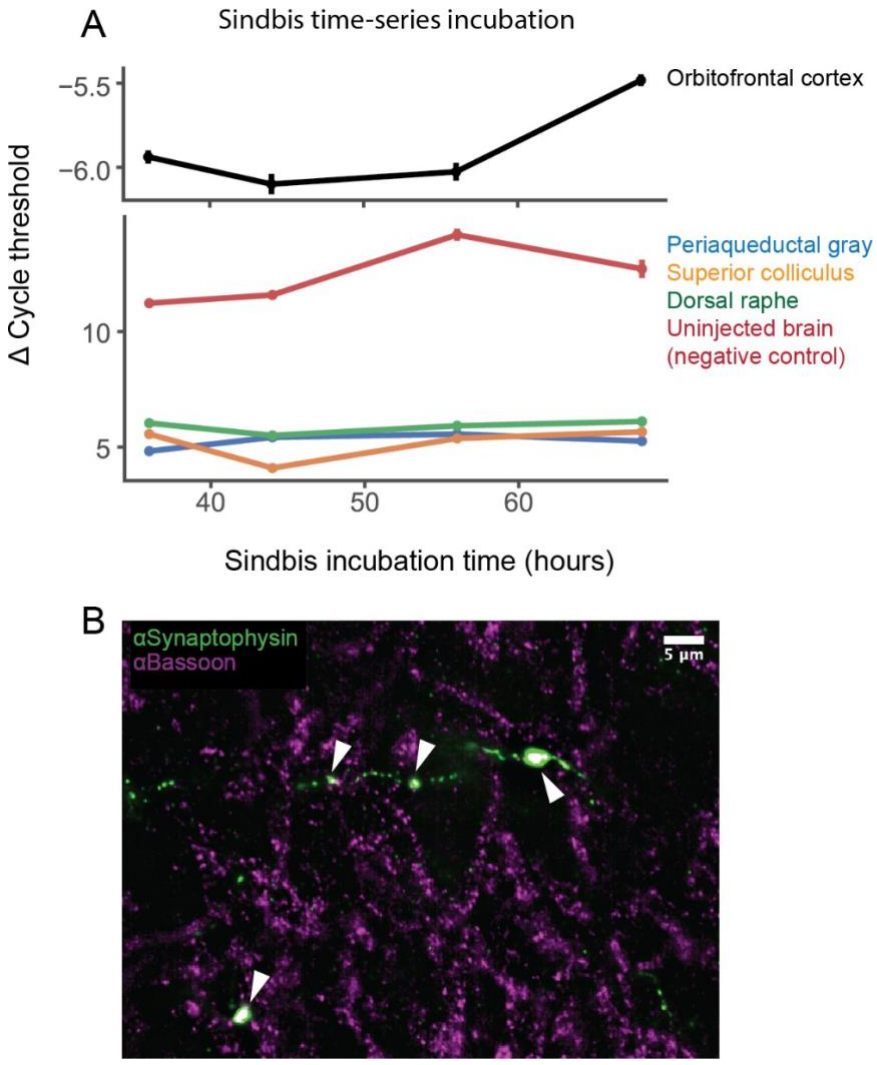


Figure 10 Sindbis time-series and synaptophysin control experiment

(A) Sindbis expression is quantified *via* qPCR of GFP in the orbitofrontal cortex (i.e., the injection site) as well as in distal projection targets. Delta CT values were normalized to the housekeeping gene actin and remained stable from 36 to 68 hours. **(B)** Staining of pre-synaptic marker Bassoon to ensure that striatal collaterals represented terminal synapses and were not passing fibers.

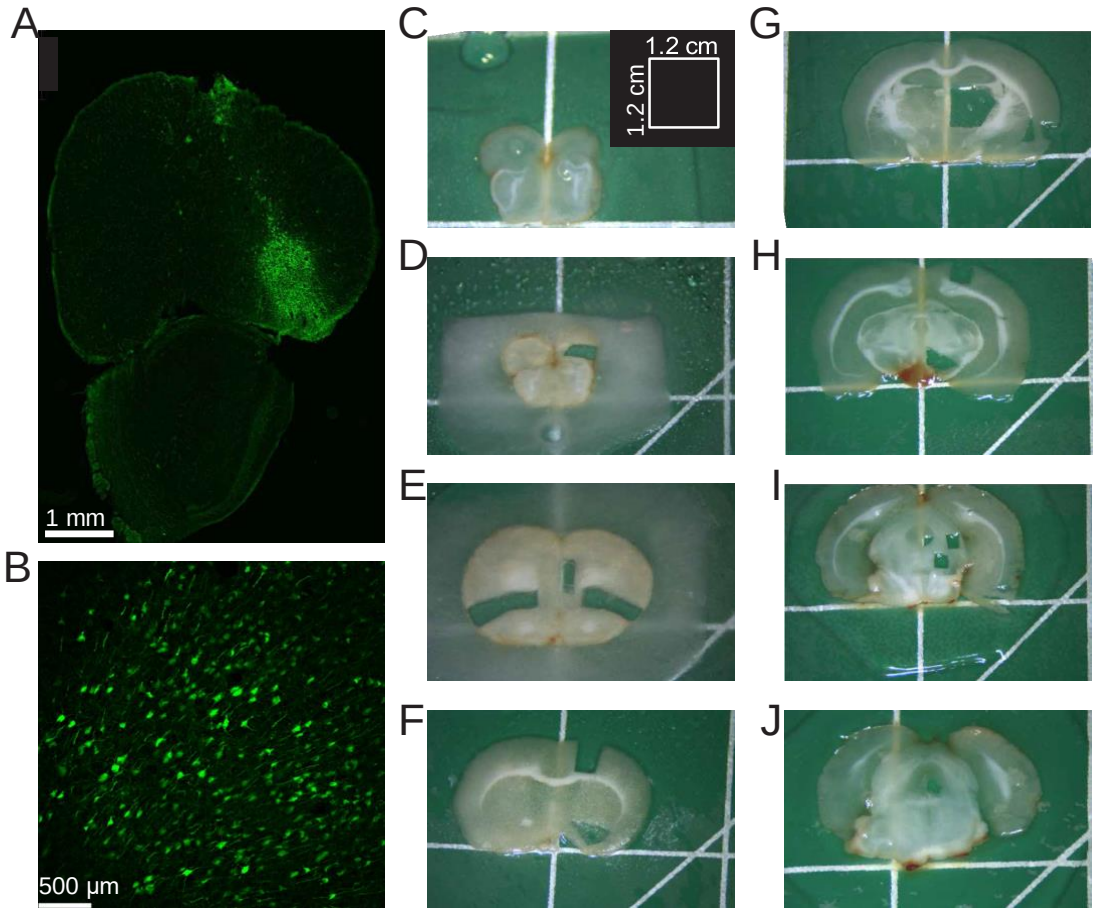


Figure 11 Sindbis injection and hand dissection of projection targets**.**

(A) Example Sindbis injection in OFC at 5x demonstrating limited spread at a single point of injection. **(B)** Example Sindbis injection in OFC at 20x. **(C)** Slice of OFC immediately anterior to (D). Example of hand dissected projection targets in **(D)** OFC, **(E)** contralateral OFC (left) and OFC (right), **(F)** striatum **(G)** thalamus **(H)** ventral tegmental area, **(I)** periaqueductal grey, superior colliculus, and mesencephalic locomotor region (not analyzed in this dataset) and **(J)** dorsal raphe.

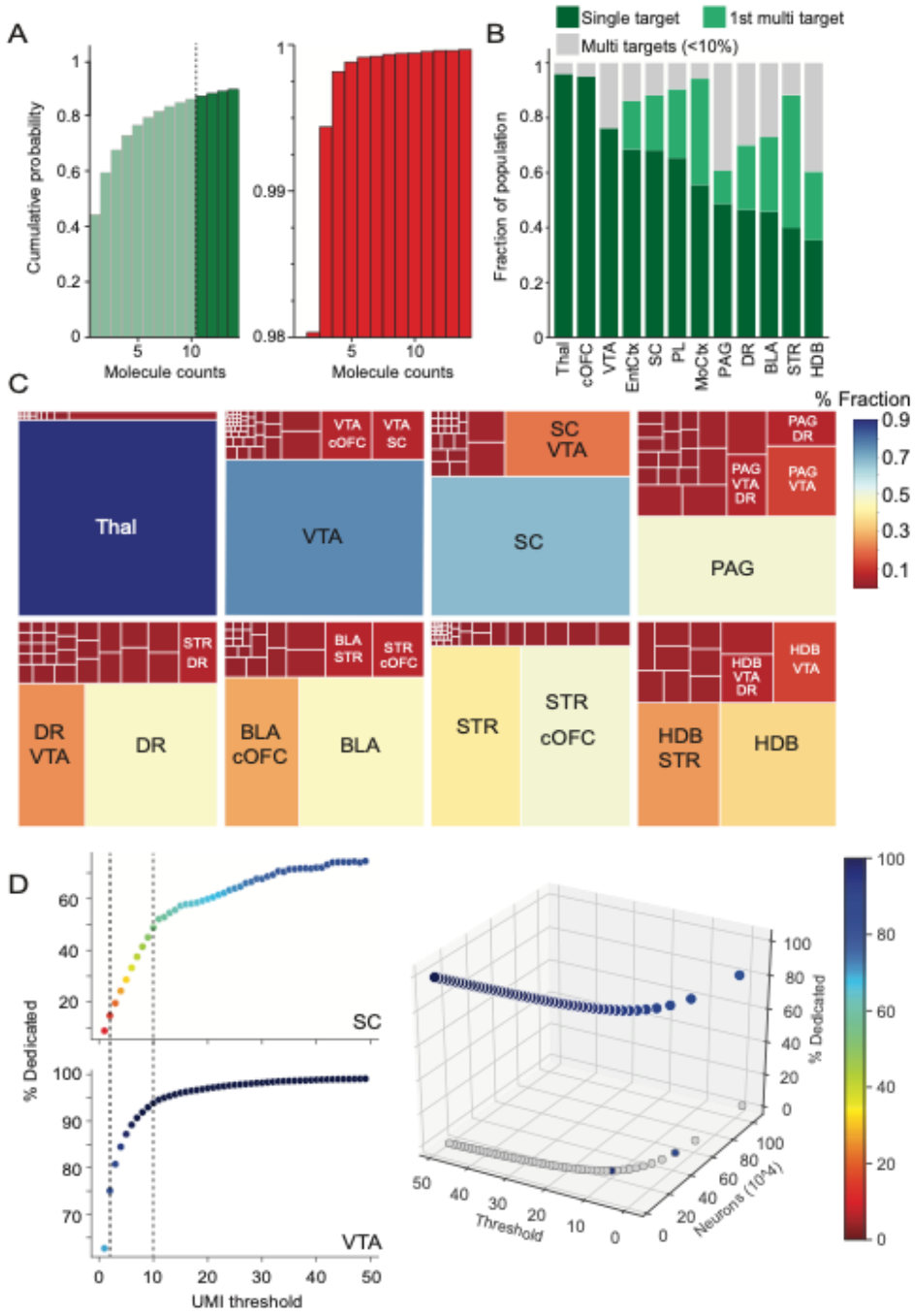


Figure 12 The fraction of projection neurons that target a single area is boosted using a higher threshold that minimizes noise

(A) The distribution of barcode molecule counts in target regions (left, green distribution) using a more conservative threshold. To minimize false projection patterns (e.g., false positives), we chose a conservative threshold (dotted line) of at least 11 barcode molecules, filtering barcodes due to noise. The false negative rate is < 0.001% using this conservative threshold (right, red distribution). **(B)** The fraction of single target projection patterns compared to the next most common multi-target projection pattern(s) above 10% for each target. Projection patterns below 10% are pooled and visualized in gray. **(C)** The fraction of single-target and multi-target projection patterns above 5% for each target region. Projection patterns below 5%

are not labelled for visualization. **(D)** Fraction of dedicated neurons as a function of UMI threshold for the SC and VTA (left). A 3D plot demonstrating the fraction of dedicated neurons in relation to the UMI threshold and total number of neurons in the dataset (right). *Thal, thalamus; cOFC, contralateral orbitofrontal cortex; VTA, ventral tegmental area; EntCtx, entorhinal cortex; SC, superior colliculus; MoCtx, motor cortex; PL, prelimbic area; BLA, basolateral amygdala; DR, dorsal raphe; STR, striatum; PAG, periaqueductal gray; HDB, horizontal diagonal band.*

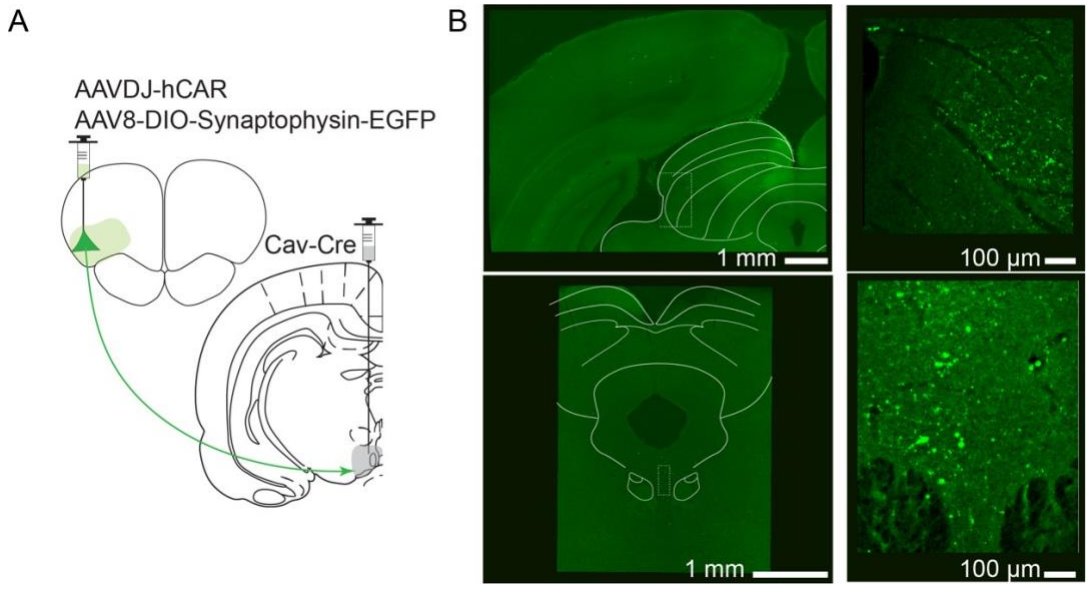


Figure 13 Bifurcating projection neuron from OFC to VTA and DR, and OFC to VTA and SC validated with synaptophysin**.**

(A) Cre-dependent synaptophysin injected in the OFC and Cav2-Cre injected in the VTA. **(B)** GFP fused synaptophysin visible in the superior colliculus and dorsal raphe. Note, that terminals are only visible in the ventro-lateral subregion of superior colliculus.

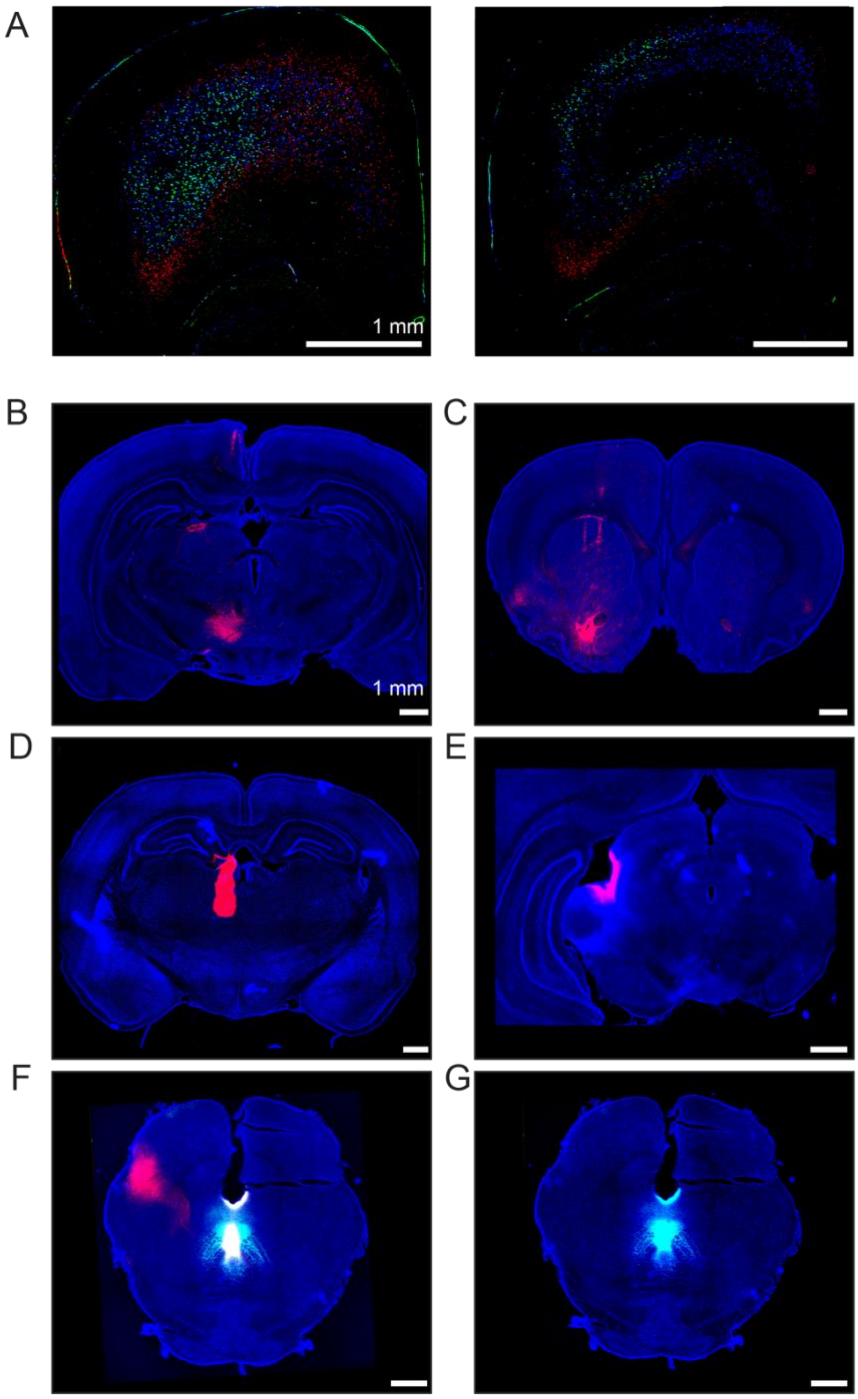


Figure 14 Injection sites for retrograde tracing

(A) Striatum (red), ventral tegmental area (blue), and superior colliculus (green) projection neurons in orbitofrontal cortex at ~ 4.6 AP (left) and ~4.2 AP (right). **(B-G).** Injection sites for retrograde tracing in (B) ventral tegmental area (C) striatum (D) thalamus (E) superior colliculus and (F) dorsal raphe. Retrograde tracers included CTB-594 (red) and CTB 647 (cyan).

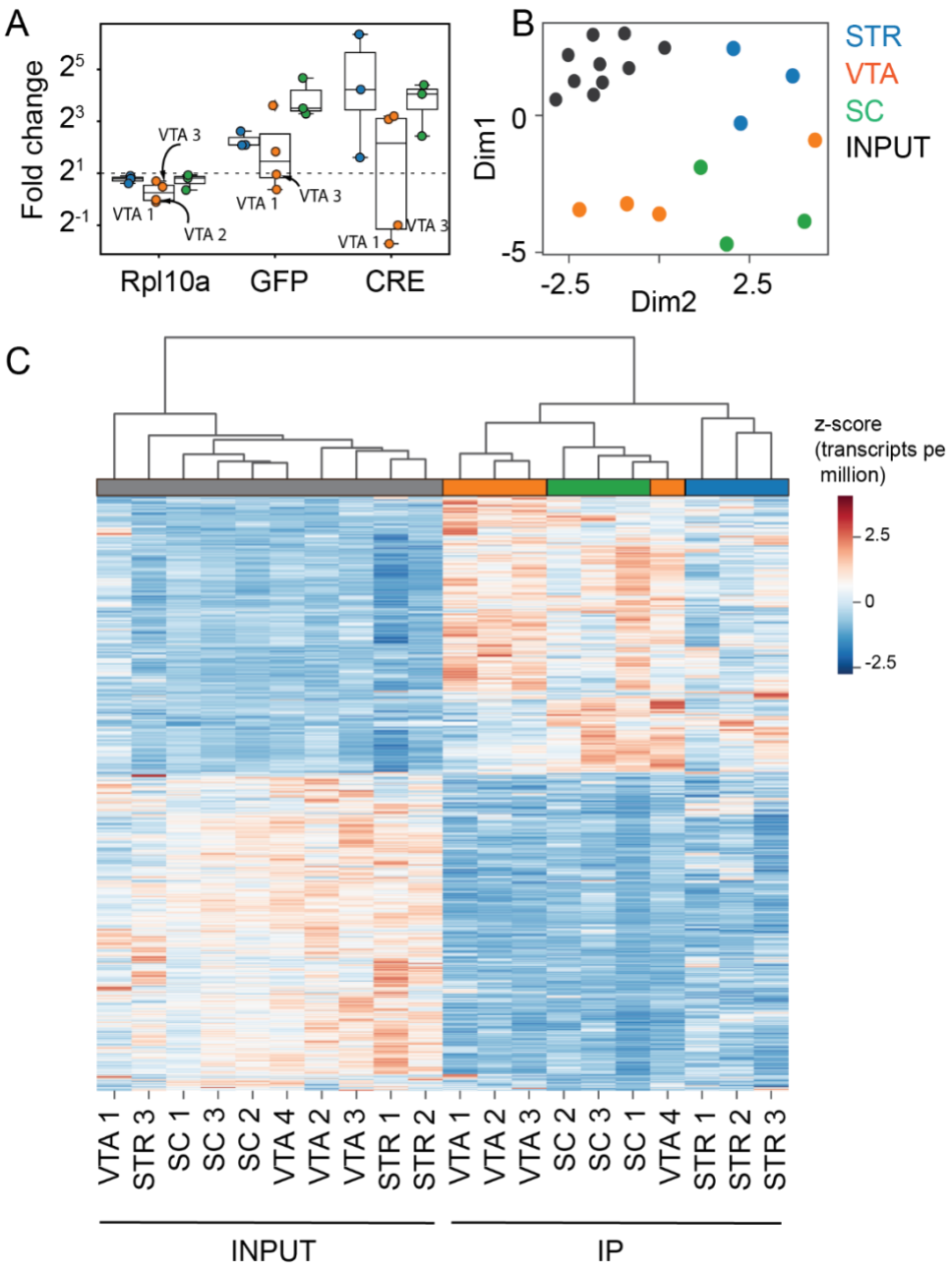


Figure 15 Samples largely cluster according to projection target

(A) Fold-change for control genes for RPL10a, GFP, and Cre. Despite a slight enrichment in GFP, we excluded VTA1 and VTA3 due to a minimal (less than 2 fold) enrichment in GFP and a depletion in Cre. **(B)** Multi-dimensional scaling reveals that samples largely cluster according to their projection target. **(C)** Likewise, hierarchical clustering of all significantly enriched genes demonstrates that samples largely cluster according to their projection target. Input samples cluster together and have similar depletions and enrichments.

Chapter 3

A Proposed Temporal Matching Mechanism for Subcortical Target Selection

Abstract

Cortex has exquisite spatial structure, with different layers corresponding to distinct cell types. Layer 5 projection neurons are the major output to subcortex and have extensive dendritic trees that span the depth of cortex, allowing them to integrate diverse sources of information. In orbitofrontal cortex, we identified distinct projection cell types that largely send information to a single subcortical target. Further, these output sublayers are ordered according to the distance of their projection target, in a "deeper farther" pattern. This pattern is reminiscent of the "inside-out" principle, wherein the laminar position of cortical neurons is established from the staggered birth of different neuron types. This led to the hypothesis that a "deeper-farther" spatial pattern might also reflect the temporal logic of development, with evolutionarily older subcortical regions receiving the first wave of connections for the deepest sublaminae of cortex. Here, I outline a possible developmental mechanism that could link sublaminar position and projection type. Such

an ordered mechanism underlying mesoscale target selection would support the existence of core, conserved projection patterns through which cortex exerts its control over subcortex.

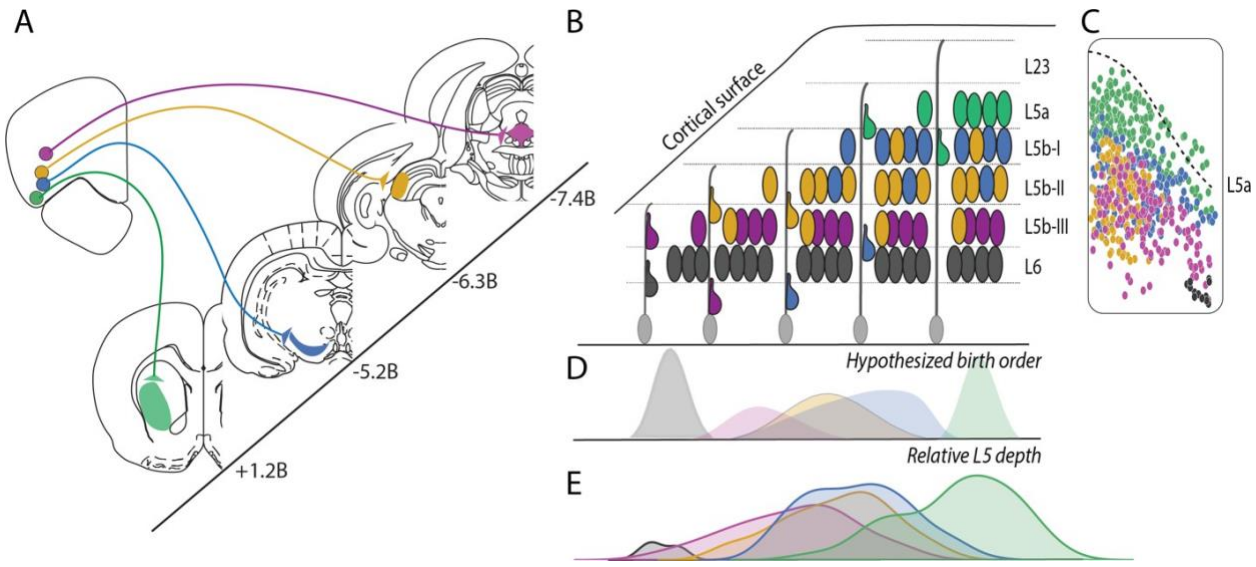


Figure 16 "Deeper farther" sublaminar organization in rat orbitofrontal cortex.

(A) Cartoon of the observation that the more superficially located projection neurons target more rostral structures, whereas deeper projection neurons target more caudal structures. **(B)** Proposed generation of projection types according to the "inside out" developmental rule and **(D)** hypothesized birth order. **(C)** Cell soma positions of projection neurons in rat orbitofrontal cortex (n=3) **(E)** Histogram of relative L5 depth from neurons quantified in (C).

Introduction

Through the millennia and pressures of evolution, cortical circuitry has been shaped to promote survival. This initial architecture preserves core conserved functions and sets the stage for animals to rapidly learn (Zador, 2019). Such an architecture is faced with the challenge of being able to flexibly support sophisticated computations, as well as communicate with subcortex, a phylogenetically older nervous system (Lake et al., 2016b; Sherman and Usrey, 2021). One possibility is that these functions are segregated across different cortical layers made up of distinct

projection types (Harris and Shepherd, 2015; Musall et al., 2021). Residing in the deep layers of cortex, subcortically projecting neurons are the major descending output from cortex and target diverse regions from midbrain to hindbrain to spinal cord (Baker et al., 2018; Leone et al., 2008; O'Leary and Koester, 1993; O'Leary and Terashima, 1988). Though often considered to be a single cell type, their diverse projection patterns have long defied this simplistic view (Kita and Kita, 2012; Lévesque et al., 1996b, 1996a; Winnubst et al., 2019). Determining the principles of cortical connectivity is critical for understanding brain function and constraining circuit computations.

Within cortex, only a small fraction of projection neurons mediate communication from cortex to subcortex (Gabbott and Stewart, 1987; Gabbott et al., 2005a; Zhang et al., 2021). Such a bottleneck is suggestive of an evolutionary mechanism to filter information from developing cortex to older subcortical circuits, but how many different projection types exist, and their overarching logic remains unclear. Recent large-scale circuit mapping has allowed for a bird's eye view of subcortical projection patterns at a single neuron resolution (Kebschull et al., 2016a; Muñoz-Castañeda et al., 2020; Winnubst et al., 2019). In rat orbitofrontal cortex, we found that subcortical projection neurons not only preferentially targeted a single region, but also were ordered into previously unappreciated sublayers; neurons positioned more superficially projected to more rostral regions and those positioned deeper projected to more caudal regions (Figure 16). This "deeper further" organization revealed a refined relationship between (local) sublaminar position and long-range target selection that has also been reported in motor cortex.

How such different projection patterns arise is an open question, but this sublaminar organization suggests an overarching rule for target selection. The inside-out birth of projection types is one such rule, wherein the six layers of cortex are formed from the staggered birth of mesoscale projections, that is, broad projection classes (McConnell and Kaznowski, 1991b; Rakic, 1974). The refinement of this rule to include sublaminar projection patterns is intuitive, but still requires an addendum to explain macroscale rather than mesoscale target selection. From our limited observations in rat orbitofrontal cortex, we hypothesize that there is relationship between a neuron's spatial distribution and its choice of synaptic partners. Such a finding raises the intriguing possibility that a global process manages the development of core conserved projection patterns.

The proposed relationship between laminar depth and subcerebral projection patterns

We reasoned that as the environmental factors present at a neuron's birth determine its laminar fate, these might also shape its projection pattern. However, as subcortical projection neurons are initially thought to be largely identical (O'Leary and Koester, 1993; O'Leary and Terashima, 1988), target specificity might arise not from the neurons themselves, but their projection targets. That is, subcortical regions might be primed for innervation during select time windows that align with waves of newborn neurons (Altman and Bayer, 1978, 1980, 1981; Jensen and Altman, 1982). Such temporal matching might naturally result in groups of projection neurons that preferentially innervate specific structures. If a region is resource limited and can only support a finite number of synapses, then early arriving neurons could attain a "first to market" advantage and block further innervation by late arriving neurons. Here, we extend this temporal logic to our local findings in orbitofrontal cortex and make three major assumptions: (1) subcortical projection neurons initiate

a common developmental program (2) neurons that are born earlier also initiate their developmental program earlier and (3) brain regions are resource limited and can only support a finite number of synaptic connections.

Subcortical projection neurons are thought to adhere to a generic developmental program that generates a stereotyped branching pattern, wherein a spinal axon is extended prior to the formation of common collateral branches in a caudal-to-rostral order (Arlotta et al., 2005; Lai et al., 2008; McConnell, 1988; O'Leary and Stanfield, 1985). As a result, it stands to reason that early born

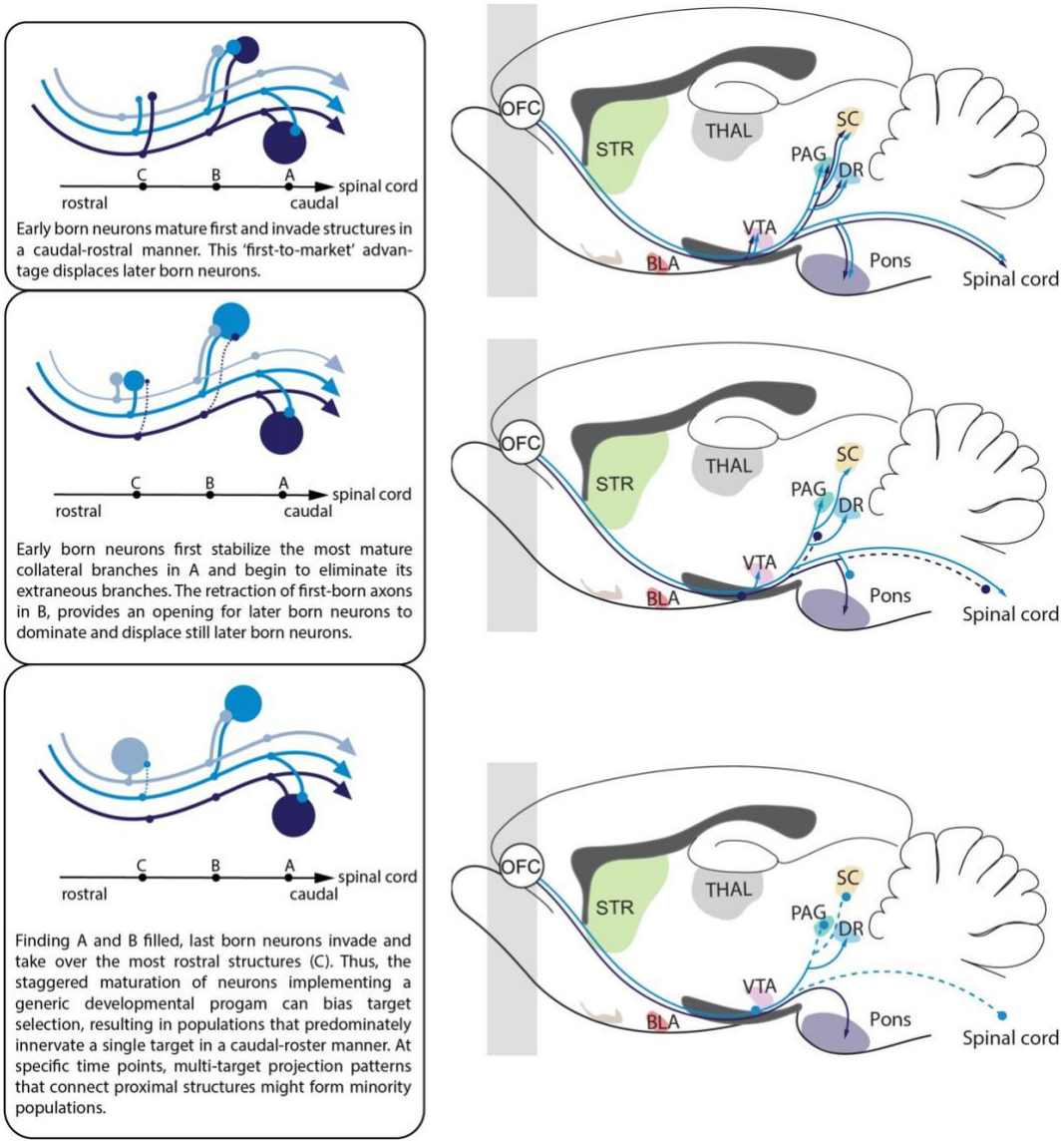


Figure 17 Proposed matching mechanism for subcortical target selection

neurons have the most time to stabilize in these caudal structures and begin eliminating immature branches in more rostral structures. If we then assume that a structure is only able to support a limited number of synapses, then these caudal structures become unavailable to further innervation. As a corollary, because all subcortical neurons adhere to the same caudal-to-rostral path, the elimination of immature branches by early born neurons provides an opportunity for later

born neurons to assume more territory in the next most caudal structure. Thus, we suggest that there are cascading developmental windows wherein neurons are biased to innervate the most caudal structure that is available. This represents a stochastic but ordered mechanism for the organization of distinct populations of single target subcerebral projection neurons into overlapping sublayers (Figure 17). In addition, this potentially also explains the preference of multi-target projections to innervate neighboring structures in orbitofrontal cortex. That is, neurons on the border of developmental windows might retain a handful of mature collaterals to nearby structures.

Evolutionary and functional implications for segregated subcortical outputs

In contrast to older evolutionary hypotheses (Fournier et al., 2015; Masterton et al., 2018; Reiner, 1991; Suryanarayana et al., 2017), it was recently proposed that the deep layer projection classes were a recent addition and inserted into the cortical framework (Shepherd and Rowe, 2017; Sherman and Usrey, 2021). Though it still debated how the six-layered mammalian cortex is derived from the three-layered reptilian cortex (Karten, 2013; Shepherd and Rowe, 2017), this hypothesis has intriguing functional implications. First, this hypothesis might underlie the subdivision of subcortical projection neurons into distinct transcriptomic types, versus the continuous transcriptomic gradient of intra-telencephalic projection neuron types (Yao et al., 2021; Zhang et al., 2021). Second, this hypothesis implies that a bottleneck layer which filters information from neocortex to subcortex is evolutionarily advantageous and selectively connects target subsets, rather than all possible targets. While recent transcriptomic studies have cast doubt on whether there are reptilian analogs of mammalian projection neuron types (Tosches et al.,

2018), the anatomical division-of-labor in rat orbitofrontal cortex is a structural pre-requisite for a subcortical sorting principle. Despite incredible representational diversity across single neuron responses in frontal cortex, anatomically distinct cell types in rat orbitofrontal cortex can be mapped to decision variables, such as decision confidence and previous trial outcome. In orbitofrontal cortex, these target-defined projection neurons are thought to serve as parallel streams of (occasionally redundant) information that mediate dissociable aspects of cognition.

Sublaminar organization of cell types beyond rat orbitofrontal cortex

In mouse motor cortex, there is a similar coarse relationship between laminar depth and subcortical projection patterns (Muñoz-Castañeda et al., 2020; Zhang et al., 2021). An analogous "deeper further" laminar structure is difficult to assess in mouse motor cortex because these subcortical projection neurons demonstrate a one-to-many logic. However, these subcortical projection neurons represent a heterogenous population of distinct transcriptomic and anatomic subtypes (as opposed to the continuous intra-telencephalic transcriptomic types), such as the cortico-spinal subtype localized in superficial L5b, and cortico-medullary subtype localized in deep L5b (Economo et al., 2018b). As cortico-spinal neurons have the longest axons that reach the end of the spinal cord during the first or second post-natal week of life in mice and rats (Gianino et al., 1999; Joosten et al., 1989; Stanfield, 1992), this suggests that distance-ordered sublayers might instead reflect a correlated variable, such as the caudal-to-rostral maturation of subcortical targets (Altman and Bayer, 1978, 1980, 1981; Bayer, 1980; Fentress et al., 1981; Finlay and Darlington, 1995). This higher order logic might be more disordered in mouse motor cortex because first, many circuit plans in rats tend to be more subtle in mice, for instance the strictness of the inside-out developmental rule, tonotopic organization of auditory cortex (personal communication), the

columnar segregation of thalamic inputs to barrel cortex (Rakic, 1974; Sato et al., 2007; Shepherd and Svoboda, 2005). Second, target specificity is markedly increased from the intra-telencephalic projection subtypes in L2/3 to the cortico-thalamic projection neurons in L6, this raises the possibility that there might be other global trends, for instance, a rostral-to-caudal gradient matching the thickening of cortex (Jensen and Altman, 1982). A loosening of organizational principles from frontal to motor cortex might imply that subcortical projections in prefrontal cortex are more tightly regulated. This could be due to either region-specific factors (Bhaduri et al., 2021; O'Leary and Sahara, 2008) or from a stochastic process such as temporal matching, wherein early developing regions have their first choice of synaptic targets and other later developing regions scramble to compete for a continuously depleting pool. Interestingly, single neuron activity in motor cortex is much more difficult to interpret and map to motor related parameters, an observation that might result from the more complex subcortical projection patterns (Churchland and Shenoy, 2007; Kita and Kita, 2012; Vyas et al., 2020).

Conclusions and outlook

In summary, we propose that the connectivity patterns of subcortical projection neurons are guided by a stochastic temporal matching process in development. Such a mechanism side steps the need for genetic programs that deterministically encode single neuron projection patterns, but does not exclude the possibility that there might be refined genetic subtypes (Arlotta et al., 2005; Lai et al., 2008; Sahni et al., 2021a, 2021b). Though in rat orbitofrontal cortex, anatomic projection types represent molecularly distinct subclasses, this might be the result of activity-dependent differences that arise much later in development and possibly reflect their distinct functional roles.

The gap in between circuit-driven systems and developmental neuroscience underscores a need for more studies at a cellular resolution. Early in development, it is generally thought that all subcerebral projection neurons are largely indistinguishable (O'Leary and Koester, 1993; O'Leary and Stanfield, 1985). Though time-dependent and region-specific factors play an important role in shaping cell fate (Bishop et al., 2000; McConnell, 1988; O'Leary and Sahara, 2008), precisely what triggers the selective elimination of collateral branches and when a cell type becomes immutable is unknown. The advent of single neuron resolution technologies has made it possible to revisit fundamental questions in development that shape cortical function.

Chapter 4

Mesolimbic dopamine encodes subjective value and predicts time investment decisions

Suelynn Ren[*1,2], Torben Ott[*1,3], Apoorva Arora[*1], Victoria Vega[1,3], Thiago Gouvêa, Adam Kepecs[1]

[1] Department of Neuroscience and Deparment of Psychiatry, Washington University in St Louis

[2] Medical Scientist Training Program, Washington University in St Louis

[3]Bernstein Center for Computational Neuroscience Berlin Humboldt University of Berlin *equal contribution

Abstract

Value is thought to drive many types of decisions. According to learning theory, value can be inferred from reinforcement history through an algorithmic computation ('model-inferred value'). In contrast, in behavioral economics, appropriately designed tasks can elicit choice patterns that reveal subjective preferences ('revealed choice value'). However, the relationship between these distinct measures of subjective value remains unclear, which limits studying its neural basis. Here, we designed a probabilistic reward learning task that brought together these two approaches. Rats chose between two options that were probabilistically 'baited' with rewards, varying across blocks. To earn a reward, rats committed to their choice by investing time for uncertain, delayed rewards. Choice value could be inferred from reward and choice history using generalized linear models or fitting reinforcement learning models. These model-inferred choice values robustly predicted rats' choice behavior, as well as the magnitude of time investment. In turn, shorter time investments predicted choosing the lower-valued ('model-inconsistent'), rather than the higher-valued ('model-consistent') option. Thus, time investment reflected the rats' subjective choice valuation, beyond model-inferred value alone, thereby behaviorally revealing choice value. To investigate the neural processes underlying choice valuation, we monitored dopamine release in the ventral striatum using fiber photometry with virally expressed, genetic dopamine sensors. We observed phasic dopamine release at the time of choice, which strongly predicted trial-by-trial time investment seconds in advance, but was not correlated with model-inferred choice value based on reward history. Thus, mesolimbic dopamine encodes the subjective valuation of choice options that can be behaviorally read-out in single-trial time investment decisions.

Highlights

- Time investment is tightly linked to traditional measures of model inferred value
- Striatal dopamine release at the time of choice predicts time investment, seconds in advance
- Time investment is a novel trial-by-trial behavioral read out of subjective value

Introduction

From picking dinner to college courses, the choices we make reflect our past experiences and individual preferences. The value of a choice is inherently subjective and can fluctuate from moment to moment, yet we need to identify objective measures to study its neural basis. Different approaches from learning theory and behavioral economics naturally capture distinct contributions to value (Padoa-Schioppa and Schoenbaum, 2015). According to learning theory, the values of various options can be inferred from reinforcement history and pitted against one another to determine an optimal choice (Sutton and Barto, 2018). In the real world, however, many choices, like apples and oranges, cannot be directly compared and there is no "right" answer. Capturing such subjective preferences is central to behavioral economics and can be revealed from choice patterns ('revealed choice value') (McFadden, 2001; Padoa-Schioppa, 2011). Yet, this approach lacks a rigorous framework that can be used to fractionate underlying choice processes. A unified approach that bridges these distinct measures of value is needed to better understand ongoing choice behavior.

Within the brain, dopamine is thought to broadcast representations of value critical for learning and motivated behavior (Berridge and Robinson, 1998; Hamid et al., 2016; Schultz et al., 1997, 2015; Tobler et al., 2005; Wyvell and Berridge, 2000). Prior to outcome, the cue-evoked dopamine response invigorates approach behavior and reflects anticipated value (Burke et al., 2018; Flagel et al., 2007; Mohebi et al., 2019), considering varied dimensions from outcome magnitude and timing (Hamid et al., 2016; Howe et al., 2013; Lak et al., 2014a; Roesch et al., 2009; Saddoris et al., 2015; Wei et al., 2021) to homeostatic state (Cannon et al., 2004; Stouffer et al., 2015). Such findings have bolstered a role for dopamine in signaling subjective value, but its relationship to choice remains unclear. For instance, dopamine release in the nucleus accumbens has been shown to signal the highest valued option, rather than the chosen option for effort related choices (Hollon et al., 2014). Selection of the seemingly less valuable option has numerous possible interpretations from an impulsive choice (Freels et al., 2020; Gao et al.; Puumala and Sirviö, 1998; Zeeb et al., 2009) to information seeking one (Bromberg-Martin and Hikosaka, 2009; Fink and Smith, 1980). A single trial behavioral readout that reflects one's commitment to a choice without altering its prospective value could help clarify the role of dopamine.

Like money, time is a limited and valuable resource. We previously demonstrated that humans, rats, and mice strategically invest their time to obtain uncertain future rewards in proportion to their confidence in receiving a pay out (Ott et al.). We therefore reasoned that single trial time investment decisions could be used as a behavioral read out of subjective value. Here, we adapted a classic two-armed bandit task (Lau and Glimcher, 2005) and brought together two distinct measures of value from learning theory and behavioral economics. As expected, established models robustly predicted rats' choice behavior. These model-inferred choice values further

predicted the magnitude of post-choice time investment decisions such that lower-valued choices yielded shorter and high-valued choices yielded longer time investments. To investigate the neural processes underlying choice valuation, we also monitored dopamine release in the ventral striatum using fiber photometry with virally expressed, genetic dopamine sensors (Patriarchi et al., 2018, 2020). We observed phasic dopamine release at the time of choice, which strongly predicted trial-by-trial time investment seconds in advance, but was not correlated with model-inferred choice value based on reinforcement history. Thus, mesolimbic dopamine encodes the subjective valuation of choice options that can be behaviorally read-out in single trial time investment decisions.

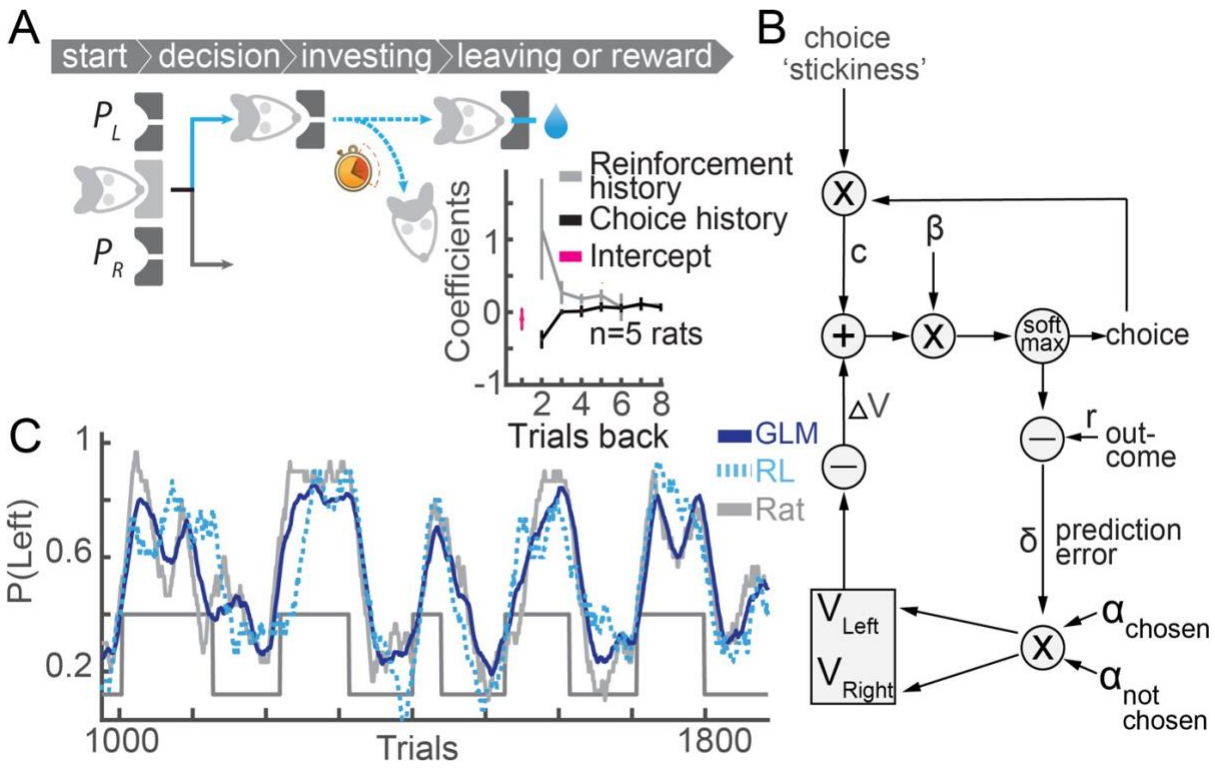


Figure 18 Single-trial time investment reflects subjective value and is predicted by mesolimbic dopamine.

(A) 2-armed bandit task in which rats commit to a choice and then invest their time. (B) Reinforcement learning (RL) model that learns choice values. (C) Observed choice behavior across interleaved blocks for a single rat and corresponding fits for the linear and RL model. (D) Choice values inferred from both the linear and RL models predicted the rat's choice behavior.

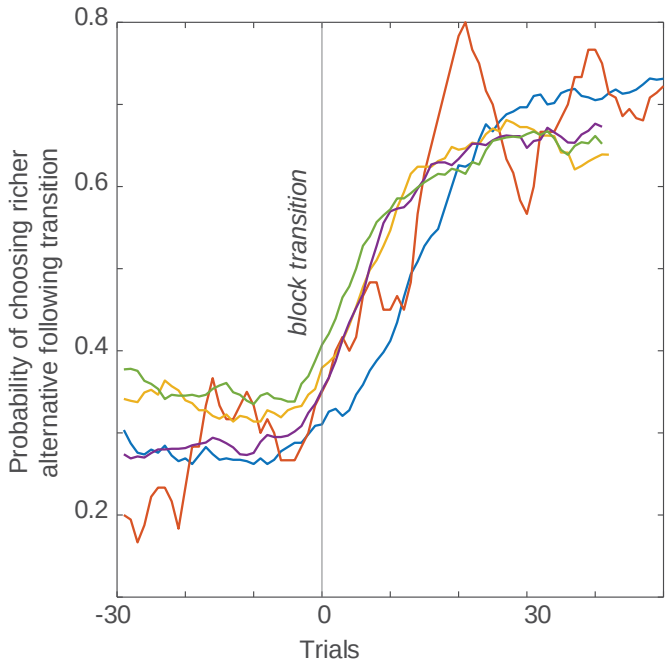


Figure 19 Rats' choice behavior before and after a block transition

Individual rats' ongoing behavior is shown in different colors, pre- and post- block transition.

Results

In a novel variant of the classic two-armed bandit task, rats (n = 5) made choices (left or right) and invested time in their choice to obtain an uncertain, delayed reward (Figure 18A). In each trial, water reward became available at a choice port with a fixed probability and stayed available ('baited') until it was collected. Reward probabilities for each port (p_{high} = 0.4, p_{low} = 0.1) were alternated across short, un-cued blocks (random length, ~75-150 trials). After choosing a baited choice port, reward was delivered after a random delay, sampled from a truncated exponential distribution (0.5 s-8 s, time constant 1.5 s). To encourage the efficient use of time, we used a long

inter-trial-interval that was adjusted for individual rats (4-10 s). The rats' willingness to wait for a reward provided a graded, single-trial time investment in their choice, which we hypothesized reflected subjective choice value. To assess the relationship between time investment and choice value, we first inferred trial-by-trial choice value ('model-inferred value') using nested reinforcement learning (RL) models and estimated parameters via maximum-likelihood fitting (Figure 20) (Delgado et al., 2011; Sutton and Barto, 2018). Cross-validated choice data, as well as model-selection criteria (AIC and BIC), were best predicted by a RL model with an average pseudo-R^2 of 0.38 (SD = 0.21) and included the parameters learning rate (α), inverse temperature of a soft-max decision rule (β), a 'stickiness' or perseveration parameter that captured short-term tendencies to repeat or alternate choices, and a forgetting parameter ($\alpha_{\text{not chosen}}$) that simulated the decay of the unchosen option's value (Figure 18, 20). To test the robustness of these RL model-inferred choice values, we also used a linear regression approach that has been previously described (Lau and Glimcher, 2005) and considered the effects of past rewards and choices (Figure 18A, inset) (1). Both the RL and linear models captured the rats' choice behavior and yielded similar choice value estimates on a trial-by-trial basis (Figure 18C, D). Further, the rats' ongoing choice behavior is like that of non-human primates, with the linear model coefficients and choices after a transition demonstrating similar patterns across species (Figure 18, Figure 19).

Using these model-inferred choice values, we then identified trials in which the rats' choice violated model-inferred value predictions, that is, trials in which rats chose the lower valued (i.e., 'model-negative) rather than the higher valued (i.e., 'model-positive) option. Rats invested time in proportion to model-inferred value for model-consistent choices, but inversely for model-inconsistent choices, i.e., time investment predicted the percentage of model-consistent choices

(Figure 21A-C). In other words, time investment reflected the rats' degree of violating model-inferred choice value, arguing that time investment provides a scalar, trial-by-trial behavioral measure of subjective value ('revealed choice value').

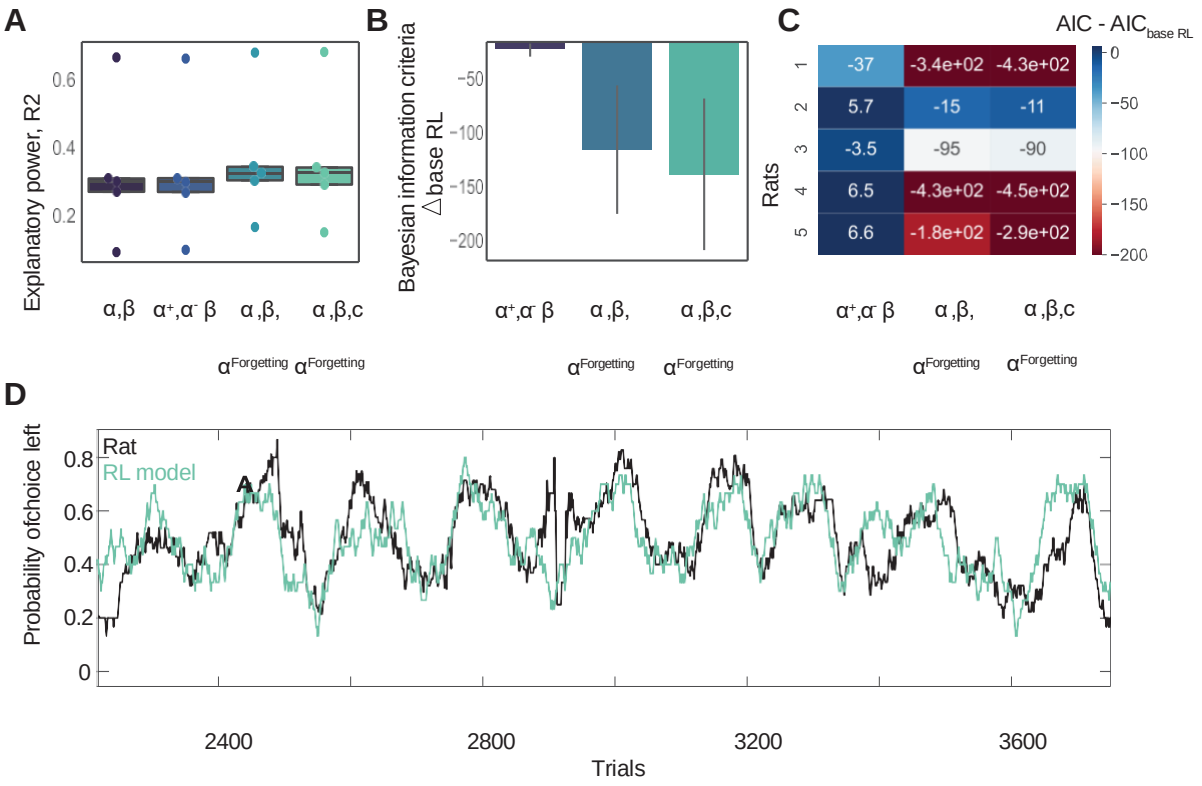


Figure 20 Reinforcement learning model selection

Nested reinforcement models were compared across various metric including (A) explanatory power (B) Bayesian information criterion, and (C) Akaike's Information Criterion. (D) The rats' choices were smoothed using a moving mean and a window of thirty trials (black). The four parameter (α, β, c, αForgetting) reinforcement learning model closely captures choice behavior.

We next probed the relationship between dopamine release, model-inferred choice value, and time investment. We recorded dopamine release via a genetically-encoded fluorescent sensor in the ventral striatum in rats (n = 6) chronically implanted with optical fibers. To correct for motion artifacts in freely moving rats, we co-recorded from a neural activity-independent red fluorophore, thus providing a reference channel to de-trend the dopamine signal (Schmack et al., 2021). To

correct for slow drifts in fluorescence across a session (e.g., bleaching), baseline fluorescent signal was calculated from the 1-minute moving median and subtracted from the raw data to yield a normalized fluorescence signal, dF/F. As expected, reward evoked strong dopaminergic responses in the ventral striatum (Figure 21D). At the time of choice, however, dopamine predicted single-trial time investment, seconds in advance, rather than reflecting model-inferred choice value (Figure 21F, G). These results suggest that ventral striatal dopamine encodes subjective choice value that informs future investment decisions, a hallmark of economic decision-making.

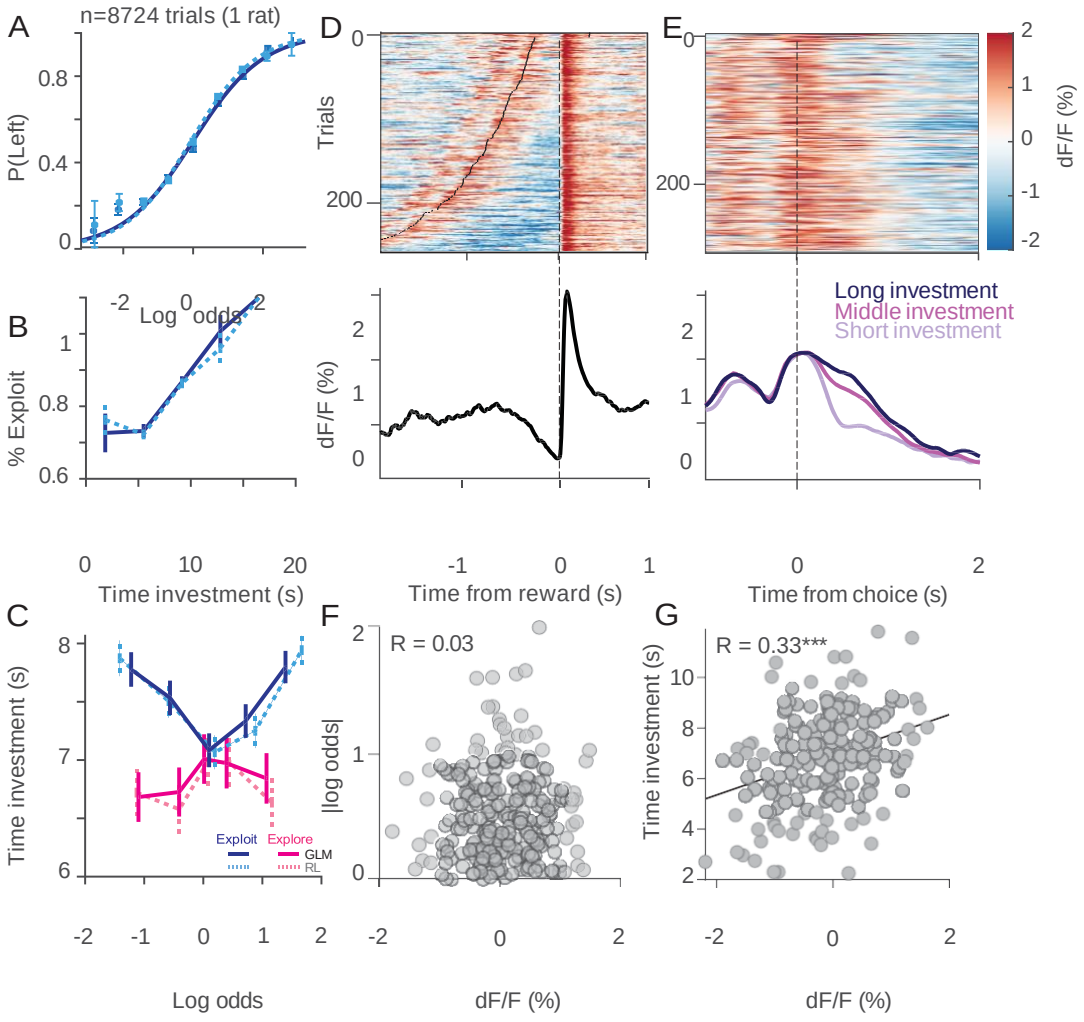


Figure 21 Time investment is a single trial behavioral measure of revealed choice that is predicted by striatal

dopamine.

(A) Choice values inferred from both the linear and RL models predicted the rat's choice behavior. (B) Time allocation increased in proportion to the fraction of trials in which the higher-valued ('model-consistent') option was chosen. (C) Time investment strategies showed opposite patterns in model-consistent (pink; m^{pos}) versus model-inconsistent (blue; m^{neg}) trials for both linear (solid) and RL (dotted) derived choice values. (D) Dopamine release in the ventral striatum is strongly evoked by reward. (E) Dopamine release at the time of choice predicts time investment seconds in advance. (F) Dopamine release is not correlated with model-inferred value. (G) Dopamine release is strongly correlated to single-trial time investment.

As striatal dopamine predicted subjective value better than reinforcement history, we next sought to determine what information is captured in time investment decisions, but not model-inferred value. First, we regressed out the effect of model-inferred value on time investment by comparing the residuals from time investment and log odds and the residuals from model-inferred value and choice. For simplicity, we took all left choices as a reference such that by design, a log-odds positive, left choice is model positive and a log-odds negative, left choice is model negative. We find a significant inverse correlation specifically for model-negative choices, choices that were opposite to the model prediction. Rats "under-stay" or waited for a shorter time for left choices, when the model predicted a higher valued right choice. As both models largely capture choice patterns, the seemingly less valued option might be chosen because of an additional process, such as fluctuating satiation or internal noise. Such an evolving process might be described as a

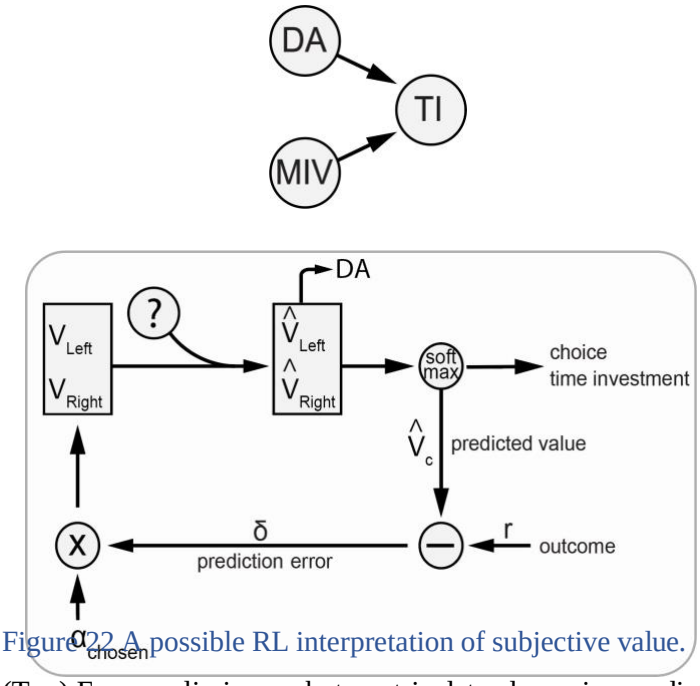


Figure 22 A possible RL interpretation of subjective value. (Top) From preliminary photometric data, dopamine predicts time investment, but not model inferred value. (Bottom) This could be explained by a time-dependent noise process such as a random walk.

random walk that could be incorporated into the reinforcement learning model (Figure 22).

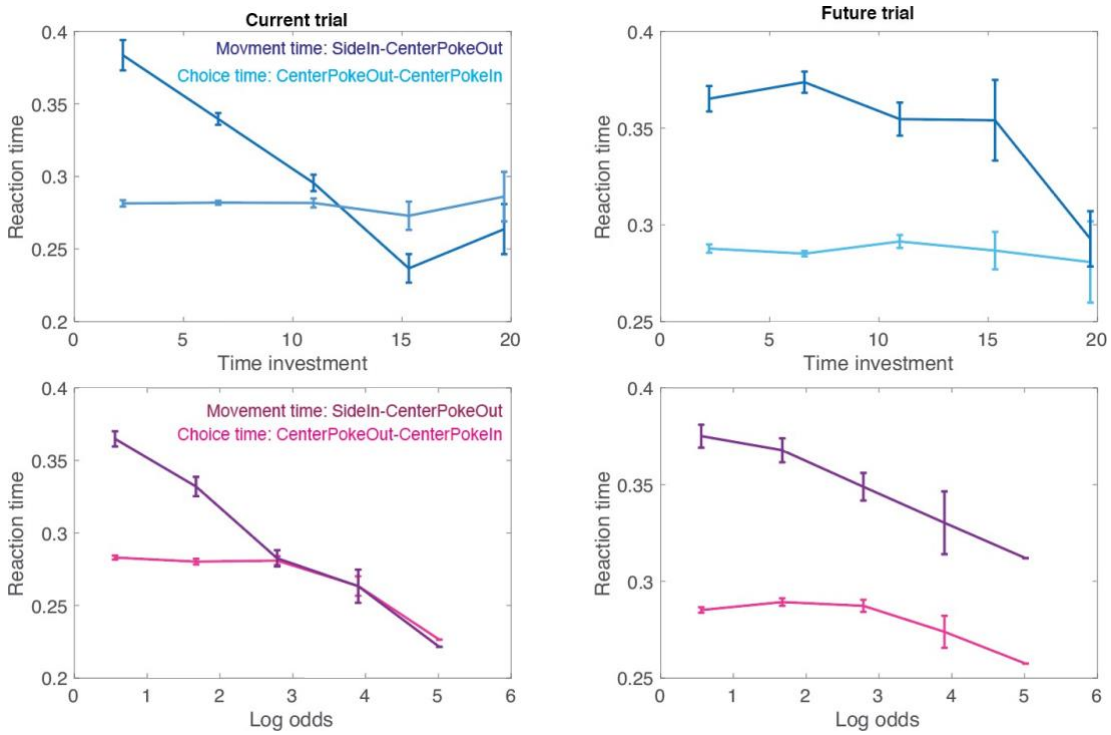


Figure 23 Both time investment and model-inferred value relate to reaction time.

(Left) The movement and choice time on the current trial plotted against the time investment (Top) and model inferred value (Bottom). (Right) Future movement and choice time plotted against time investment (Top) and model inferred value (Bottom).

Next, because dopamine is thought to play a role in influencing motivation and invigorating approach behavior, we explored the relationship between single trial time investments and different types of reaction times. We focused on: (1) choice time, from 'center port in' to 'center port out' (2) movement time, from 'center port out' to 'choice port in', and (3) latency time, from the end of the last trial to 'center port in' of the current trial. We find that both time investment and model inferred value have markedly similar relationships to choice and movement times, such that rats are much faster in high expected value trials and high time investment trials. This is mainly an effect on movement time, as there is no significant change across choice times for differently valued trials (Figure 23). While this is consistent with a role of dopamine in invigorating approach

(du Hoffmann and Nicola, 2014; Shadmehr et al., 2019), these results suggest that this effect can be explained from model-inferred value. Looking for a different estimate of motivation, we therefore turned to latency time, the time that it takes for a rat to initiate the next trial. In contrast to the reaction times that are calculated when the rat is engaged in the task, single trial time investments and model inferred value diverge for latency times (Figure 24). Though there is no relationship with model-inferred values, a longer trial latency typically correlates with longer single trial time investments. Such a finding is consistent with the possibility that the rat is willing to spend its time in periods when the time is an abundant resource (e.g., less valuable), either because the rat is sated or because the global reward rate is lower. Future analysis will work towards clarifying precisely what information might be captured by time investment, but not model inferred value.

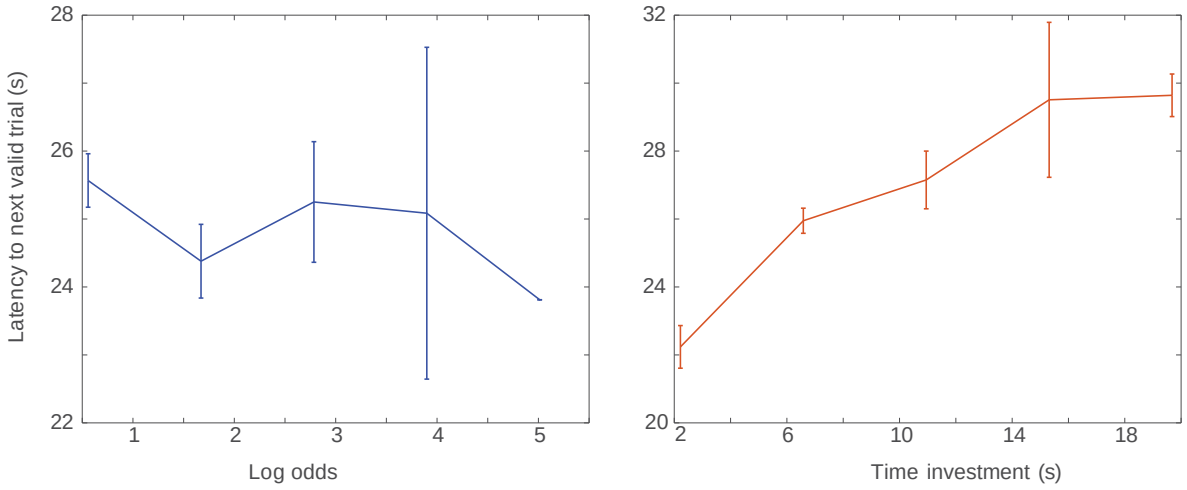


Figure 24 Time investment predicts latency to initiate next trial.

(left) Latency to next trial plotted against model-inferred value and (right) single trial time investments.

Methods

Animals

Adult male Long Evans rats (~300–500 g) were used for the study (Envigo). Rats were group-housed and maintained on a reverse 12 h light/dark cycle. All procedures were carried out in accordance with National Institutes of Health standards and were approved by the Cold Spring Harbor Laboratory Institutional Animal Care and Use Committee and Washington University in St Louis.

Photometry experiment and analysis

Rats were anesthetized with 1-3% isoflurane, injected with pre-operative buprenorphine (0.03 mg/kg) and placed in a Kopf stereotactic setup. For injections, we used pulled 5 uL calibrated glass micropipettes, cut to an opening diameter of ~10-20 um. Craniotomies were made with a dental drill and dura cut with a 27 gauge needle. In post-operative recovery, rats were provided with oral carprofen tablets and observed either daily for a minimum of 7 days, or to euthanasia.

We used a viral approach to express the genetically encoded optical dopamine sensor, GRABDA. Rats (n=6) were injected with a 300 nl of 1:1 AAV9-hSyn-DA4.3 (~5 x 1012 titre) and AAV1-CAG-tdTomato (~1 x 1011 titre) in the ventral striatum (1.6AP; 1.75ML; -6.9 and -7.1DV). During the same surgery optical fibers (400um core) were inserted at -6.9DV and cemented in place. Data were collected >three weeks later, to allow for GRABDA expression.

We used a custom dual color fiber photometry acquisition set-up to simultaneously record dopamine activity from green fluorescent GRABDA and a control fluorophore tdTomato. The

excitation beams for the green and red channels were produced by a 470 nm and 565 nm LED light source, and launched into a 400 um core, 0.48 NA fiber patch cable. Fluorescence and excitation and detection were performed through a multi-modal optical fiber. To ensure proper separation of the green and red channel, the flurescence signals were amplitude modulate dby sinusoidally varying command voltage of the LED driver with two different frequencies (531 Hx and 211 Hz) and demodulated prior to data processing. Photometry data were aquired using a data acquisition card (PC!e-6321 National Instruments) and synchronized with the behavioral task using Bpod and custom software written in Matlab.

Fluorescence signals were expressed as the relative fluorescence, normalized to the 1-2 second baseline period in which there is no task specified movement. In addition, the red control fluorophore which is not tied to neural activity serves as an estimate of movement or cable induced artifacts. To correct for these dopamine independent changes in fluorescence, we subtracted the shared variance across the red and green channels.

Behavior and training

A rectangular behavioral box contained three ports aligned on one wall which were equipped with LEDs, infrared photodiodes, and phototransistors. Interruption of the infrared photobeam was used to determine port entries and exits. The two side ports (left and right choice ports) were equipped with valve controlled water spouts for reward delivery. Water valves, LEDs and phototransistors were controlled by the behavioral measurement system Bpod and Pulse Pal.

Five naive rats were first trained to self-initiate trials via holding their snouts in the center port for a minimum of 0.4 seconds. Rats were then able to retrieve a water reward from either the left or right port. Ports were illuminated to signal availability and phases of the trial, such that at the start of a trial, only the center port was illuminated. After trial initiation, both the left and right ports were illuminated. An early withdrawal from the center poke was punished with a 3 second time-out and in this punishment phase, no ports were illuminated. After a 1 week period of habituation and trial initiation practice, rats were tested on the foraging task. Reinforcers were independently drawn for the right and left ports with independent probabilities that were assigned in a blockwise manner. As such, on any given trial, it is possible that both left and right choices might be rewarded, neither might be rewarded, or only a single choice might be rewarded. If a reward was "armed" for the unchosen option, it remained available (i.e., baited) until that option was chosen. Block transitions were not signaled and a changeover delay or response was not used. The block length was drawn from a normal distribution with a mean of 120 trials and a variance of 10 trials.

After stable performance levels (visual inspection of matching behavior), we gradually introduced randomly delayed rewards. For correct choices, we withheld reward delivery for a random time between 0.6 - 8 seconds, drawn from a truncated exponential distribution with a time constant of 1.5 seconds. Rats had to maintain their snout or keep poking in the choice port during the entire time point; to avoid false detections of leaving decisions, a grace period of 0.3 seconds was set such that if the rat exited and re-entered the choice port within 0.3 seconds, it did not constitute a leaving decision. There was no feedback on non-rewarded trials. Further, we gradually increased the inter-trial interval to 4-15 seconds for each rat in order to lower the average reward rate and in turn the opportunity cost, thereby reinforcing time as a commodity.

This approach provided us with a continuous time investment as well as a binary choice on each non-rewarded trial. Each session lasted about 2-3 hours, in which rats typically performed 300-600 trials with at least 3 block transitions.

Analysis of Behavioral Data

We analyzed 41,151 trials from five rats across 95 sessions. We excluded sessions where rats performed less than 100 trials. For estimates of mode-inferred value, we split the behavioral data into training (even sessions) and test (odd sessions) sets. Log odds reflects the strength of choice evidence which is explicitly formulated in the logistic regression and determined as the softmax of left and right values in the reinforcement learning model. Model consistent refers to choices that were predicted by the model, whereas model inconsistent refers to choices that were not predicted by the model.

Measures of model-inferred value: logistic regression

First, we fit steady state choice behavior using a response-by-response model that considers the effects of past reinforcers and choices as separate linear processes, as has been described elsewhere. Briefly, the following logistic regressions has been shown to capture ongoing choice behavior,

Where r is the reinforcement (1 if rewarded or 0 if un-rewarded) and c is the choice. Note that due to the formulation, 1 is a left choice and -1 is a right choice. The alpha and beta coefficients measure the influence of past reinforcers and choices, and the intercept term gamma captures left

or right bias that is not attributable to reinforcement or choice history. We fit this model to the training data using the fitglm function in Matlab. To select the best candidate model, that is the combination of lengths of reinforcement and choice histories, we used Akaike's Information Criterion (AIC). We computed AIC values for a family of models in which we varied the lengths of reinforcer and choice histories, up to 50 trials each. Importantly, when multiple candidate models were similar, that is, when relative differences between AIC values were less than 2, we selected the most parsimonious model.

Measures of model-inferred value: reinforcement learning

We next fit choice data with a standard reinforcement learning model (Delgado et al., 2011; Sutton and Barto, 2018). The model uses a sequence of choices and outcomes to estimate the value of each option for every trial. The expected values were initialized at 0, and after each trial, the value of the chosen option was updated according to the following rule. We compared several reinforcement learning model variants, but the base model consisted of two parameters: the learning rate (alpha) and temperature (beta) that controls choice stochasticity. We then iteratively added parameters that have been classically used to describe matching behavior including c, a choice stickiness parameter that modulates how often a choice is repeated (positive) or alternated (negative), a negative learning rate that alters updating for an un-rewarded choice, and a forgetting term that decays the value of the unchosen option. The data likelihood was taken as the product of the log choice probabilities and free parameters were estimated by maximum likelihood using the fminsearch function in Matlab. Parameters were constrained such that alpha, alpha forgetting, and alpha negative parameters were in the range of 0 to 1 and beta in the range of 0 to 20. We used several different measures to assess fit on the test set including: explanatory power, Bayesian

Information Criterion (BIC) and Akaike's Information Criterion (AIC). Though the final selected model was selected on the basis of the relative BIC and AIC scores, it should be noted that adding these parameters did not significantly improve fit as assessed via a likelihood ratio test. This is presumably because most of the predictive power is from reinforcement history.

Chapter 5

A novel task to probe the role projection types in meta-learning

Introduction

Once thought to be a homogenous population, ventral tegmental area dopamine neurons are increasingly being subdivided into discrete subtypes that receive differential input (Beier et al., 2015b, 2019; Saunders et al., 2018; Watabe-Uchida et al., 2012). As such, subtypes likely receive information that is unique and subserves a circuit-specific function, though principles governing cortical-subcortical information transfer remain poorly understood (Han et al., 2018b; Kebschull et al., 2016b). As with other cortical areas, the orbitofrontal cortex is composed of projection neurons that target diverse subcortical structures (Chandler et al., 2013; Gabbott et al., 2005b; Ohara et al., 2003; Takahashi et al., 2009b). However, orbitofrontal cortex projection neurons appear to be segregated by their target structure (Chapter 2). This strict anatomical separation potentially reflects a functional division-of-labor, serving as a means to selectively route

information. The orbitofrontal cortex is reciprocally connected to the ventral tegmental area and is believed to play a role in relaying abstract and subjective value information (Gabbott et al., 2005b; Takahashi et al., 2009b). Remarkably, though the orbitofrontal cortex sends sparse, direct projections to the ventral tegmental area, precisely what information is routed, and for what purpose, is an open question (Takahashi et al., 2009b). In turn, ventral tegmental area dopamine neurons send dense afferent input to the orbitofrontal cortex, but precisely identifying the impact of dopaminergic modulation and reconciling it with both normal and pathological choice behavior, requires a mechanistic understanding of decision-making circuitry (Chandler et al., 2013; Ohara et al., 2003).

Humans and animals must make decisions and adapt their behavior in an ever-changing and uncertain world. For instance, a sudden change in a previously stable environment might be an indication of new information that requires immediate action. On the other hand, one should not make rash decisions based on transient fluctuations in an unstable environment. Indeed, uncertainty has long been hypothesized to drive learning by modulating the degree of updating. However, it is unknown if animals track the second order statistics of their environments or, how this information affects choice behavior (Dayan and Yu, 2002). Ventral tegmental dopamine neurons are thought to integrate such uncertainty information, modulating their firing rate based on confidence estimates and risk (Fiorillo, 2011; Jo et al., 2018; Lak et al., 2020; Naudé et al., 2016; Stauffer et al., 2016). It is largely unknown, however, how such information is routed to the ventral tegmental area dopamine neurons. The orbitofrontal cortex is a strong candidate for such information (Cromwell et al., 2018; Grabenhorst et al., 2019; Lak et al., 2014b; Masset et al., 2020; Padoa-Schioppa, 2009; Padoa-Schioppa and Assad, 2006b; Pastor-Bernier et al., 2021).

Single neuron responses in orbitofrontal cortex are reported to reflect the statistical parameters of uncertain outcomes, including mean value and variance (O'Neill and Schultz, 2010). Intrigued by such findings in the literature and by our own data that these subcortically projecting neurons are highly dedicated, we hypothesized that the ventral tegmental projection neurons might route uncertainty information.

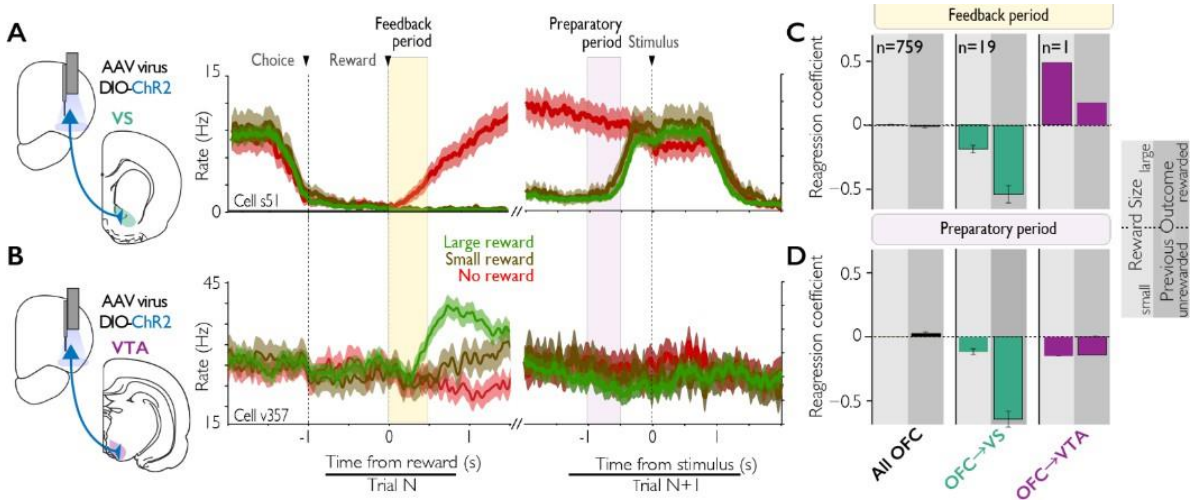


Figure 25 Optogenetic tagging of VS and VTA-projecting OFC neurons.

Optogenetically identified tetrode recordings form OFC-VS **(A)** and OFC-VTA **(B)** neurons were made using the dual-virus approach to deliver ChR2. Average activity of an OFC-VS neuron following reward delivery (feedback period) and in the leadup to the subsequent trial (preparatory period). OFC-VS neurons reliably show strong representations of negative outcomes. Average activity of an OFC-VTA neuron, showing graded representations of reward size. Unlike OFC-VS neurons, OFC-VTA representations are not sustained through the subsequent trial. **(C,D)** Average regression coefficients for Reward Size Preference and Previous Outcome Preference (equivalent to behavioral bias measures in Fig. 2) for all OFC neurons, OFC-VS, and OFC-VS projection neurons. *VS, ventral striatum; VTA, ventral tegmental area.*

Methods and results

Our work support the hypothesis that different projection populations within orbitofrontal cortex encode different forms of value information. Preliminary recordings from optogenetically tagged projection neurons to the striatum (n=19) demonstrate high sustained firing following unrewarded trials (Figure 25). In contrast, subcortically projecting neurons to the ventral tegmental area (n=1) demonstrate transient firing that is positive and graded, reflecting reward size. The ventral tegmental projection neuron responses did not reflect perceptual uncertainty, but their sensitivity to outcome uncertainty remains an open question.

To assess this possibility, we developed a novel reward biased, perceptual decision-making task in rats which were required to track the second order statistics of their environment (Figure 26). The core design is a two-alternative forced choice paradigm in which rats make a binary decision based on sensory (auditory) information. Rats must determine which of two auditory click trains delivered binaurally has the greater number of clicks. Rats perform close to perfect on the easiest trials and their accuracy varies with the strength of sensory evidence and correct choices are rewarded with water after a fixed reward delay. As opposed to a free choice task, such a perceptual task provides a normative framework, constraining choice behavior. The left and right reward sizes for a single trial are drawn from independent normal distributions and change in a block-wise manner. One port is changeable, whereas the other port is static. In an “unequal” block, the distribution for the static port has a small mean and variance. The distribution for the changeable port can either be a large mean and variance, or a large mean and small variance. At an un-signaled

time point, the block transitions to an “equal” or control block wherein the changeable port’s distribution matches that of the static port. Such a task design has an intuitive prediction for the speed of transition; it is easier to detect a change in a more stable versus a stochastic environment.

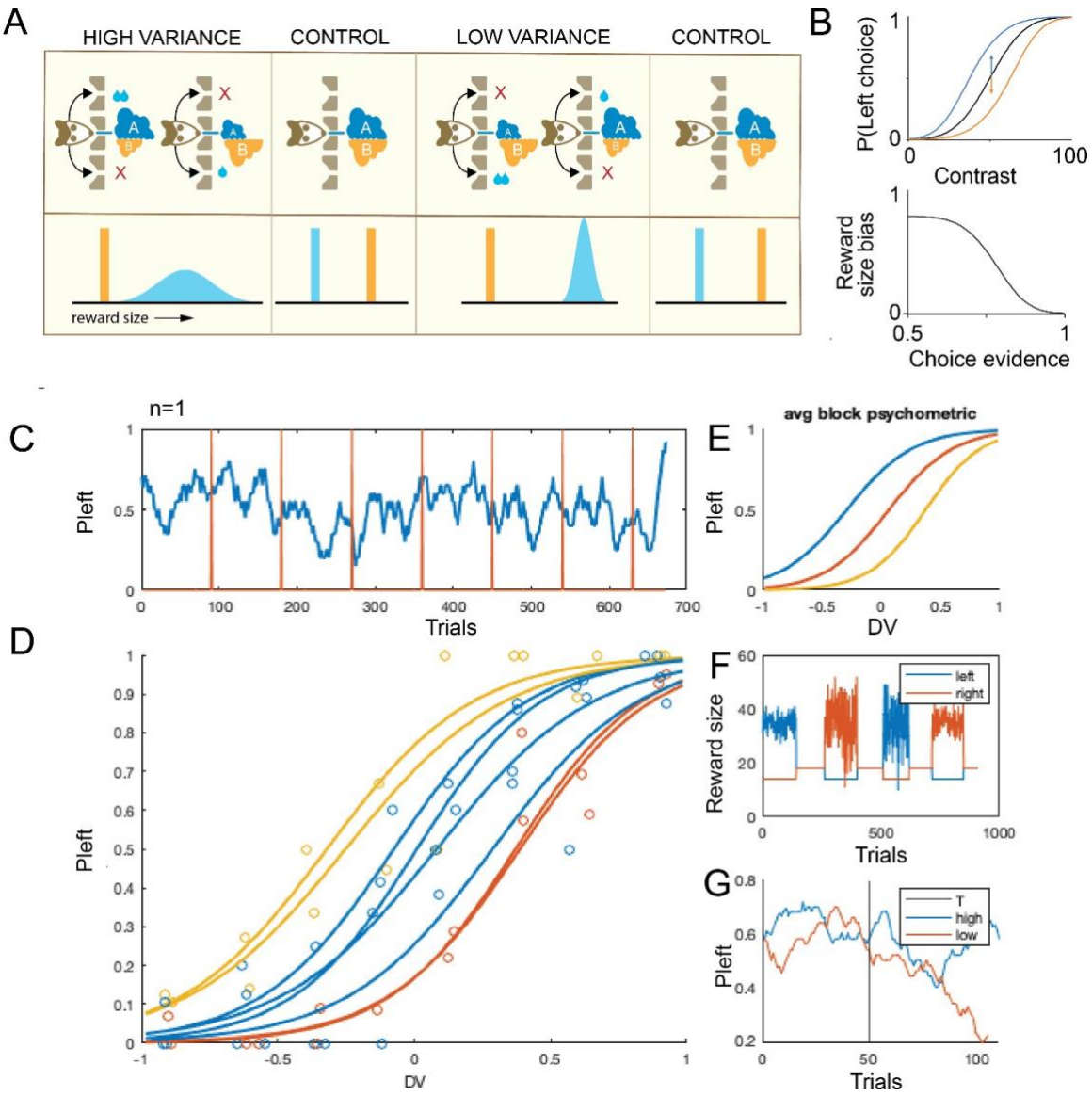


Figure 26 Variance task design.

(A) Rats are trained on a perceptual two-alternative forced choice task in which the reward size of the "changeable" port is drawn from a normal distribution with a low to high variance. In this pilot form of the task, the reward size for the static port is fixed and the mean reward sizes flip across blocks. **(B)** An adaptive shift in the psychometric towards the higher valued port (blue, left high; orange right high) was initially used to assess speed of learning following an unsigned block transition. **(C)** Overview of choice behavior in a pilot run for n=1 rat. **(D)** For initial testing, equal magnitude blocks were included so ensure that the rat's behavior returned to baseline. Including variance in the reward sizes did not disrupt choice behavior, and the shape and lapse rate of the variance (yellow, left high; red, right high) versus control blocks (blue) were comparable. **(E)** Average psychometric curves averaged across blocks. **(F)** Reward sizes per trial across blocks for the left (blue) and right (red) port. **(G)** Choice behavior 50 trials pre- and post- block transition.

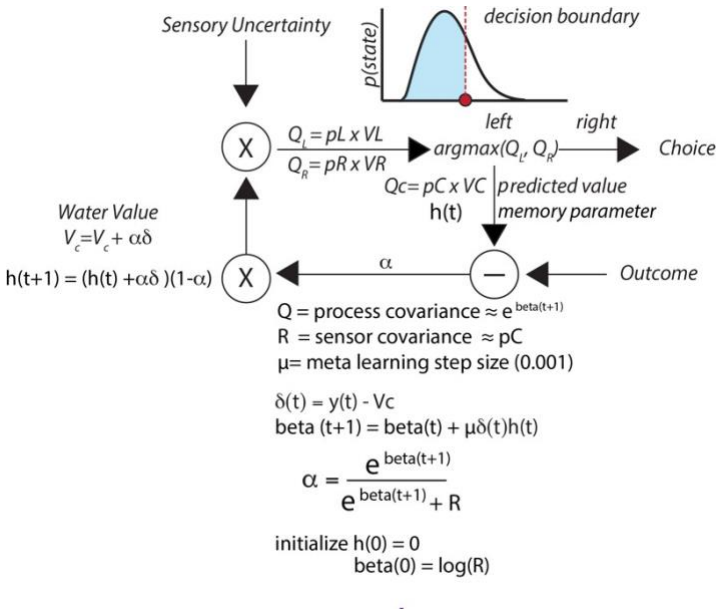


Figure 27 K1 rule for a dynamic learning rate in a reinforcement learning model.

To make behavioral predictions from first principles, we made different reinforcement learning models, one with a static and the other a dynamic learning rate (Figure 27). In the dynamic learning rate model, a Kalman filter is used to adjust learning rate based on outcome uncertainty such that in more variable environments (i.e., high mean, high variance distribution) the learning rate is low and in more static environments, the learning rate is high. In this dynamic learning rate model, the Q values for the left and right ports reflect integrated value, a combination of sensory uncertainty and value. The learning rate is then updated using a Kalman filter implemented through the K1 rule (Sutton, 1992). The K1 rule describes a dynamic programming form of stochastic gradient descent which boosts the amount of learning that happens if the current weight change is correlated with recent weight changes. This model confirms our intuition that faster transitions take place when a change occurs in a stable environment versus a more stochastic one (Figure 28). In contrast,

if rats do not adjust their learning rate, there should be no differences in the speed of transition across different environments.

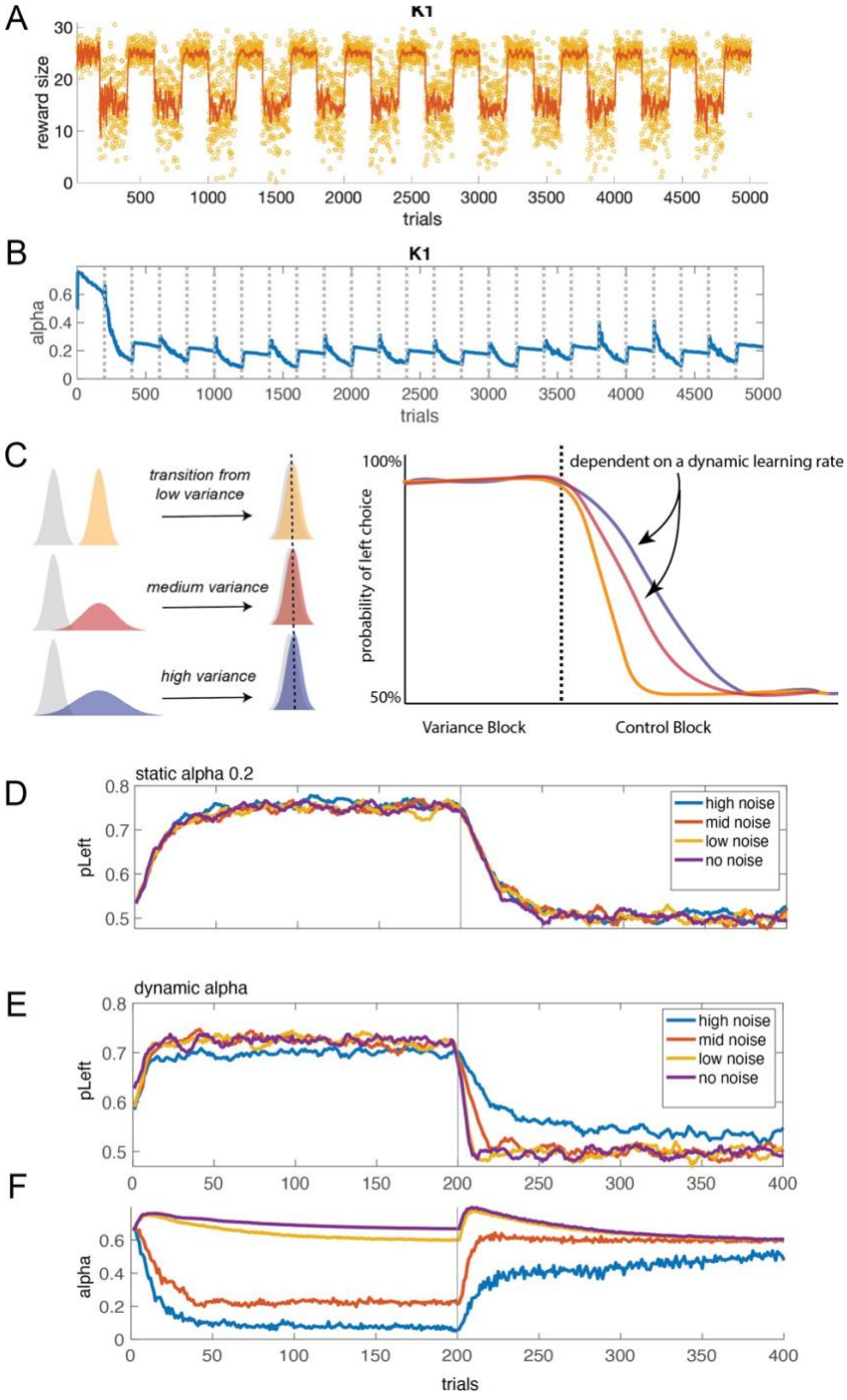


Figure 28 K1 RL model predictions

(A) Behavior of the K1 RL model in an environment with varying reward sizes, but no perceptual uncertainty. **(B)** After initialization, the learning rate remains relatively stable in low variance environments and then rapidly decreases in high variance environments. **(C)** Possible block transitions in the variance task and prediction that the speed of learning depends on a dynamic learning rate. **(D)** A RL model with a static learning rate predicts no differences in the speed of learning across different block transitions. **(E)** K1 RL model reproduces the intuition that speed of learning is slower in a high versus low variance environment. **(F)** Read outs of the learning rate on a single trial basis suggest that this is due to low learning rates in high variance environments.

Preliminary behavioral data supports the hypothesis that rats (n=1) adjust their learning rate based on outcome uncertainty. In this pilot study, there is zero variance in the control block. Despite added outcome variance, this did not disrupt choice behavior and the psychometric curve retained its shape and lapse rate across the different blocks. To assess choice behavior at the time of a block transition, we specifically looked at choices for difficult trials (Figure 29, top). Rats were slower to adjust their choice behavior in uncertain trials when switching from a high versus a low variance environment, an observation that follows from basic statistical principles. Interestingly, we also observed an effect on reaction time specifically for difficult trials (Figure 29, bottom). In control blocks, reaction time is modulated by trial difficulty such that rats spend more time in the center port for difficult versus easy choices. Surprisingly, this effect vanishes in a high variance environment and might also be a behavioral signature of the block transition.

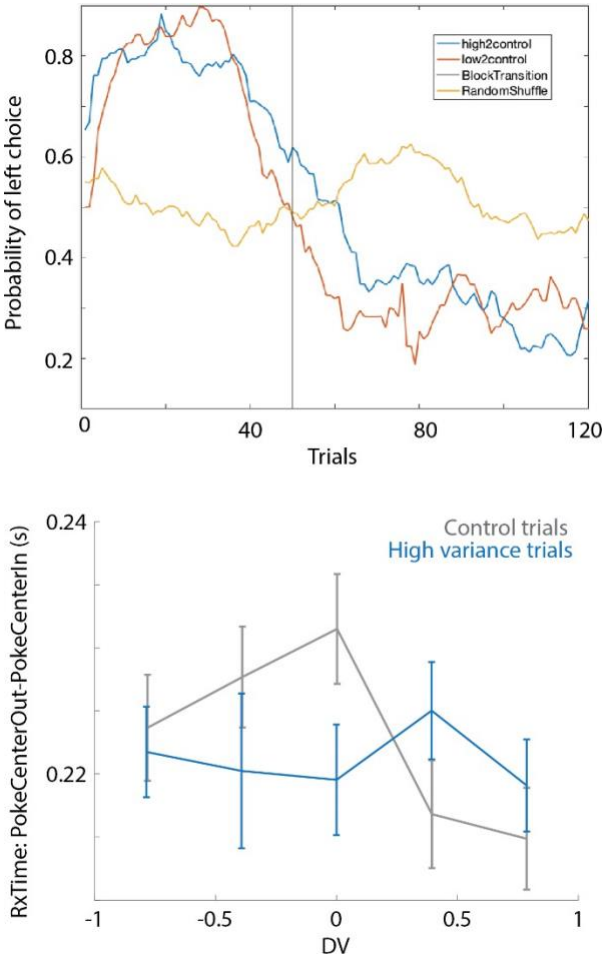


Figure 29 Preliminary behavioral data at the time of a block transition.

(Top) Slower speed of learning when moving from a high variance environment to the control. **(**Bottom**)** Blunted increase in reaction time for difficult trials in high variance environments.

To capture a more definitive readout of learning the block transition, we propose adding an additional "switch port" (Figure 30). Inferring the amount of evidence needed to recognize a transition and mapping these behaviors back to neural data is challenging as there are limited trials. As such, we adapted the task design to include a self-report of the block transition. After the block transition, the rat can re-initialize the unequal block by poking into a fourth "switch" port. As the "equal" block has a lower average reward rate, the rat should be incentivized to return to the "unequal" higher reward environment.

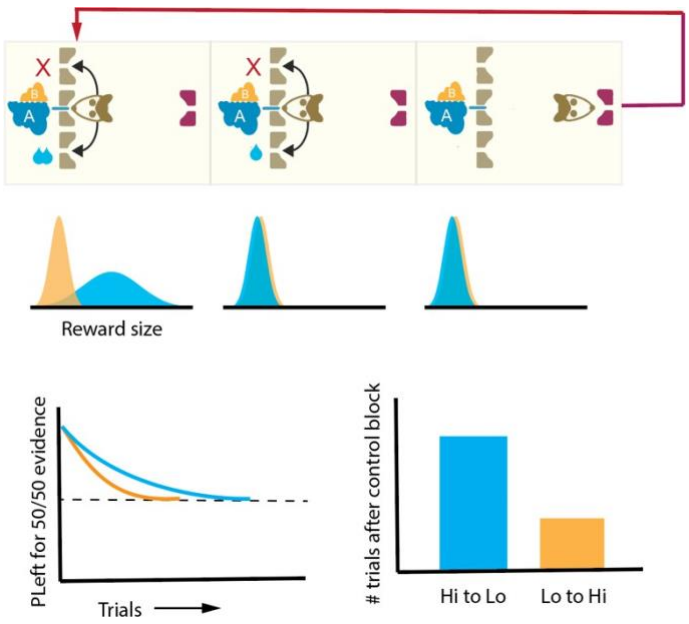


Figure 30 Proposed modification to future variance task to include a self-report of the block transition.

At an unsigned block transition, the reward environment becomes leaner and the port distributions are matched. The rat can re-initialize a richer environment by poking into a fourth "switch port." We expect that this will provide a clearer read-out for when the rat realizes it is in a different environment.

Future work

This preliminary behavioral data and computational framework sets the stage for future studies to test for static and dynamic representations of learning rate that could support meta-learning. In future, we will use dual-color head-mounted miniscopes (nVue from Inscopix) to simultaneously image excitatory pyramidal cells and target-defined projection cells in the OFC during freely moving behavior. We will test the hypothesis that ventral tegmental area projection neurons

encode distributions of reward size, that is, carries information about both magnitude and variance. Further, we will be able to determine to what degree the diversity of OFC neuronal responses map to distinct projection cell types. If decision variables do not map onto projection cell types and we observe mixed information across circuits, our large-scale recordings will enable shifting to population encoding of decision processes.

Chapter 6

Conclusion

Summary

The architecture of mammalian cortex has been shaped over thousands of years and can support varied functions soon after birth (Bhattasali et al., 2022; Yang and Molano-Mazón, 2021; Zador, 2019). As such, the circuit architecture imposes an evolutionary constraint on different computations, but an overarching projection logic has remained elusive. Here, I established the subcortical output architecture of rat orbitofrontal cortex (OFC) at a single neuron resolution. I found that most subcortically projecting neurons adhere to a one-neuron-one-target projection logic and revealed a previously unappreciated relationship between subcortical target and sublaminar position (a “deeper-farther” rule). The existence of a highly structured circuit bolsters the hypothesis that anatomically distinct projection types represent functionally distinct information channels. To test the functional roles of different projection types in OFC, I next

developed two task variants that enable testing for (1) static and dynamic learning rate representations by manipulating the outcome distribution and (2) subjective value representations that could drive dopaminergic signaling and in turn, momentary choice preferences. In future, these tasks can be used to assess for stable and task-general neural representations across different projection types in orbitofrontal cortex. Taken together, this work furthers a cell type-specific understanding of how OFC representations are routed to subcortical targets to drive healthy decision-making.

Future outlook

The orbitofrontal cortex (OFC) has taken center stage in studies of choice behavior (Cromwell et al., 2018; Kimmel et al., 2020; Masset et al., 2020; Padoa-Schioppa and Assad, 2006a). Though recently conceptualized as the hub of the brain's valuation system, we lack a quantitative understanding of what facets of valuation and learning OFC supports and how the relevant computations are produced by neural circuit mechanisms. In part, this is attributable to the immense diversity of neurobiological cell types from which diverse and complex response properties might arise.

The cell type diversity in the brain is staggering and at times, disheartening. It is often unclear what is "stamp collecting" and how to interpret overwhelming data from high throughput circuit mapping and sequencing. However, it is important to note that the end goal is not to identify endless cell types. Rather, it is to find the right level of granularity and establish a biologically grounded framework from which we can both constrain and infer circuit functions. Cell types are the lens that brings anatomy, genetics, and circuit function into focus on a single plane, yet little

is known about the cell-type-specific organization of OFC. Such an approach is critical to further a mechanistic understanding of decision-making and better understand the neural processes that give rise to behavior. It is critical to take a multi-disciplinary approach and run the table, from receptor interactions and cell types to high dimensional behaviors to computational models. Such a unified framework sets the stage for a better understanding as to how information is distributed and encoded, what are the core conserved circuits in frontal cortex, and how such circuits might be changed as a function of experience.

References

Adesnik, H., and Naka, A. (2018). Cracking the Function of Layers in the Sensory Cortex. Neuron *100*, 1028–1043. https://doi.org/10.1016/j.neuron.2018.10.032.

Akintunde, A., and Buxton, D.F. (1992). Origins and collateralization of corticospinal, corticopontine, corticorubral and corticostriatal tracts: a multiple retrograde fluorescent tracing study. Brain Research *586*, 208–218. https://doi.org/10.1016/0006-8993(92)91629-S.

Alcamo, E.A., Chirivella, L., Dautzenberg, M., Dobreva, G., Fariñas, I., Grosschedl, R., and McConnell, S.K. (2008). Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron *57*, 364–377. https://doi.org/10.1016/j.neuron.2007.12.012.

Allaway, K.C., Muñoz, W., Tremblay, R., Sherer, M., Herron, J., Rudy, B., Machold, R., and Fishell, G. (2020). Cellular birthdate predicts laminar and regional cholinergic projection topography in the forebrain. ELife *9*, e63249. https://doi.org/10.7554/eLife.63249.

Altman, J., and Bayer, S.A. (1978). Development of the diencephalon in the rat. II. Correlation of the embryonic development of the hypothalamus with the time of origin of its neurons. J Comp Neurol *182*, 973–993. https://doi.org/10.1002/cne.901820512.

Altman, J., and Bayer, S.A. (1980). Development of the brain stem in the rat. I. Thymidine-radiographic study of the time of origin of neurons of the lower medulla. Journal of Comparative Neurology *194*, 1–35. https://doi.org/10.1002/cne.901940102.

Altman, J., and Bayer, S.A. (1981). Development of the brain stem in the rat. V. Thymidine-radiographic study of the time of origin of neurons in the midbrain tegmentum. Journal of Comparative Neurology *198*, 677–716. https://doi.org/10.1002/cne.901980409.

Altshuler, L.L., Bookheimer, S.Y., Townsend, J., Proenza, M.A., Eisenberger, N., Sabb, F., Mintz, J., and Cohen, M.S. (2005). Blunted Activation in Orbitofrontal Cortex During Mania: A Functional Magnetic Resonance Imaging Study. Biological Psychiatry *58*, 763–769. https://doi.org/10.1016/j.biopsych.2005.09.012.

Angevine, J.B., and Sidman, R.L. (1961). Autoradiographic Study of Cell Migration during Histogenesis of Cerebral Cortex in the Mouse. Nature *192*, 766–768. https://doi.org/10.1038/192766b0.

Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M.D., et al. (2016). The origin and evolution of cell types. Nat Rev Genet *17*, 744–757. https://doi.org/10.1038/nrg.2016.127.

Arlotta, P., Molyneaux, B.J., Chen, J., Inoue, J., Kominami, R., and Macklis, J.D. (2005). Neuronal subtype-specific genes that control corticospinal motor neuron development in vivo. Neuron *45*, 207–221. https://doi.org/10.1016/j.neuron.2004.12.036.

Baker, A., Kalmbach, B., Morishima, M., Kim, J., Juavinett, A., Li, N., and Dembrow, N. (2018). Specialized Subpopulations of Deep-Layer Pyramidal Neurons in the Neocortex: Bridging Cellular Properties to Functional Consequences. J Neurosci *38*, 5441–5455. https://doi.org/10.1523/JNEUROSCI.0150-18.2018.

Bariselli, S., Miyazaki, N.L., Creed, M.C., and Kravitz, A.V. (2020). Orbitofrontal-striatal potentiation underlies cocaine-induced hyperactivity. Nat Commun *11*, 3996. https://doi.org/10.1038/s41467-020-17763-8.

Bates, C.A., and Killackey, H.P. (1984). The emergence of a discretely distributed pattern of corticospinal projection neurons. Brain Res *315*, 265–273. https://doi.org/10.1016/0165-3806(84)90161-5.

Bayer, S.A. (1980). Quantitative 3H-thymidine radiographic analyses of neurogenesis on the rat amygdala. Journal of Comparative Neurology *194*, 845–875. https://doi.org/10.1002/cne.901940409.

Bechara, A., and Damasio, H. (2002). Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. Neuropsychologia *40*, 1675–1689. https://doi.org/10.1016/s0028-3932(02)00015-5.

Bechara, A., Dolan, S., and Hindes, A. (2002). Decision-making and addiction (part II): myopia for the future or hypersensitivity to reward? Neuropsychologia *40*, 1690–1705. https://doi.org/10.1016/S0028-3932(02)00016-7.

Beier, K.T., Steinberg, E.E., DeLoach, K.E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X.J., Kremer, E.J., Malenka, R.C., and Luo, L. (2015a). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input–Output Mapping. Cell *162*, 622. https://doi.org/10.1016/j.cell.2015.07.015.

Beier, K.T., Steinberg, E.E., DeLoach, K.E., Xie, S., Miyamichi, K., Schwarz, L., Gao, X.J., Kremer, E.J., Malenka, R.C., and Luo, L. (2015b). Circuit Architecture of VTA Dopamine Neurons Revealed by Systematic Input-Output Mapping. Cell *162*, 622–634. https://doi.org/10.1016/j.cell.2015.07.015.

Beier, K.T., Gao, X.J., Xie, S., DeLoach, K.E., Malenka, R.C., and Luo, L. (2019). Topological Organization of Ventral Tegmental Area Connectivity Revealed by Viral-Genetic Dissection of Input-Output Relations. Cell Reports *26*, 159-167.e6. https://doi.org/10.1016/j.celrep.2018.12.040.

Berendse, H.W., Graaf, Y.G.-D., and Groenewegen, H.J. (1992). Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. Journal of Comparative Neurology *316*, 314–347. https://doi.org/10.1002/cne.903160305.

Berke, J.D. (2018). What does dopamine mean? Nat Neurosci *21*, 787–793. https://doi.org/10.1038/s41593-018-0152-y.

Berridge, K.C., and Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research Reviews *28*, 309–369. https://doi.org/10.1016/S0165-0173(98)00019-8.

Beucke, J.C., Sepulcre, J., Talukdar, T., Linnman, C., Zschenderlein, K., Endrass, T., Kaufmann, C., and Kathmann, N. (2013). Abnormally High Degree Connectivity of the Orbitofrontal Cortex in Obsessive-Compulsive Disorder. JAMA Psychiatry *70*, 619–629. https://doi.org/10.1001/jamapsychiatry.2013.173.

Beul, S.F., and Hilgetag, C.C. (2015). Towards a "canonical" agranular cortical microcircuit. Frontiers in Neuroanatomy *8*. .

Beyeler, A., Chang, C.-J., Silvestre, M., Lévêque, C., Namburi, P., Wildes, C.P., and Tye, K.M. (2018). Organization of Valence-Encoding and Projection-Defined Neurons in the Basolateral Amygdala. Cell Reports *22*, 905–918. https://doi.org/10.1016/j.celrep.2017.12.097.

Bhaduri, A., Sandoval-Espinosa, C., Otero-Garcia, M., Oh, I., Yin, R., Eze, U.C., Nowakowski, T.J., and Kriegstein, A.R. (2021). An atlas of cortical arealization identifies dynamic molecular signatures. Nature *598*, 200–204. https://doi.org/10.1038/s41586-021-03910-8.

Bhattasali, N.X., Zador, A.M., and Engel, T.A. (2022). Neural Circuit Architectural Priors for Embodied Control. ArXiv:2201.05242 [Cs, q-Bio].

Bicks, L.K., Koike, H., Akbarian, S., and Morishita, H. (2015). Prefrontal Cortex and Social Cognition in Mouse and Man. Frontiers in Psychology *6*. .

Bishop, K.M., Goudreau, G., and O'Leary, D.D.M. (2000). Regulation of Area Identity in the Mammalian Neocortex by Emx2 and Pax6. Science https://doi.org/10.1126/science.288.5464.344.

Bourassa, J., and Deschênes, M. (1995). Corticothalamic projections from the primary visual cortex in rats: a single fiber study using biocytin as an anterograde tracer. Neuroscience *66*, 253–263. https://doi.org/10.1016/0306-4522(95)00009-8.

Bourassa, J., Pinault, D., and Deschênes, M. (1995). Corticothalamic Projections from the Cortical Barrel Field to the Somatosensory Thalamus in Rats: A Single-fibre Study Using Biocytin as an Anterograde Tracer. European Journal of Neuroscience *7*, 19–30. https://doi.org/10.1111/j.1460-9568.1995.tb01016.x.

Brand, M., Kalbe, E., Labudda, K., Fujiwara, E., Kessler, J., and Markowitsch, H.J. (2005). Decision-making impairments in patients with pathological gambling. Psychiatry Res *133*, 91–99. https://doi.org/10.1016/j.psychres.2004.10.003.

Bremner, J.D., Vythilingam, M., Vermetten, E., Nazeer, A., Adil, J., Khan, S., Staib, L.H., and Charney, D.S. (2002). Reduced volume of orbitofrontal cortex in major depression. Biological Psychiatry *51*, 273–279. https://doi.org/10.1016/S0006-3223(01)01336-1.

Britanova, O., de Juan Romero, C., Cheung, A., Kwan, K.Y., Schwark, M., Gyorgy, A., Vogel, T., Akopov, S., Mitkovski, M., Agoston, D., et al. (2008). Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. Neuron *57*, 378–392. https://doi.org/10.1016/j.neuron.2007.12.028.

Bromberg-Martin, E.S., and Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. Neuron *63*, 119–126. https://doi.org/10.1016/j.neuron.2009.06.009.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. Neuron *68*, 815–834. https://doi.org/10.1016/j.neuron.2010.11.022.

Brown, S.P., and Hestrin, S. (2009). Intracortical circuits of pyramidal neurons reflect their long-range axonal targets. Nature *457*, 1133–1136. https://doi.org/10.1038/nature07658.

Burke, C.J., Soutschek, A., Weber, S., Raja Beharelle, A., Fehr, E., Haker, H., and Tobler, P.N. (2018). Dopamine Receptor-Specific Contributions to the Computation of Value. Neuropsychopharmacol. *43*, 1415–1424. https://doi.org/10.1038/npp.2017.302.

Callaway, E.M., Dong, H.-W., Ecker, J.R., Hawrylycz, M.J., Huang, Z.J., Lein, E.S., Ngai, J., Osten, P., Ren, B., Tolias, A.S., et al. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. Nature *598*, 86–102. https://doi.org/10.1038/s41586-021-03950-0.

Cannon, C.M., Abdallah, L., Tecott, L.H., During, M.J., and Palmiter, R.D. (2004). Dysregulation of Striatal Dopamine Signaling by Amphetamine Inhibits Feeding by Hungry Mice. Neuron *44*, 509–520. https://doi.org/10.1016/j.neuron.2004.10.009.

Caviness, V.S., and Rakic, P. (1978). Mechanisms of Cortical Development: A View From Mutations in Mice. Annu. Rev. Neurosci. *1*, 297–326. https://doi.org/10.1146/annurev.ne.01.030178.001501.

Chagnac-Amitai, Y., Luhmann, H.J., and Prince, D.A. (1990). Burst generating and regular spiking layer 5 pyramidal neurons of rat neocortex have different morphological features. J Comp Neurol *296*, 598–613. https://doi.org/10.1002/cne.902960407.

Chamberlain, S.R., Fineberg, N.A., Menzies, L.A., Blackwell, A.D., Bullmore, E.T., Robbins, T.W., and Sahakian, B.J. (2007). Impaired cognitive flexibility and motor inhibition in unaffected first-degree relatives of patients with obsessive-compulsive disorder. Am J Psychiatry *164*, 335–338. https://doi.org/10.1176/ajp.2007.164.2.335.

Chandler, D.J., Lamperski, C.S., and Waterhouse, B.D. (2013). Identification and distribution of projections from monoaminergic and cholinergic nuclei to functionally differentiated subregions

of prefrontal cortex. Brain Research *1522*, 38–58. https://doi.org/10.1016/J.BRAINRES.2013.04.057.

Chen, B., Schaevitz, L.R., and McConnell, S.K. (2005). Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. Proceedings of the National Academy of Sciences *102*, 17184–17189. https://doi.org/10.1073/pnas.0508732102.

Chen, B., Wang, S.S., Hattox, A.M., Rayburn, H., Nelson, S.B., and McConnell, S.K. (2008). The Fezf2–Ctip2 genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. PNAS *105*, 11382–11387. https://doi.org/10.1073/pnas.0804918105.

Chen, X., Sun, Y.-C., Zhan, H., Kebschull, J.M., Fischer, S., Matho, K., Huang, Z.J., Gillis, J., and Zador, A.M. (2019). High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ Sequencing. Cell *179*, 772-786.e19. https://doi.org/10.1016/j.cell.2019.09.023.

Churchland, M.M., and Shenoy, K.V. (2007). Temporal Complexity and Heterogeneity of Single-Neuron Activity in Premotor and Motor Cortex. Journal of Neurophysiology *97*, 4235–4257. https://doi.org/10.1152/jn.00095.2007.

Clatworthy, P.L., Lewis, S.J.G., Brichard, L., Hong, Y.T., Izquierdo, D., Clark, L., Cools, R., Aigbirhio, F.I., Baron, J.-C., Fryer, T.D., et al. (2009). Dopamine Release in Dissociable Striatal Subregions Predicts the Different Effects of Oral Methylphenidate on Reversal Learning and Spatial Working Memory. J. Neurosci. *29*, 4690–4696. https://doi.org/10.1523/JNEUROSCI.3266-08.2009.

Coizet, V., Heilbronner, S.R., Carcenac, C., Mailly, P., Lehman, J.F., Savasta, M., David, O., Deniau, J.-M., Groenewegen, H.J., and Haber, S.N. (2017). Organization of the Anterior Limb of the Internal Capsule in the Rat. J Neurosci *37*, 2539–2554. https://doi.org/10.1523/JNEUROSCI.3304-16.2017.

Constantinople, C.M., and Bruno, R.M. (2013). Deep Cortical Layers Are Activated Directly by Thalamus. Science *340*, 1591–1594. https://doi.org/10.1126/science.1236425.

Constantinople, C.M., Piet, A.T., Bibawi, P., Akrami, A., Kopec, C., and Brody, C.D. (2019a). Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. ELife *8*, e49744. https://doi.org/10.7554/eLife.49744.

Constantinople, C.M., Piet, A.T., and Brody, C.D. (2019b). An Analysis of Decision under Risk in Rats. Current Biology *29*, 2066-2074.e5. https://doi.org/10.1016/j.cub.2019.05.013.

Cromwell, H.C., Tremblay, L., and Schultz, W. (2018). Neural encoding of choice during a delayed response task in primate striatum and orbitofrontal cortex. Exp Brain Res *236*, 1679–1688. https://doi.org/10.1007/s00221-018-5253-z.

Dahmen, D., Layer, M., Deutz, L., Dąbrowska, P.A., Voges, N., von Papen, M., Brochier, T., Riehle, A., Diesmann, M., Grün, S., et al. (2022). Global organization of neuronal activity only requires unstructured local connectivity. ELife *11*, e68422. https://doi.org/10.7554/eLife.68422.

Dalley, J.W., Everitt, B.J., and Robbins, T.W. (2011). Impulsivity, Compulsivity, and Top-Down Cognitive Control. Neuron *69*, 680–694. https://doi.org/10.1016/j.neuron.2011.01.020.

Dani, V.S., and Nelson, S.B. (2009). Intact Long-Term Potentiation but Reduced Connectivity between Neocortical Layer 5 Pyramidal Neurons in a Mouse Model of Rett Syndrome. J. Neurosci. *29*, 11263–11270. https://doi.org/10.1523/JNEUROSCI.1019-09.2009.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. Nature *441*, 876–879. https://doi.org/10.1038/nature04766.

Dayan, P., and Yu, A.J. (2002). Expected and Unexpected Uncertainty: ACh and NE in the Neocortex. In Advances in Neural Information Processing Systems, (MIT Press), p.

DeFelipe, Javier, and Edward G. Jones. (1988). Cajal on the cerebral cortex: An annotated translation of the complete writings. (Oxford Unive. Press).

Delgado, M.R., Phelps, E.A., and Robbins, T.W. (2011). Decision Making, Affect, and Learning: Attention and Performance XXIII (OUP Oxford).

Di Ciano, P., Robbins, T.W., and Everitt, B.J. (2008). Differential Effects of Nucleus Accumbens Core, Shell, or Dorsal Striatal Inactivations on the Persistence, Reacquisition, or Reinstatement of Responding for a Drug-Paired Conditioned Reinforcer. Neuropsychopharmacol *33*, 1413–1425. https://doi.org/10.1038/sj.npp.1301522.

Dorris, M.C., and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. Neuron *44*, 365–378. https://doi.org/10.1016/j.neuron.2004.09.009.

Duncker, L., and Sahani, M. (2021). Dynamics on the manifold: Identifying computational dynamical activity from neural population recordings. Current Opinion in Neurobiology *70*, 163–170. https://doi.org/10.1016/j.conb.2021.10.014.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018a). Distinct descending motor cortex pathways and their roles in movement. Nature *563*, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Economo, M.N., Viswanathan, S., Tasic, B., Bas, E., Winnubst, J., Menon, V., Graybuck, L.T., Nguyen, T.N., Smith, K.A., Yao, Z., et al. (2018b). Distinct descending motor cortex pathways and their roles in movement. Nature *563*, 79–84. https://doi.org/10.1038/s41586-018-0642-9.

Ersche, K.D., Fletcher, P.C., Roiser, J.P., Fryer, T.D., London, M., Robbins, T.W., and Sahakian, B.J. (2006). Differences in orbitofrontal activation during decision-making between methadone-maintained opiate users, heroin users and healthy volunteers. Psychopharmacology (Berl) *188*, 364–373. https://doi.org/10.1007/s00213-006-0515-z.

Fentress, J.C., Stanfield, B.B., and Cowan, W.M. (1981). Observations on the development of the striatum in mice and rats. Anat Embryol *163*, 275–298. https://doi.org/10.1007/BF00315705.

Fernandes, B.S., Williams, L.M., Steiner, J., Leboyer, M., Carvalho, A.F., and Berk, M. (2017). The new field of 'precision psychiatry.' BMC Medicine *15*, 80. https://doi.org/10.1186/s12916-017-0849-x.

Fink, J.S., and Smith, G.P. (1980). Mesolimbic and mesocortical dopaminergic neurons are necessary for normal exploratory behavior in rats. Neuroscience Letters *17*, 61–65. https://doi.org/10.1016/0304-3940(80)90062-2.

Finlay, B.L., and Darlington, R.B. (1995). Linked regularities in the development and evolution of mammalian brains. Science *268*, 1578–1584. https://doi.org/10.1126/science.7777856.

Fiorillo, C.D. (2011). Transient activation of midbrain dopamine neurons by reward risk. Neuroscience *197*, 162–171. https://doi.org/10.1016/j.neuroscience.2011.09.037.

Fishell, G. (1995). Striatal precursors adopt cortical identities in response to local cues. Development *121*, 803–812. https://doi.org/10.1242/dev.121.3.803.

Fishell, G., and Heintz, N. (2013). The Neuron Identity Problem: Form Meets Function. Neuron *80*, 602–612. https://doi.org/10.1016/j.neuron.2013.10.035.

Flagel, S.B., Watson, S.J., Robinson, T.E., and Akil, H. (2007). Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. Psychopharmacology *191*, 599–607. https://doi.org/10.1007/s00213-006-0535-8.

Fleming, S.M., and Dolan, R.J. (2012). The neural basis of metacognitive ability. Philos Trans R Soc Lond B Biol Sci *367*, 1338–1349. https://doi.org/10.1098/rstb.2011.0417.

Fletcher, L.N., and Williams, S.R. (2019). Neocortical Topology Governs the Dendritic Integrative Capacity of Layer 5 Pyramidal Neurons. Neuron *101*, 76-90.e4. https://doi.org/10.1016/j.neuron.2018.10.048.

Fournier, J., Müller, C.M., and Laurent, G. (2015). Looking for the roots of cortical sensory computation in three-layered cortices. Current Opinion in Neurobiology *31*, 119–126. https://doi.org/10.1016/j.conb.2014.09.006.

Frantz, G.D., and McConnell, S.K. (1996). Restriction of Late Cerebral Cortical Progenitors to an Upper-Layer Fate. Neuron *17*, 55–61. https://doi.org/10.1016/S0896-6273(00)80280-9.

Freels, T.G., Gabriel, D.B.K., Lester, D.B., and Simon, N.W. (2020). Risky decision-making predicts dopamine release dynamics in nucleus accumbens shell. Neuropsychopharmacol. *45*, 266–275. https://doi.org/10.1038/s41386-019-0527-0.

Fuster, J.M. (2001). The prefrontal cortex--an update: time is of the essence. Neuron *30*, 319–333. https://doi.org/10.1016/s0896-6273(01)00285-9.

Gabbott, P.L.A., and Stewart, M.G. (1987). Distribution of neurons and glia in the visual cortex (area 17) of the adult albino rat: A quantitative description. Neuroscience *21*, 833–845. https://doi.org/10.1016/0306-4522(87)90040-6.

Gabbott, P.L.A., Warner, T.A., Jays, P.R.L., Salway, P., and Busby, S.J. (2005a). Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. J Comp Neurol *492*, 145–177. https://doi.org/10.1002/cne.20738.

Gabbott, P.L.A., Warner, T.A., Jays, P.R.L., Salway, P., and Busby, S.J. (2005b). Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. The Journal of Comparative Neurology *492*, 145–177. https://doi.org/10.1002/cne.20738.

Gallagher, M., McMahan, R.W., and Schoenbaum, G. (1999). Orbitofrontal Cortex and Representation of Incentive Value in Associative Learning. J. Neurosci. *19*, 6610–6614. https://doi.org/10.1523/JNEUROSCI.19-15-06610.1999.

Gandal, M.J., Leppa, V., Won, H., Parikshak, N.N., and Geschwind, D.H. (2016). The road to precision psychiatry: translating genetics into disease mechanisms. Nat Neurosci *19*, 1397–1407. https://doi.org/10.1038/nn.4409.

Gao, Z., Wang, H., Lu, C., Lu, T., Froudist-Walsh, S., Chen, M., Wang, X.-J., Hu, J., and Sun, W. The neural basis of delayed gratification. Science Advances *7*, eabg6611. https://doi.org/10.1126/sciadv.abg6611.

Gardner, M.P., Conroy, J.C., Styer, C.V., Huynh, T., Whitaker, L.R., and Schoenbaum, G. (2018). Medial orbitofrontal inactivation does not affect economic choice. ELife *7*, e38963. https://doi.org/10.7554/eLife.38963.

Gardner, M.P.H., Conroy, J.S., Shaham, M.H., Styer, C.V., and Schoenbaum, G. (2017). Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. Neuron *96*, 1192-1203.e4. https://doi.org/10.1016/j.neuron.2017.10.026.

Gergues, M.M., Han, K.J., Choi, H.S., Brown, B., Clausing, K.J., Turner, V.S., Vainchtein, I.D., Molofsky, A.V., and Kheirbek, M.A. (2020). Circuit and molecular architecture of a ventral hippocampal network. Nat Neurosci *23*, 1444–1452. https://doi.org/10.1038/s41593-020-0705-8.

Gianino, S., Stein, S.A., Li, H., Lu, X., Biesiada, E., Ulas, J., and Xu, X.M. (1999). Postnatal growth of corticospinal axons in the spinal cord of developing mice. Developmental Brain Research *112*, 189–204. https://doi.org/10.1016/S0165-3806(98)00168-0.

Gilbert, C.D., and Kelly, J.P. (1975). The projections of cells in different layers of the cat's visual cortex. Journal of Comparative Neurology *163*, 81–105. https://doi.org/10.1002/cne.901630106.

Girgis, R.R., Minshew, N.J., Melhem, N.M., Nutche, J.J., Keshavan, M.S., and Hardan, A.Y. (2007). Volumetric alterations of the orbitofrontal cortex in autism. Progress in Neuro-Psychopharmacology and Biological Psychiatry *31*, 41–45. https://doi.org/10.1016/j.pnpbp.2006.06.007.

Gokce, O., Stanley, G., Treutlein, B., Neff, N.F., Camp, G.J., Malenka, R.C., Rothwell, P.E., Fuccillo, M.V., Südhof, T.C., and Quake, S.R. (2016). Cellular Taxonomy of the Mouse

Striatum as Revealed by Single-Cell RNA-Seq. Cell Rep *16*, 1126–1137. https://doi.org/10.1016/j.celrep.2016.06.059.

Gordon, J.A. (2016). On being a circuit psychiatrist. Nat Neurosci *19*, 1385–1386. https://doi.org/10.1038/nn.4419.

Gottesman, I.I., and Gould, T.D. (2003). The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions. AJP *160*, 636–645. https://doi.org/10.1176/appi.ajp.160.4.636.

Gottfried, J.A. (2003). Encoding Predictive Reward Value in Human Amygdala and Orbitofrontal Cortex. Science *301*, 1104–1107. https://doi.org/10.1126/science.1087919.

Gouwens, N.W., Sorensen, S.A., Berg, J., Lee, C., Jarsky, T., Ting, J., Sunkin, S.M., Feng, D., Anastassiou, C.A., Barkan, E., et al. (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nat Neurosci *22*, 1182–1195. https://doi.org/10.1038/s41593-019-0417-0.

Grabenhorst, F., Tsutsui, K.-I., Kobayashi, S., and Schultz, W. (2019). Primate prefrontal neurons signal economic risk derived from the statistics of recent reward experience. ELife *8*, e44838. https://doi.org/10.7554/eLife.44838.

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

Gremel, C.M., Chancey, J.H., Atwood, B.K., Luo, G., Neve, R., Ramakrishnan, C., Deisseroth, K., Lovinger, D.M., and Costa, R.M. (2016). Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. Neuron *90*, 1312–1324. https://doi.org/10.1016/j.neuron.2016.04.043.

Gribnau, A.A.M., de Kort, E.J.M., Dederen, P.J.W.C., and Nieuwenhuys, R. (1986). On the development of the pyramidal tract in the rat. Anat Embryol *175*, 101–110. https://doi.org/10.1007/BF00315460.

Groman, S.M., and Jentsch, J.D. (2012). Cognitive control and the dopamine D2-like receptor: a dimensional understanding of addiction. Depression and Anxiety *29*, 295–306. https://doi.org/10.1002/da.20897.

Groman, S.M., Keistler, C., Keip, A.J., Hammarlund, E., DiLeone, R.J., Pittenger, C., Lee, D., and Taylor, J.R. (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. Neuron *103*, 734-746.e3. https://doi.org/10.1016/j.neuron.2019.05.042.

Guan, D., Armstrong, W.E., and Foehring, R.C. (2015). Electrophysiological properties of genetically identified subtypes of layer 5 neocortical pyramidal neurons: Ca2+ dependence and differential modulation by norepinephrine. Journal of Neurophysiology *113*, 2014–2032. https://doi.org/10.1152/jn.00524.2014.

Gueguen, M.C.M., Schweitzer, E.M., and Konova, A.B. (2021). Computational theory-driven studies of reinforcement learning and decision-making in addiction: What have we learned? Curr Opin Behav Sci *38*, 40–48. https://doi.org/10.1016/j.cobeha.2020.08.007.

Hamid, A.A., Pettibone, J.R., Mabrouk, O.S., Hetrick, V.L., Schmidt, R., Vander Weele, C.M., Kennedy, R.T., Aragona, B.J., and Berke, J.D. (2016). Mesolimbic dopamine signals the value of work. Nat Neurosci *19*, 117–126. https://doi.org/10.1038/nn.4173.

Han, Y., Kebschull, J.M., Campbell, R.A.A., Cowan, D., Imhof, F., Zador, A.M., and Mrsic-Flogel, T.D. (2018a). The logic of single-cell projections from visual cortex. Nature *556*, 51–56. https://doi.org/10.1038/nature26159.

Han, Y., Kebschull, J.M., Campbell, R.A.A., Cowan, D., Imhof, F., Zador, A.M., and Mrsic-Flogel, T.D. (2018b). The logic of single-cell projections from visual cortex. Nature *556*, 51–56. https://doi.org/10.1038/nature26159.

Harada, M., Pascoli, V., Hiver, A., Flakowski, J., and Lüscher, C. (2021). Corticostriatal Activity Driving Compulsive Reward Seeking. Biological Psychiatry *90*, 808–818. https://doi.org/10.1016/j.biopsych.2021.08.018.

Harris, K.D., and Shepherd, G.M.G. (2015). The neocortical circuit: themes and variations. Nat Neurosci *18*, 170–181. https://doi.org/10.1038/nn.3917.

Harris, J.A., Mihalas, S., Hirokawa, K.E., Whitesell, J.D., Choi, H., Bernard, A., Bohn, P., Caldejon, S., Casal, L., Cho, A., et al. (2019). Hierarchical organization of cortical and thalamic connectivity. Nature *575*, 195–202. https://doi.org/10.1038/s41586-019-1716-z.

Hattox, A.M., and Nelson, S.B. (2007). Layer V Neurons in Mouse Cortex Projecting to Different Targets Have Distinct Physiological Properties. Journal of Neurophysiology *98*, 3330–3340. https://doi.org/10.1152/jn.00397.2007.

Heerey, E.A., Bell-Warren, K.R., and Gold, J.M. (2008). Decision-Making Impairments in the Context of Intact Reward Sensitivity in Schizophrenia. Biol Psychiatry *64*, 62–69. https://doi.org/10.1016/j.biopsych.2008.02.015.

Hirai, Y., Morishima, M., Karube, F., and Kawaguchi, Y. (2012). Specialized Cortical Subnetworks Differentially Connect Frontal Cortex to Parahippocampal Areas. J. Neurosci. *32*, 1898–1913. https://doi.org/10.1523/JNEUROSCI.2810-11.2012.

Hirokawa, J., Vaughan, A., Masset, P., Ott, T., and Kepecs, A. (2019). Frontal cortex neuron types categorically encode single decision variables. Nature *576*, 446–451. https://doi.org/10.1038/s41586-019-1816-9.

Hocker, D.L., Brody, C.D., Savin, C., and Constantinople, C.M. (2021). Subpopulations of neurons in lOFC encode previous and current rewards at time of choice. ELife *10*, e70129. https://doi.org/10.7554/eLife.70129.

du Hoffmann, J., and Nicola, S.M. (2014). Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens. J Neurosci *34*, 14349–14364. https://doi.org/10.1523/JNEUROSCI.3492-14.2014.

Hollon, N.G., Arnold, M.M., Gan, J.O., Walton, M.E., and Phillips, P.E.M. (2014). Dopamine-associated cached values are not sufficient as the basis for action selection. Proceedings of the National Academy of Sciences *111*, 18357–18362. https://doi.org/10.1073/pnas.1419770111.

Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579. https://doi.org/10.1038/nature12475.

Huang, L., Kebschull, J.M., Fürth, D., Musall, S., Kaufman, M.T., Churchland, A.K., and Zador, A.M. (2020). BRICseq Bridges Brain-wide Interregional Connectivity to Neural Activity and Gene Expression in Single Animals. Cell *182*, 177-188.e27. https://doi.org/10.1016/j.cell.2020.05.029.

Insel, T.R. (2014). The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry. AJP *171*, 395–397. https://doi.org/10.1176/appi.ajp.2014.14020138.

Insel, T.R., and Quirion, R. (2005). Psychiatry as a Clinical Neuroscience Discipline. JAMA *294*, 2221–2224. https://doi.org/10.1001/jama.294.17.2221.

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

Jazayeri, M., and Ostojic, S. (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. ArXiv:2107.04084 [q-Bio].

Jensen, K.F., and Altman, J. (1982). The contribution of late-generated neurons to the callosal projection in rat: a study with prenatal x-irradiation. J Comp Neurol *209*, 113–122. https://doi.org/10.1002/cne.902090202.

Jo, Y.S., Heymann, G., and Zweifel, L.S. (2018). Dopamine Neurons Reflect the Uncertainty in Fear Generalization. Neuron *100*, 916-925.e3. https://doi.org/10.1016/j.neuron.2018.09.028.

Joosten, E.A.J., Gribnau, A.A.M., and Dederen, P.J.W.C. (1989). Postnatal development of the corticospinal tract in the rat. Anat Embryol *179*, 449–456. https://doi.org/10.1007/BF00319587.

Karten, H.J. (2013). Neocortical Evolution: Neuronal Circuits Arise Independently of Lamination. Current Biology *23*, R12–R15. https://doi.org/10.1016/j.cub.2012.11.013.

Kasper, E.M., Larkman, A.U., Lübke, J., and Blakemore, C. (1994). Pyramidal neurons in layer 5 of the rat visual cortex. I. Correlation among cell morphology, intrinsic electrophysiological properties, and axon targets. J Comp Neurol *339*, 459–474. https://doi.org/10.1002/cne.903390402.

Kebschull, J.M., Garcia da Silva, P., Reid, A.P., Peikon, I.D., Albeanu, D.F., and Zador, A.M. (2016a). High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. Neuron *91*, 975–987. https://doi.org/10.1016/j.neuron.2016.07.036.

Kebschull, J.M., Garcia Da Silva, P., Reid, A.P., Peikon, I.D., Albeanu, D.F., and Zador Correspondence, A.M. (2016b). High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. https://doi.org/10.1016/j.neuron.2016.07.036.

Kepecs, A., and Fishell, G. (2014). Interneuron cell types are fit to function. Nature *505*, 318–326. https://doi.org/10.1038/nature12983.

Keramati, M., and Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. ELife *3*, e04811. https://doi.org/10.7554/eLife.04811.

Killackey, H.P., Koralek, K.A., Chiaia, N.L., and Rhodes, R.W. (1989). Laminar and areal differences in the origin of the subcortical projection neurons of the rat somatosensory cortex. J Comp Neurol *282*, 428–445. https://doi.org/10.1002/cne.902820309.

Kim, H.R., Malik, A.N., Mikhael, J.G., Bech, P., Tsutsui-Kimura, I., Sun, F., Zhang, Y., Li, Y., Watabe-Uchida, M., Gershman, S.J., et al. (2020). A Unified Framework for Dopamine Signals across Timescales. Cell *183*, 1600-1616.e25. https://doi.org/10.1016/j.cell.2020.11.013.

Kim, M.-H., Znamenskiy, P., Iacaruso, M.F., and Mrsic-Flogel, T.D. (2018). Segregated Subnetworks of Intracortical Projection Neurons in Primary Visual Cortex. Neuron *100*, 1313-1321.e6. https://doi.org/10.1016/j.neuron.2018.10.023.

Kim, Y., Yang, G.R., Pradhan, K., Venkataraju, K.U., Bota, M., García Del Molino, L.C., Fitzgerald, G., Ram, K., He, M., Levine, J.M., et al. (2017). Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. Cell *171*, 456-469.e22. https://doi.org/10.1016/j.cell.2017.09.020.

Kimmel, D.L., Elsayed, G.F., Cunningham, J.P., and Newsome, W.T. (2020). Value and choice as separable and stable representations in orbitofrontal cortex. Nat Commun *11*, 3466. https://doi.org/10.1038/s41467-020-17058-y.

Kita, T., and Kita, H. (2012). The Subthalamic Nucleus Is One of Multiple Innervation Sites for Long-Range Corticofugal Axons: A Single-Axon Tracing Study in the Rat. J. Neurosci. *32*, 5990–5999. https://doi.org/10.1523/JNEUROSCI.5717-11.2012.

Klein, S., Staring, M., Murphy, K., Viergever, M.A., and Pluim, J.P.W. (2010). elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging *29*, 196–205. https://doi.org/10.1109/TMI.2009.2035616.

Klingler, E., Tomasello, U., Prados, J., Kebschull, J.M., Contestabile, A., Galiñanes, G.L., Fièvre, S., Santinha, A., Platt, R., Huber, D., et al. (2021). Temporal controls over inter-areal cortical projection neuron fate diversity. Nature *599*, 453–457. https://doi.org/10.1038/s41586-021-04048-3.

Lai, T., Jabaudon, D., Molyneaux, B.J., Azim, E., Arlotta, P., Menezes, J.R.L., and Macklis, J.D. (2008). SOX5 Controls the Sequential Generation of Distinct Corticofugal Neuron Subtypes. Neuron *57*, 232–247. https://doi.org/10.1016/j.neuron.2007.12.023.

Lak, A., Stauffer, W.R., and Schultz, W. (2014a). Dopamine prediction error responses integrate subjective value from different reward dimensions. Proceedings of the National Academy of Sciences *111*, 2343–2348. https://doi.org/10.1073/pnas.1321596111.

Lak, A., Costa, G.M., Romberg, E., Koulakov, A.A., Mainen, Z.F., and Kepecs, A. (2014b). Orbitofrontal Cortex Is Required for Optimal Waiting Based on Decision Confidence. Neuron *84*, 190–201. https://doi.org/10.1016/j.neuron.2014.08.039.

Lak, A., Okun, M., Moss, M.M., Gurnani, H., Farrell, K., Wells, M.J., Reddy, C.B., Kepecs, A., Harris, K.D., and Carandini, M. (2020). Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value. Neuron *105*, 700-711.e6. https://doi.org/10.1016/j.neuron.2019.11.018.

Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung, H.-L., Chen, S., et al. (2016a). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science *352*, 1586–1590. https://doi.org/10.1126/science.aaf1204.

Lake, B.M., Ullman, T.D., Tenenbaum, J.B., and Gershman, S.J. (2016b). Building Machines That Learn and Think Like People. ArXiv:1604.00289 [Cs, Stat].

Larkum, M.E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic Integration in Tuft Dendrites of Layer 5 Pyramidal Neurons: A New Unifying Principle. Science https://doi.org/10.1126/science.1171958.

Lau, B., and Glimcher, P.W. (2005). Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys. J Exp Anal Behav *84*, 555–579. https://doi.org/10.1901/jeab.2005.110-04.

Lawrence, N.S., Wooderson, S., Mataix-Cols, D., David, R., Speckens, A., and Phillips, M.L. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in obsessive-compulsive disorder. Neuropsychology *20*, 409–419. https://doi.org/10.1037/0894-4105.20.4.409.

Leone, D.P., Srinivasan, K., Chen, B., Alcamo, E., and McConnell, S.K. (2008). The determination of projection neuron identity in the developing cerebral cortex. Current Opinion in Neurobiology *18*, 28–35. https://doi.org/10.1016/j.conb.2008.05.006.

Lévesque, M., Charara, A., Gagnon, S., Parent, A., and Deschênes, M. (1996a). Corticostriatal projections from layer V cells in rat are collaterals of long-range corticofugal axons. Brain Research *709*, 311–315. https://doi.org/10.1016/0006-8993(95)01333-4.

Lévesque, M., Gagnon, S., Parent, A., and Deschênes, null (1996b). Axonal arborizations of corticostriatal and corticothalamic fibers arising from the second somatosensory area in the rat. Cereb Cortex *6*, 759–770. https://doi.org/10.1093/cercor/6.6.759.

Levy, R., and Dubois, B. (2006). Apathy and the Functional Anatomy of the Prefrontal Cortex–Basal Ganglia Circuits. Cerebral Cortex *16*, 916–928. https://doi.org/10.1093/cercor/bhj043.

Li, H., Horns, F., Wu, B., Xie, Q., Li, J., Li, T., Luginbuhl, D.J., Quake, S.R., and Luo, L. (2017). Classifying Drosophila Olfactory Projection Neuron Subtypes by Single-cell RNA Sequencing. Cell *171*, 1206-1220.e22. https://doi.org/10.1016/j.cell.2017.10.019.

Li, S.-J., Vaughan, A., Sturgill, J.F., and Kepecs, A. (2018). A Viral Receptor Complementation Strategy to Overcome CAV-2 Tropism for Efficient Retrograde Targeting of Neurons. Neuron *98*, 905-917.e5. https://doi.org/10.1016/j.neuron.2018.05.028.

Lichtenberg, N.T., Pennington, Z.T., Holley, S.M., Greenfield, V.Y., Cepeda, C., Levine, M.S., and Wassum, K.M. (2017). Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. J. Neurosci. *37*, 8374–8384. https://doi.org/10.1523/JNEUROSCI.0486-17.2017.

Lichtenberg, N.T., Lee, B., Kashtelyan, V., Chappa, B.S., Girma, H.T., Green, E.A., Kantor, S., Lagowala, D.A., Myers, M.A., Potemri, D., et al. (2018). Rat behavior and dopamine release are modulated by conspecific distress. ELife *7*, e38090. https://doi.org/10.7554/eLife.38090.

Lui, J.H., Nguyen, N.D., Grutzner, S.M., Darmanis, S., Peixoto, D., Wagner, M.J., Allen, W.E., Kebschull, J.M., Richman, E.B., Ren, J., et al. (2021). Differential encoding in prefrontal cortex projection neuron classes across cognitive tasks. Cell *184*, 489-506.e26. https://doi.org/10.1016/j.cell.2020.11.046.

Lynd-Balta, E., and Haber, S.N. (1994). The organization of midbrain projections to the ventral striatum in the primate. Neuroscience *59*, 609–623. https://doi.org/10.1016/0306-4522(94)90181-3.

Malvaez, M., Shieh, C., Murphy, M.D., Greenfield, V.Y., and Wassum, K.M. (2019). Distinct cortical–amygdala projections drive reward value encoding and retrieval. Nat Neurosci *22*, 762–769. https://doi.org/10.1038/s41593-019-0374-7.

Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. The Journal of Physiology *500*, 409–440. https://doi.org/10.1113/jphysiol.1997.sp022031.

Märtin, A., Calvigioni, D., Tzortzi, O., Fuzik, J., Wärnberg, E., and Meletis, K. (2019). A Spatiomolecular Map of the Striatum. Cell Rep *29*, 4320-4333.e5. https://doi.org/10.1016/j.celrep.2019.11.096.

Mason, A., and Larkman, A. (1990). Correlations between morphology and electrophysiology of pyramidal neurons in slices of rat visual cortex. II. Electrophysiology. J. Neurosci. *10*, 1415–1428. https://doi.org/10.1523/JNEUROSCI.10-05-01415.1990.

Masset, P., Ott, T., Lak, A., Hirokawa, J., and Kepecs, A. (2020). Behavior- and Modality-General Representation of Confidence in Orbitofrontal Cortex. Cell *182*, 112-126.e18. https://doi.org/10.1016/j.cell.2020.05.022.

Masterton, R.B., Bitterman, M.E., Campbell, C.B.G., and Hotton, N. (2018). Evolution of Brain and Behavior in Vertebrates (Routledge).

McConnell, S.K. (1988). Fates of visual cortical neurons in the ferret after isochronic and heterochronic transplantation. J Neurosci *8*, 945–974. .

McConnell, S.K., and Kaznowski, C.E. (1991a). Cell Cycle Dependence of Laminar Determination in Developing Neocortex. Science *254*, 282–285. https://doi.org/10.1126/science.1925583.

McConnell, S.K., and Kaznowski, C.E. (1991b). Cell Cycle Dependence of Laminar Determination in Developing Neocortex. Science *254*, 282–285. https://doi.org/10.1126/science.1925583.

McFadden, D. (2001). Economic Choices. American Economic Review *91*, 351–378. https://doi.org/10.1257/aer.91.3.351.

McGeorge, A.J., and Faull, R.L.M. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. Neuroscience *29*, 503–537. https://doi.org/10.1016/0306-4522(89)90128-0.

Migliore, M., and Shepherd, G.M. (2005). An integrated approach to classifying neuronal phenotypes. Nat Rev Neurosci *6*, 810–818. https://doi.org/10.1038/nrn1769.

Mikhael, J.G., Kim, H.R., Uchida, N., and Gershman, S.J. (2022). The role of state uncertainty in the dynamics of dopamine. Current Biology *32*, 1077-1087.e9. https://doi.org/10.1016/j.cub.2022.01.025.

Miller, E.K. (2000). The prefontral cortex and cognitive control. Nat Rev Neurosci *1*, 59–65. https://doi.org/10.1038/35036228.

Miller, J.A., Gouwens, N.W., Tasic, B., Collman, F., van Velthoven, C.T., Bakken, T.E., Hawrylycz, M.J., Zeng, H., Lein, E.S., and Bernard, A. (2020a). Common cell type nomenclature for the mammalian brain. ELife *9*, e59928. https://doi.org/10.7554/eLife.59928.

Miller, K.J., Botvinick, M.M., and Brody, C.D. (2020b). Value Representations in the Rodent Orbitofrontal Cortex Drive Learning, not Choice.

Mohebi, A., Pettibone, J.R., Hamid, A.A., Wong, J.-M.T., Vinson, L.T., Patriarchi, T., Tian, L., Kennedy, R.T., and Berke, J.D. (2019). Dissociable dopamine dynamics for learning and motivation. Nature *570*, 65–70. https://doi.org/10.1038/s41586-019-1235-y.

Molyneaux, B.J., Arlotta, P., Menezes, J.R.L., and Macklis, J.D. (2007). Neuronal subtype specification in the cerebral cortex. Nat Rev Neurosci *8*, 427–437. https://doi.org/10.1038/nrn2151.

Montague, P.R., Dolan, R.J., Friston, K.J., and Dayan, P. (2012). Computational psychiatry. Trends in Cognitive Sciences *16*, 72–80. https://doi.org/10.1016/j.tics.2011.11.018.

Morishima, M., Morita, K., Kubota, Y., and Kawaguchi, Y. (2011). Highly Differentiated Projection-Specific Cortical Subnetworks. J. Neurosci. *31*, 10380–10391. https://doi.org/10.1523/JNEUROSCI.0772-11.2011.

Mukamel, E.A., and Ngai, J. (2019). Perspectives on defining cell types in the brain. Curr Opin Neurobiol *56*, 61–68. https://doi.org/10.1016/j.conb.2018.11.007.

Muñoz-Castañeda, R., Zingg, B., Matho, K.S., Wang, Q., Chen, X., Foster, N.N., Narasimhan, A., Li, A., Hirokawa, K.E., Huo, B., et al. (2020). Cellular Anatomy of the Mouse Primary Motor Cortex.

Murphy, M.J.M., and Deutch, A.Y. (2018). Organization of afferents to the orbitofrontal cortex in the rat. J Comp Neurol *526*, 1498–1526. https://doi.org/10.1002/cne.24424.

Murugan, M., Jang, H.J., Park, M., Miller, E.M., Cox, J., Taliaferro, J.P., Parker, N.F., Bhave, V., Hur, H., Liang, Y., et al. (2017). Combined Social and Spatial Coding in a Descending Projection from the Prefrontal Cortex. Cell *171*, 1663-1677.e16. https://doi.org/10.1016/j.cell.2017.11.002.

Musall, S., Sun, X.R., Mohan, H., An, X., Gluf, S., Drewes, R., Osten, P., and Churchland, A.K. (2021). Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. 2021.09.27.461599. https://doi.org/10.1101/2021.09.27.461599.

Namboodiri, V.M.K., Otis, J.M., van Heeswijk, K., Voets, E.S., Alghorazi, R.A., Rodriguez-Romaguera, J., Mihalas, S., and Stuber, G.D. (2019). Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. Nat Neurosci *22*, 1110–1121. https://doi.org/10.1038/s41593-019-0408-1.

Naudé, J., Tolu, S., Dongelmans, M., Torquet, N., Valverde, S., Rodriguez, G., Pons, S., Maskos, U., Mourot, A., Marti, F., et al. (2016). Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking. Nat Neurosci *19*, 471–478. https://doi.org/10.1038/nn.4223.

Nectow, A.R., Moya, M.V., Ekstrand, M.I., Mousa, A., McGuire, K.L., Sferrazza, C.E., Field, B.C., Rabinowitz, G.S., Sawicka, K., Liang, Y., et al. (2017). Rapid Molecular Profiling of Defined Cell Types Using Viral TRAP. Cell Rep *19*, 655–667. https://doi.org/10.1016/j.celrep.2017.03.048.

Nieuwenhuys, R. (1994). The neocortex. An overview of its evolutionary development, structural organization and synaptology. Anat Embryol (Berl) *190*, 307–337. https://doi.org/10.1007/BF00187291.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci *4*, 95–102. https://doi.org/10.1038/82959.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature *508*, 207–214. https://doi.org/10.1038/nature13186.

Ohara, P.T., Granato, A., Moallem, T.M., Wang, B.-R., Tillet, Y., and Jasmin, L. (2003). Dopaminergic input to GABAergic neurons in the rostral agranular insular cortex of the rat. Journal of Neurocytology *32*, 131–141. .

O'Leary, D.D., and Stanfield, B.B. (1985). Occipital cortical neurons with transient pyramidal tract axons extend and maintain collaterals to subcortical but not intracortical targets. Brain Res *336*, 326–333. https://doi.org/10.1016/0006-8993(85)90661-4.

O'Leary, D.D.M., and Koester, S.E. (1993). Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. Neuron *10*, 991–1006. https://doi.org/10.1016/0896-6273(93)90049-W.

O'Leary, D.D.M., and Sahara, S. (2008). Genetic regulation of arealization of the neocortex. Curr Opin Neurobiol *18*, 90–100. https://doi.org/10.1016/j.conb.2008.05.011.

O'Leary, D.D.M., and Terashima, T. (1988). Cortical axons branch to multiple subcortical targets by interstitial axon budding: Implications for target recognition and "waiting periods." Neuron *1*, 901–910. https://doi.org/10.1016/0896-6273(88)90147-X.

O'Neill, M., and Schultz, W. (2010). Coding of Reward Risk by Orbitofrontal Neurons Is Mostly Distinct from Coding of Reward Value. Neuron *68*, 789–800. https://doi.org/10.1016/j.neuron.2010.09.031.

Otsuka, T., and Kawaguchi, Y. (2011). Cell Diversity and Connection Specificity between Callosal Projection Neurons in the Frontal Cortex. J. Neurosci. *31*, 3862–3870. https://doi.org/10.1523/JNEUROSCI.5795-10.2011.

Ott, T., Bosc, M., Sanders, J.L., Masset, P., and Kepecs, A. Near-optimal time allocation in proportion to confidence in humans, rats, and mice. In Preparation.

Padoa-Schioppa, C. (2009). Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. J. Neurosci. *29*, 14004–14014. https://doi.org/10.1523/JNEUROSCI.3751-09.2009.

Padoa-Schioppa, C. (2011). Neurobiology of Economic Choice: A Good-Based Model. Annual Review of Neuroscience *34*, 333–359. https://doi.org/10.1146/annurev-neuro-061010-113648.

Padoa-Schioppa, C., and Assad, J.A. (2006a). Neurons in the orbitofrontal cortex encode economic value. Nature *441*, 223–226. https://doi.org/10.1038/nature04676.

Padoa-Schioppa, C., and Assad, J.A. (2006b). Neurons in the orbitofrontal cortex encode economic value. Nature *441*, 223–226. https://doi.org/10.1038/nature04676.

Padoa-Schioppa, C., and Schoenbaum, G. (2015). Dialogue on economic choice, learning theory, and neuronal representations. Curr Opin Behav Sci *5*, 16–23. https://doi.org/10.1016/j.cobeha.2015.06.004.

Paolino, A., Fenlon, L.R., Suárez, R., and Richards, L.J. (2018). Transcriptional control of long-range cortical projections. Current Opinion in Neurobiology *53*, 57–65. https://doi.org/10.1016/j.conb.2018.05.005.

Pascoli, V., Hiver, A., Van Zessen, R., Loureiro, M., Achargui, R., Harada, M., Flakowski, J., and Lüscher, C. (2018). Stochastic synaptic plasticity underlying compulsion in a model of addiction. Nature *564*, 366–371. https://doi.org/10.1038/s41586-018-0789-4.

Pastor-Bernier, A., Stasiak, A., and Schultz, W. (2021). Reward-specific satiety affects subjective value signals in orbitofrontal cortex during multicomponent economic choice. Proceedings of the National Academy of Sciences *118*, e2022650118. https://doi.org/10.1073/pnas.2022650118.

Patriarchi, T., Cho, J.R., Merten, K., Howe, M.W., Marley, A., Xiong, W.-H., Folk, R.W., Broussard, G.J., Liang, R., Jang, M.J., et al. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. Science *360*, eaat4422. https://doi.org/10.1126/science.aat4422.

Patriarchi, T., Mohebi, A., Sun, J., Marley, A., Liang, R., Dong, C., Puhger, K., Mizuno, G.O., Davis, C.M., Wiltgen, B., et al. (2020). An expanded palette of dopamine sensors for multiplex imaging in vivo. Nat Methods *17*, 1147–1155. https://doi.org/10.1038/s41592-020-0936-3.

Petreanu, L., Mao, T., Sternson, S.M., and Svoboda, K. (2009). The subcellular organization of neocortical excitatory connections. Nature *457*, 1142–1145. https://doi.org/10.1038/nature07709.

Petrovic, P., and Castellanos, F.X. (2016). Top-Down Dysregulation—From ADHD to Emotional Instability. Frontiers in Behavioral Neuroscience *10*. .

Puumala, T., and Sirviö, J. (1998). Changes in activities of dopamine and serotonin systems in the frontal cortex underlie poor choice accuracy and impulsivity of rats in an attention task. Neuroscience *83*, 489–499. https://doi.org/10.1016/S0306-4522(97)00392-8.

Rakic, P. (1974). Neurons in Rhesus Monkey Visual Cortex: Systematic Relation between Time of Origin and Eventual Disposition. Science *183*, 425–427. https://doi.org/10.1126/science.183.4123.425.

Ramaswamy, S., and Markram, H. (2015). Anatomy and physiology of the thick-tufted layer 5 pyramidal neuron. Frontiers in Cellular Neuroscience *9*. .

Reiner, A. (1991). A Comparison of Neurotransmitter-Specific and Neuropeptide-Specific Neuronal Cell Types Present in the Dorsal Cortex in Turtles with Those Present in the Isocortex in Mammals: Implications for the Evolution of Isocortex; pp. 73–82. BBE *38*, 73–82. https://doi.org/10.1159/000114380.

Ren, J., Isakova, A., Friedmann, D., Zeng, J., Grutzner, S.M., Pun, A., Zhao, G.Q., Kolluru, S.S., Wang, R., Lin, R., et al. (2019). Single-cell transcriptomes and whole-brain projections of serotonin neurons in the mouse dorsal and median raphe nuclei. ELife *8*, e49424. https://doi.org/10.7554/eLife.49424.

Renteria, R., Baltz, E.T., and Gremel, C.M. (2018). Chronic alcohol exposure disrupts top-down control over basal ganglia action selection to produce habits. Nat Commun *9*, 211. https://doi.org/10.1038/s41467-017-02615-9.

Rigotti, M., Barak, O., Warden, M.R., Wang, X.-J., Daw, N.D., Miller, E.K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature *497*, 585–590. https://doi.org/10.1038/nature12160.

Robinson, D.L., Venton, B.J., Heien, M.L.A.V., and Wightman, R.M. (2003). Detecting Subsecond Dopamine Release with Fast-Scan Cyclic Voltammetry in Vivo. Clinical Chemistry *49*, 1763–1773. https://doi.org/10.1373/49.10.1763.

Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E., and Schoenbaum, G. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. J Neurosci *29*, 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009.

Rojas-Piloni, G., Guest, J.M., Egger, R., Johnson, A.S., Sakmann, B., and Oberlaender, M. (2017). Relationships between structure, in vivo function and long-range axonal target of cortical pyramidal tract neurons. Nat Commun *8*, 870. https://doi.org/10.1038/s41467-017-00971-0.

Rutledge, R.B., Lazzaro, S.C., Lau, B., Myers, C.E., Gluck, M.A., and Glimcher, P.W. (2009). Dopaminergic Drugs Modulate Learning Rates and Perseveration in Parkinson's Patients in a Dynamic Foraging Task. J. Neurosci. *29*, 15104–15114. https://doi.org/10.1523/JNEUROSCI.3524-09.2009.

Saddoris, M.P., Sugam, J.A., Stuber, G.D., Witten, I.B., Deisseroth, K., and Carelli, R.M. (2015). Mesolimbic Dopamine Dynamically Tracks, and Is Causally Linked to, Discrete Aspects of Value-Based Decision Making. Biological Psychiatry *77*, 903–911. https://doi.org/10.1016/j.biopsych.2014.10.024.

Sahni, V., Shnider, S.J., Jabaudon, D., Song, J.H.T., Itoh, Y., Greig, L.C., and Macklis, J.D. (2021a). Corticospinal neuron subpopulation-specific developmental genes prospectively indicate mature segmentally specific axon projection targeting. Cell Rep *37*, 109843. https://doi.org/10.1016/j.celrep.2021.109843.

Sahni, V., Itoh, Y., Shnider, S.J., and Macklis, J.D. (2021b). Crim1 and Kelch-like 14 exert complementary dual-directional developmental control over segmentally specific corticospinal axon projection targeting. Cell Reports *37*. https://doi.org/10.1016/j.celrep.2021.109842.

Sato, T.R., Gray, N.W., Mainen, Z.F., and Svoboda, K. (2007). The Functional Microarchitecture of the Mouse Barrel Cortex. PLOS Biology *5*, e189. https://doi.org/10.1371/journal.pbio.0050189.

Saunders, B.T., Richard, J.M., Margolis, E.B., and Janak, P.H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nature Neuroscience *21*, 1072–1083. https://doi.org/10.1038/s41593-018-0191-4.

Schlaggar, B.L., and O'Leary, D.D. (1991). Potential of visual cortex to develop an array of functional units unique to somatosensory cortex. Science *252*, 1556–1560. https://doi.org/10.1126/science.2047863.

Schmack, K., Bosc, M., Ott, T., Sturgill, J., and Kepecs, A. (2021). Striatal dopamine mediates hallucination-like perception in mice. Science *372*, eabf4740. https://doi.org/10.1126/science.abf4740.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599. https://doi.org/10.1126/science.275.5306.1593.

Schultz, W., Carelli, R.M., and Wightman, R.M. (2015). Phasic dopamine signals: from subjective reward value to formal economic utility. Current Opinion in Behavioral Sciences *5*, 147–154. https://doi.org/10.1016/j.cobeha.2015.09.006.

Senn, V., Wolff, S.B.E., Herry, C., Grenier, F., Ehrlich, I., Gründemann, J., Fadok, J.P., Müller, C., Letzkus, J.J., and Lüthi, A. (2014). Long-Range Connectivity Defines Behavioral Specificity of Amygdala Neurons. Neuron *81*, 428–437. https://doi.org/10.1016/j.neuron.2013.11.006.

Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement vigor as a reflection of subjective economic utility. Trends Neurosci *42*, 323–336. https://doi.org/10.1016/j.tins.2019.02.003.

Shai, A.S., Anastassiou, C.A., Larkum, M.E., and Koch, C. (2015). Physiology of Layer 5 Pyramidal Neurons in Mouse Primary Visual Cortex: Coincidence Detection through Bursting. PLOS Computational Biology *11*, e1004090. https://doi.org/10.1371/journal.pcbi.1004090.

Shepherd, G.M.G. (2013). Corticostriatal connectivity and its role in disease. Nat Rev Neurosci *14*, 278–291. https://doi.org/10.1038/nrn3469.

Shepherd, G.M., and Rowe, T.B. (2017). Neocortical Lamination: Insights from Neuron Types and Evolutionary Precursors. Front Neuroanat *11*, 100. https://doi.org/10.3389/fnana.2017.00100.

Shepherd, G.M.G., and Svoboda, K. (2005). Laminar and Columnar Organization of Ascending Excitatory Projections to Layer 2/3 Pyramidal Neurons in Rat Barrel Cortex. J. Neurosci. *25*, 5670–5679. https://doi.org/10.1523/JNEUROSCI.1173-05.2005.

Sherman, S.M., and Usrey, W.M. (2021). Cortical control of behavior and attention from an evolutionary perspective. Neuron *109*, 3048–3054. https://doi.org/10.1016/j.neuron.2021.06.021.

Sias, A.C., Morse, A.K., Wang, S., Greenfield, V.Y., Goodpaster, C.M., Wrenn, T.M., Wikenheiser, A.M., Holley, S.M., Cepeda, C., Levine, M.S., et al. (2021). A bidirectional corticoamygdala circuit for the encoding and retrieval of detailed reward memories. ELife *10*, e68617. https://doi.org/10.7554/eLife.68617.

Singer, B.F., Fadanelli, M., Kawa, A.B., and Robinson, T.E. (2018). Are Cocaine-Seeking "Habits" Necessary for the Development of Addiction-Like Behavior in Rats? J. Neurosci. *38*, 60–73. https://doi.org/10.1523/JNEUROSCI.2458-17.2017.

Sosulski, D.L., Bloom, M.L., Cutforth, T., Axel, R., and Datta, S.R. (2011). Distinct representations of olfactory information in different cortical centres. Nature *472*, 213–216. https://doi.org/10.1038/nature09868.

Soudais, C., Laplace-Builhe, C., Kissa, K., and Kremer, E.J. (2001). Preferential transduction of neurons by canine adenovirus vectors and their efficient retrograde transport in vivo. FASEB J *15*, 2283–2285. https://doi.org/10.1096/fj.01-0321fje.

Spellman, T., Svei, M., Kaminsky, J., Manzano-Nieves, G., and Liston, C. (2021). Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. Cell *184*, 2750-2766.e17. https://doi.org/10.1016/j.cell.2021.03.047.

Stanfield, B.B. (1992). The development of the corticospinal projection. Progress in Neurobiology *38*, 169–202. https://doi.org/10.1016/0301-0082(92)90039-H.

Stanfield, B.B., and O'Leary, D.D.M. (1985). Fetal occipital cortical neurons transplanted to the rostral cortex can extend and maintain a pyramidal tract axon. Nature *313*, 135–137. https://doi.org/10.1038/313135a0.

Stanfield, B.B., O'Leary, D.D.M., and Fricks, C. (1982). Selective collateral elimination in early postnatal development restricts cortical distribution of rat pyramidal tract neurones. Nature *298*, 371–373. https://doi.org/10.1038/298371a0.

Stauffer, W.R., Lak, A., Kobayashi, S., and Schultz, W. (2016). Components and characteristics of the dopamine reward utility signal. Journal of Comparative Neurology *524*, 1699–1711. https://doi.org/10.1002/cne.23880.

Steinberg, E.E., Christoffel, D.J., Deisseroth, K., and Malenka, R.C. (2015). Illuminating circuitry relevant to psychiatric disorders with optogenetics. Current Opinion in Neurobiology *30*, 9–16. https://doi.org/10.1016/j.conb.2014.08.004.

Steiner, A.P., and Redish, A.D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. Nat Neurosci *17*, 995–1002. https://doi.org/10.1038/nn.3740.

Stouffer, M.A., Woods, C.A., Patel, J.C., Lee, C.R., Witkovsky, P., Bao, L., Machold, R.P., Jones, K.T., de Vaca, S.C., Reith, M.E.A., et al. (2015). Insulin enhances striatal dopamine release by activating cholinergic interneurons and thereby signals reward. Nat Commun *6*, 8543. https://doi.org/10.1038/ncomms9543.

Suryanarayana, S.M., Robertson, B., Wallén, P., and Grillner, S. (2017). The Lamprey Pallium Provides a Blueprint of the Mammalian Layered Cortex. Current Biology *27*, 3264-3277.e5. https://doi.org/10.1016/j.cub.2017.09.034.

Sutton, R. (1992). Gain Adaptation Beats Least Squares? In In Proceedings of the 7th Yale Workshop on Adaptive and Learning Systems, pp. 161–166.

Sutton, R.S., and Barto, A.G. (2018). Reinforcement Learning, second edition: An Introduction (MIT Press).

Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A., and Schoenbaum, G. (2009a). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. Neuron *62*, 269–280. https://doi.org/10.1016/j.neuron.2009.03.005.

Takahashi, Y.K., Roesch, M.R., Stalnaker, T.A., Haney, R.Z., Calu, D.J., Taylor, A.R., Burke, K.A., and Schoenbaum, G. (2009b). The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. Neuron *62*, 269–280. https://doi.org/10.1016/J.NEURON.2009.03.005.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci *14*, 1590–1597. https://doi.org/10.1038/nn.2957.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci *19*, 335–346. https://doi.org/10.1038/nn.4216.

Terra, H., Bruinsma, B., de Kloet, S.F., van der Roest, M., Pattij, T., and Mansvelder, H.D. (2020). Prefrontal Cortical Projection Neurons Targeting Dorsomedial Striatum Control Behavioral Inhibition. Curr Biol *30*, 4188-4200.e5. https://doi.org/10.1016/j.cub.2020.08.031.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive Coding of Reward Value by Dopamine Neurons. Science *307*, 1642–1645. https://doi.org/10.1126/science.1105370.

Tosches, M.A., and Laurent, G. (2019). Evolution of neuronal identity in the cerebral cortex. Current Opinion in Neurobiology *56*, 199–208. https://doi.org/10.1016/j.conb.2019.04.009.

Tosches, M.A., Yamawaki, T.M., Naumann, R.K., Jacobi, A.A., Tushev, G., and Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. Science *360*, 881–888. https://doi.org/10.1126/science.aar4237.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. Nature *398*, 704–708. https://doi.org/10.1038/19525.

Usoskin, D., Furlan, A., Islam, S., Abdo, H., Lönnerberg, P., Lou, D., Hjerling-Leffler, J., Haeggström, J., Kharchenko, O., Kharchenko, P.V., et al. (2015). Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nat Neurosci *18*, 145–153. https://doi.org/10.1038/nn.3881.

Vandaele, Y., Cantin, L., Serre, F., Vouillac-Mendoza, C., and Ahmed, S.H. (2016). Choosing Under the Influence: A Drug-Specific Mechanism by Which the Setting Controls Drug Choices in Rats. Neuropsychopharmacology *41*, 646–657. https://doi.org/10.1038/npp.2015.195.

Varazzani, C., San-Galli, A., Gilardeau, S., and Bouret, S. (2015). Noradrenaline and Dopamine Neurons in the Reward/Effort Trade-Off: A Direct Electrophysiological Comparison in Behaving Monkeys. J. Neurosci. *35*, 7866–7877. https://doi.org/10.1523/JNEUROSCI.0454-15.2015.

Volkow, N.D., Fowler, J.S., and Wang, G.-J. (2003). The addicted human brain: insights from imaging studies. J Clin Invest *111*, 1444–1451. https://doi.org/10.1172/JCI200318533.

Volkow, N.D., Tomasi, D., Wang, G.-J., Fowler, J.S., Telang, F., Goldstein, R.Z., Alia-Klein, N., Woicik, P., Wong, C., Logan, J., et al. (2011). Positive emotionality is associated with baseline metabolism in orbitofrontal cortex and in regions of the default network. Mol Psychiatry *16*, 818–825. https://doi.org/10.1038/mp.2011.30.

Vyas, S., Golub, M.D., Sussillo, D., and Shenoy, K.V. (2020). Computation Through Neural Population Dynamics. Annual Review of Neuroscience *43*, 249–275. https://doi.org/10.1146/annurev-neuro-092619-094115.

Wang, Z., and McCormick, D.A. (1993). Control of firing mode of corticotectal and corticopontine layer V burst-generating neurons by norepinephrine, acetylcholine, and 1S,3R-ACPD. J Neurosci *13*, 2199–2216. .

Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A., and Uchida, N. (2012). Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. Neuron *74*, 858–873. https://doi.org/10.1016/J.NEURON.2012.03.017.

Wei, W., Mohebi, A., and Berke, J.D. (2021). Striatal dopamine pulses follow a temporal discounting spectrum. 2021.10.31.466705. https://doi.org/10.1101/2021.10.31.466705.

Weimann, J.M., Zhang, Y.A., Levin, M.E., Devine, W.P., Brûlet, P., and McConnell, S.K. (1999). Cortical Neurons Require Otx1 for the Refinement of Exuberant Axonal Projections to Subcortical Targets. Neuron *24*, 819–831. https://doi.org/10.1016/S0896-6273(00)81030-2.

Weissbourd, B., Momose, T., Nair, A., Kennedy, A., Hunt, B., and Anderson, D.J. (2021). A genetically tractable jellyfish model for systems and evolutionary neuroscience. Cell *184*, 5854-5868.e20. https://doi.org/10.1016/j.cell.2021.10.021.

Wilson, C.J. (1987). Morphology and synaptic connections of crossed corticostriatal neurons in the rat. J Comp Neurol *263*, 567–580. https://doi.org/10.1002/cne.902630408.

Winnubst, J., Bas, E., Ferreira, T.A., Wu, Z., Economo, M.N., Edson, P., Arthur, B.J., Bruns, C., Rokicki, K., Schauder, D., et al. (2019). Reconstruction of 1,000 Projection Neurons Reveals New Cell Types and Organization of Long-Range Connectivity in the Mouse Brain. Cell *179*, 268-281.e13. https://doi.org/10.1016/j.cell.2019.07.042.

Wise, R.A. (2004). Dopamine, learning and motivation. Nat Rev Neurosci *5*, 483–494. https://doi.org/10.1038/nrn1406.

Wise, R.A., and Rompre, P.-P. (1989). Brain Dopamine and Reward. Annual Review of Psychology *40*, 191–225. https://doi.org/10.1146/annurev.ps.40.020189.001203.

Wise, S.P., and Jones, E.G. (1977). Cells of origin and terminal distribution of descending projections of the rat somatic sensory cortex. J Comp Neurol *175*, 129–157. https://doi.org/10.1002/cne.901750202.

Woodworth, M.B., Greig, L.C., Liu, K.X., Ippolito, G.C., Tucker, H.O., and Macklis, J.D. (2016). Ctip1 Regulates the Balance between Specification of Distinct Projection Neuron Subtypes in Deep Cortical Layers. Cell Reports *15*, 999–1012. https://doi.org/10.1016/j.celrep.2016.03.064.

Wyvell, C.L., and Berridge, K.C. (2000). Intra-Accumbens Amphetamine Increases the Conditioned Incentive Salience of Sucrose Reward: Enhancement of Reward "Wanting" without Enhanced "Liking" or Response Reinforcement. J. Neurosci. *20*, 8122–8130. https://doi.org/10.1523/JNEUROSCI.20-21-08122.2000.

Yang, G.R., and Molano-Mazón, M. (2021). Towards the next generation of recurrent network models for cognitive neuroscience. Curr Opin Neurobiol *70*, 182–192. https://doi.org/10.1016/j.conb.2021.10.015.

Yao, Z., Liu, H., Xie, F., Fischer, S., Adkins, R.S., Aldridge, A.I., Ament, S.A., Bartlett, A., Behrens, M.M., Van den Berge, K., et al. (2021). A transcriptomic and epigenomic cell atlas of the mouse primary motor cortex. Nature *598*, 103–110. https://doi.org/10.1038/s41586-021-03500-8.

Yuste, R. (2011). Dendritic Spines and Distributed Circuits. Neuron *71*, 772–781. https://doi.org/10.1016/j.neuron.2011.07.024.

Zador, A.M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. Nat Commun *10*, 3770. https://doi.org/10.1038/s41467-019-11786-6.

Zarrinpar, A., and Callaway, E.M. (2016). Functional Local Input to Layer 5 Pyramidal Neurons in the Rat Visual Cortex. Cereb Cortex *26*, 991–1003. https://doi.org/10.1093/cercor/bhu268.

Zeeb, F.D., Robbins, T.W., and Winstanley, C.A. (2009). Serotonergic and Dopaminergic Modulation of Gambling Behavior as Assessed Using a Novel Rat Gambling Task. Neuropsychopharmacol *34*, 2329–2343. https://doi.org/10.1038/npp.2009.62.

Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nat Rev Neurosci *18*, 530–546. https://doi.org/10.1038/nrn.2017.85.

Zeng, H., Shen, E.H., Hohmann, J.G., Oh, W.S., Bernard, A., Royall, J.J., Glattfelder, K.J., Sunkin, S.M., Morris, J.A., Guillozet-Bongaarts, A.L., et al. (2012). Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. Cell *149*, 483–496. https://doi.org/10.1016/j.cell.2012.02.052.

Zhang, M., Eichhorn, S.W., Zingg, B., Yao, Z., Cotter, K., Zeng, H., Dong, H., and Zhuang, X. (2021). Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. Nature *598*, 137–143. https://doi.org/10.1038/s41586-021-03705-x.

Zingg, B., Hintiryan, H., Gou, L., Song, M.Y., Bay, M., Bienkowski, M.S., Foster, N.N., Yamashita, S., Bowman, I., Toga, A.W., et al. (2014). Neural networks of the mouse neocortex. Cell *156*, 1096–1111. https://doi.org/10.1016/j.cell.2014.02.023.