Bayesian mapping of the striatal microcircuit reveals robust asymmetries in the probabilities and distances of connections

François Cinotti[1,2] and Mark D. Humphries[1,3]

[1]**School of Psychology, University of Nottingham**
[2]**francois.cinotti@gmail.com**
[3]**mark.humphries@nottingham.ac.uk**

ABSTRACT

The striatum's complex microcircuit is made by connections within and between its D1- and D2-receptor expressing projection neurons and at least five species of interneuron. Precise knowledge of this circuit is likely essential to understanding striatum's functional roles and its dysfunction in a wide range of movement and cognitive disorders. We introduce here a Bayesian approach to mapping neuron connectivity using intracellular recording data, which lets us simultaneously evaluate the probability of connection between neuron types, the strength of evidence for it, and its dependence on distance. Using it to synthesise a complete map of the rodent striatum, we find strong evidence for two asymmetries: a selective asymmetry of projection neuron connections, with D2 neurons connecting twice as densely to other projection neurons than do D1 neurons, but neither subtype preferentially connecting to another; and a length-scale asymmetry, with interneuron connection probabilities remaining non-negligible at more than twice the distance of projection neuron connections. We further show our Bayesian approach can evaluate evidence for wiring changes, using data from the developing striatum and a mouse model of Huntington's disease. By quantifying the uncertainty in our knowledge of the microcircuit, our approach reveals a wide range of potential striatal wiring diagrams consistent with current data.

Keywords: Striatum, Dopamine receptor, Spiny Projection Neuron, medium spiny neuron, Bayesian inference, Direct pathway, Indirect pathway

INTRODUCTION

As the input of the basal ganglia circuit, the striatum has been ascribed key computational roles in action selection (Redgrave et al., 1999; Gurney et al., 2001a; Liénard and Girard, 2014), decision making (Bogacz and Gurney, 2007; Ding and Gold, 2010, 2012; Yartsev et al., 2018), and reinforcement learning (Reynolds et al., 2001; Samejima et al., 2005; Bornstein and Daw, 2011; Khamassi and Humphries, 2012; Gurney et al., 2015). Within the striatum is a microcircuit comprising the GABAergic spiny projection neurons (SPNs), the most common type of neuron in the striatum making up as much as 97% of neurons in the rat (Oorschot, 2012), and at least five species of predominantly GABAergic interneurons (Burke et al., 2017; Tepper et al., 2018). These SPNs divide into two distinct populations that express either the D1 or D2-type of dopamine receptors (Gerfen et al., 1990; Gerfen and Surmeier, 2011). Projections from the D1 and D2 SPN populations respectively form the striatonigral and striatopallidal pathways (Gerfen et al., 1990; Kreitzer, 2009; Gerfen and Surmeier, 2011), through which they influence dynamics throughout the basal ganglia and beyond. The connections made onto the D1 and D2 SPNs by the microcircuit are therefore positioned as a potentially major actor in sculpting the output of this nucleus, and thus the computations ascribed to it.

One key to understanding the role of the microcircuit in the computations of striatum is knowing the relative influence of one neuron type on another. Two broad influences of this microcircuit on the output of SPNs are well-known: the feedforward inhibition by GABAergic interneurons, and feedback inhibition in the form of lateral connections between the SPNs (Plenz, 2003; Tepper et al., 2004, 2008;

Humphries et al., 2010). Numerous models have shown how specific elements of the microcircuit play powerful roles in shaping striatal dynamics, including the connections between SPNs (Wickens et al., 1991; Alexander and Wickens, 1993; Humphries et al., 2009; Ponzi and Wickens, 2010, 2013; Stocco and Lebiere, 2014; Tomkins et al., 2014; Spreizer et al., 2017), the connections from interneurons to the SPNs (Humphries et al., 2009; Damodaran et al., 2014; Franklin and Frank, 2015), and the connections between interneurons (Hjorth et al., 2009; Lau et al., 2010; Klaus et al., 2011). But to understand how all elements of the striatum's microcircuit influence its output requires a full account of the microcircuit's wiring, which we currently lack. To address this problem, here we synthesise data from pairwise intracellular recording studies to generate a statistically-rigorous and comprehensive map of the wiring probabilities between the key neuron species of the rodent striatum.

A key issue in estimating connection densities from intracellular recording data is that recording studies report a single probability for each connection type, given by the rate of successful connections between two types of neuron. For example, there is a claimed asymmetry in the probability of connections between the D1 and D2 SPN populations, based on the rate at which D1 and D2 SPNs respond to stimulation of the other SPN type (Taverna et al., 2008; Planert et al., 2010), upon which theories and models of the striatum have been constructed (Bahuguna et al., 2015; Burke et al., 2017; Hjorth et al., 2020). But these estimates lack any measure of uncertainty, and thus it is unclear how robust are conclusions about the striatal microcircuit drawn from these data, and thus how robust are theories and models based on those pairwise estimates. Further, without such measures of uncertainty we do not know where further data is needed.

In this paper, we solve these problems by introducing a Bayesian approach to estimating the probability of connection between neuron types using pairwise intracellular recording data, which allows us to draw rigorous conclusions about the strength of evidence for claims about the microcircuit. Using this approach on data from the rodent striatum, we show that the previously reported asymmetry between the rates at which D1 and D2 neurons make connections is robust, with D2 SPNs having roughly twice the connection rate of D1 SPNs; but contrary to previous claims we also show that there is no evidence for an asymmetry in the rates at which they receive connections, and so there is no preferential target for D1 or D2 SPNs. We then demonstrate a new method for extracting information about the effect of distance using single estimates of connection rates, allowing us to estimate distance-dependent probabilities and their uncertainty. Using these methods to analyse both SPN and interneuron connectivity, we complete our Bayesian map of the connectivity of the rodent striatum. Finally, we demonstrate how our Bayesian approach lets us quantify and test changes to that microcircuit map: we test the claim that D1 SPN connections are altered in a mouse model of Huntington's disease, and find no evidence for it; and, using recent data from Krajeski et al. (2019), we show that the selective asymmetry of D1 and D2 SPNs appears at different stages during development. Such a Bayesian approach thus simultaneously evaluates the probability of connection between neuron types, its dependence on distance, and the strength of evidence for it, creating a solid foundation for theories of striatal computation.

RESULTS

Patch clamp data on connection rates between SPNs

We begin by reviewing key data on the connections within and between the D1- and D2-type SPNs, which we will also use to motivate our Bayesian approach. Previous studies by Taverna et al. (2008) and Planert et al. (2010) collected data on pairwise connections between SPN subtypes in slices obtained from both the dorsal and ventral striatum of mice. The subtype of the SPNs was determined by targeted expression of EGFP under the control of either a D1 or D2 receptor promoter sequence for Taverna et al. (2008) and of only a D1 receptor promoter for Planert et al. (2010). Non-labelled SPNs were then assumed to belong to whichever group was not meant to be labelled in this particular animal and electrophysiological criteria were used to exclude interneurons. After choosing a pair of neighbouring cells, no further apart than 50 μm, Taverna et al. (2008) detected connections in current-clamp mode by injecting a depolarising current step in the first – potentially presynaptic – neuron of the pair, and measuring depolarising postsynaptic potentials in the second one. This procedure was then repeated the other way around. Planert et al. (2010) recorded up to four neighbouring neurons simultaneously, with a larger maximum intersomatic distance of 100 μm, and detected connections by stimulating one neuron with a train of depolarising pulses to generate action potentials in the presynaptic neuron and recording the potential postsynaptic responses of the other neurons.

In both studies, the ratio of successful tests to the total number of tests (Table 1) was then reported as an estimate $\hat{p}$ for the true probability of connection between these types of neuron (plotted as the bars in Figure 1A and B), in accordance with frequentist inference about a proportion. As they are estimates, they come with a level of uncertainty about the true proportion that depends on sample size, which here is the number of pairs that were tested. In a frequentist approach, this uncertainty would usually be given by a confidence interval.

However, typically, intracellular recording studies do not report any estimate for the uncertainty surrounding their measurements of $\hat{p}$, and recent theoretical studies of striatum (Burke et al., 2017; Hjorth et al., 2020; Bahuguna et al., 2015) have simply used the raw values $\hat{p}$ from Taverna et al. (2008) and Planert et al. (2010) to construct their models, supposing in particular that the probability of D1 to D1 connections is about twice as likely as D1 to D2 connections. When we do compute confidence intervals, such as the Wilson confidence interval for binomial proportions (Brown et al., 2001) that we add ourselves in Figure 1A and B, we find that, given the relatively small sample sizes, the confidence intervals overlap quite significantly.

Bayesian inference of connection probabilities

As we have just explained, frequentist inference gives us a single point estimate $\hat{p}$ for the probability of connection, normally surrounded by a confidence interval which may be too large to be of any practical use and also, because it is flat, may give the illusion that the true value of p might be anywhere within this interval with equal probability. By contrast, Bayesian inference is more informative because it gives us a full probability density function $f_p(p)$, called the posterior, telling us exactly how likely every possible value of p actually is, given the collected data. In this way, even when confidence intervals overlap as is the case for practically all the SPN to SPN connections here (Figure 1A and B), which in a frequentist interpretation would lead us to dismiss the difference as non-significant without insight as to whether this is due to insufficient data or a true non-difference (Dienes, 2014; Makin and De Xivry, 2019), Bayesian inference gives us a much clearer picture of what the data can tell us. We introduce here a simple Bayesian approach to calculating the full posterior $f_p(p)$ for each type of connection from any pair-wise intracellular recording data.

In our approach, the random variable of interest, p, is the success rate of a binomial distribution describing the probability of an experimental result when testing pair-wise connections:

$$P(X = k|p) = \binom{n}{k} p^k (1-p)^{n-k} \tag{1}$$

where k is the number of connected pairs and n the total number of tested pairs of that type (see Table 1 for a full summary of the experimental data). In this way, the binomial distribution provides a likelihood for the data given p. According to Bayes theorem, the posterior distribution can be determined by:

$$f_{posterior}(p) \propto P(X = k|p) f_{prior}(p) \tag{2}$$

given a prior $f_{prior}(p)$, which is a probability distribution describing our initial beliefs about the possible value of p. Finding a posterior for the success rate of a binomial distribution is a well known problem in Bayesian inference and the prior distribution used in this case is a beta distribution:

$$f_{prior}(p; a, b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} \tag{3}$$

with a and b the parameters determining the shape of the prior, and $B(a,b)$ the so-called Beta function. The main advantage of this type of prior is that it is the so-called conjugate prior of binomial distributions, which means the posterior that results from combining this prior with a likelihood in the form of a binomial distribution simply turns out to be a new beta distribution with updated parameters (sparing us the trouble of renormalising the right hand side of equation 2 to get a proper probability density function):

$$f_{posterior}(p) = f_{prior}(p; a+k, b+n-k) \tag{4}$$

In other words, to determine the posterior we simply have to add the number of successful tests k to a, and the number of unsuccessful tests $n-k$ to b. For instance, if we had initial values of a and b equal to 1, and we were looking at the data concerning D1 $\rightarrow$ D1 connections obtained by Taverna et al. (2008) (see

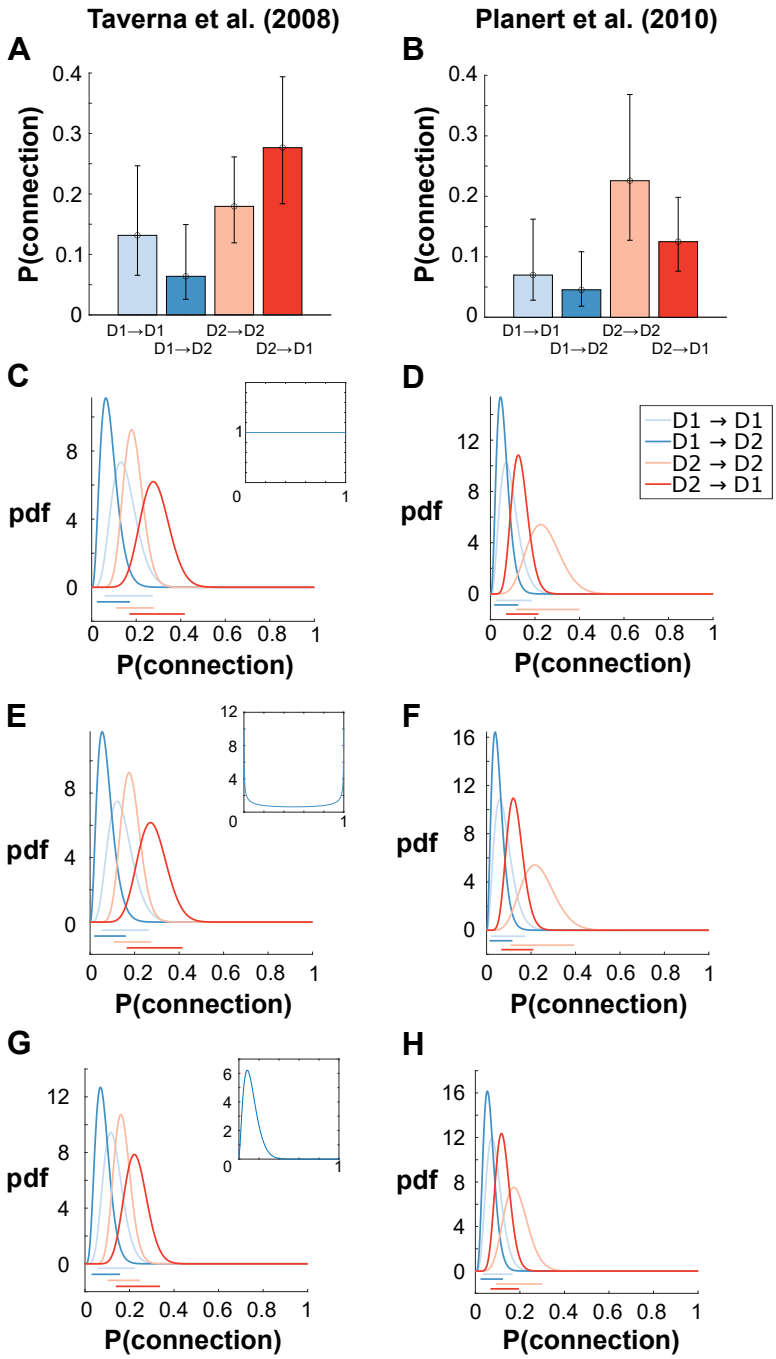

Figure 1. Probability of lateral connections between SPNs estimated using either frequentist or Bayesian methods. **A-B** Frequentist estimates of the probabilities of connection computed from intracellular recording data, and our computed 95% Wilson confidence intervals. **C-D** Posterior probability density functions for the probability of connection using a Bayesian approach. Coloured bars underneath the plot represent the 95% credibility intervals corresponding to each probability density function. Inset: shape of the prior, a uniform distribution. **E-F** Posterior probability density functions using the Jeffreys prior. **G-H** Posterior probability density functions using a prior based on previous literature with mean equal to 0.12 and variance equal to 0.005.

Study	Pair	k	n	a	b	$\hat{p}_{MAP}$	95% credibility interval
Taverna et al. (2008)	D1 SPN → D1 SPN	5	38	7.56	51.12	0.116	[0.057, 0.225]
	D1 SPN → D2 SPN	3	47	5.56	62.12	0.069	[0.030, 0.158]
	D2 SPN → D1 SPN	13	47	15.56	52.12	0.222	[0.138, 0.336]
	D2 SPN → D2 SPN	14	78	16.56	82.12	0.161	[0.101, 0.247]
	D1 SPN → SPN	8	85	10.56	95.12	0.092	[0.051, 0.164]
	D2 SPN → SPN	27	125	29.56	1116.12	0.199	[0.142, 0.272]
Planert et al. (2010)	D1 SPN → D1 SPN	3	43	5.56	58.12	0.074	[0.032, 0.167]
	D1 SPN → D2 SPN	3	66	5.56	81.12	0.054	[0.023, 0.124]
	D2 SPN → D1 SPN	10	80	12.56	88.12	0.117	[0.068, 0.196]
	D2 SPN → D2 SPN	7	31	9.56	42.12	0.172	[0.093, 0.300]
	D1 SPN → SPN	6	109	8.56	121.12	0.059	[0.030, 0.114]
	D2 SPN → SPN	17	111	19.56	112.12	0.143	[0.093, 0.214]
	FS → D1 SPN	8	9	9	2	0.889	[0.555, 0.975]
	FS → D2 SPN	6	9	7	4	0.667	[0.348, 0.878]
Gittis et al. (2010)	FS → D1 SPN	48	90	49	43	0.533	[0.431, 0.633]
	FS → D2 SPN	27	77	28	51	0.351	[0.253, 0.462]
	FS → FS	7	12	8	6	0.583	[0.316, 0.808]
	FS → PLTS *	2	21	3	20	0.095	[0.029, 0.292]
	FS → Ach	0	3	1	4	0	[0, 0.602]
	PLTS → MSN	2	60	3	59	0.033	[0.010, 0.114]
	PLTS → PLTS	0	26	1	27	0	[0, 0.13]
	PLTS → FS *	0	20	1	21	0	[0, 0.161]
	PLTS → Ach	0	10	1	11	0	[0, 0.285]
Dorst et al. (2020)	TH → Ach	13	50	14	38	0.260	[0.159, 0.396]
	Ach → TH	11	41	12	31	0.268	[0.157, 0.420]
Ibáñez-Sandoval et al. (2011)	NGF → SPN	25	29	26	5	0.862	[0.693, 0.944]
English et al. (2012)	Ach → NGF	8	14	9	7	0.571	[0.323, 0.787]
	NGF → Ach	3	14	4	12	0.214	[0.078, 0.481]

Table 1. Pairwise connection data used to build the Bayesian map of the striatum microcircuit, alongside Bayesian estimates of the connection probabilities. k: number of connected pairs in that study; n number of sampled pairs; a, b parameters of the resulting Beta distribution for the posterior of p using either the literature prior for SPN connections or a uniform prior for interneurons as explained in the main text; $\hat{p}_{MAP}$, the MAP estimate of p.
Abbreviations for neuron names: SPN : Spiny Projection Neuron; FS: Fast Spiking interneuron, PLTS: Persistent Low Threshold Spiking interneuron; Ach: Cholinergic interneuron; TH: Tyrosyne-Hydroxylase interneuron; NGF: (NPY-expressing) NeuroGliaForm interneuron.
* Data concerning FS → PLTS and PLTS → FS connections from Gittis et al. (2010) was pooled with that of Szydlowski et al. (2013).

Table 1), a would become equal to 6 and b to 34, and the shape of the resulting posterior would be the one depicted in Figure 1C (light blue curve). Depending on our assumptions, different values of a and b can be used to give the prior a desired shape. We begin with the common choice of the uniform distribution in which p could be anywhere between 0 and 1 with equal probability, achieved by setting $a = b = 1$ as in the example just given.

Study	pair	uniform	Jeffreys	literature
Taverna et al. (2008)	$D1 \to D1$	0.132	0.122	0.116
	$D1 \to D2$	0.064	0.054	0.069
	$D2 \to D1$	0.277	0.272	0.222
	$D2 \to D2$	0.179	0.175	0.161
Planert et al. (2010)	$D1 \to D1$	0.070	0.060	0.074
	$D1 \to D2$	0.045	0.038	0.054
	$D2 \to D1$	0.125	0.120	0.117
	$D2 \to D2$	0.226	0.217	0.172

Table 2. Maximum A Posteriori (MAP) estimates for the different probabilities of connection between SPNs, using different priors and different experimental studies. The "literature" prior is based on data on pairwise connections from intracellular recording studies that predated techniques for identifying types of SPN, as explained in the main text.

Using this prior in combination with the data of Taverna et al. (2008) gives us the posterior curves shown in Figure 1C. Once obtained, we can revert, if necessary, to a more frequentist standpoint by extracting from these density functions a single point estimate $\hat{p}$, typically the Maximum A Posteriori value or MAP which is simply the value of p for which $f_p(p)$ is maximum, and a credibility interval around that MAP, which is the Bayesian equivalent of a confidence interval. The MAPs concentrate around relatively low values of p and their exact values, which are given in Table 2, lie between 0.06 for D1 $\to$ D2 pairs and 0.28 for D2 $\to$ D1 pairs. The uncertainty surrounding p is given by the width of the posteriors and of the 95% credibility intervals underneath the curves. For the data from Taverna et al. (2008), the connections with the smallest credibility interval of about 0.1 are the D1 $\to$ D2 pairs while the least well resolved connections are the D2 $\to$ D1 pairs for which the credibility interval spans roughly 0.2. By contrast, when we apply the uniform prior to the data of Planert et al. (2010), the D2 $\to$ D1 connections have the smallest uncertainty and the D2 $\to$ D2 connections the largest (Figure 1D).

The fact that we used the same prior for all pairs of neuron types reflects our initial belief that there is no difference in the probability of connection between pairs. To overcome this belief requires a sufficient amount of evidence, and we can start comparing the different probabilities of connection visually by looking at how much the different posteriors overlap. Based on Figure 1C for example, it seems that in the data of Taverna et al. (2008) probabilities of connection segregate depending on the nature of the presynaptic neuron in the pair. The posteriors involving presynaptic D1 neurons overlap considerably with one another, and their region of highest density is lower than for presynaptic D2 neurons who also show great overlap, while there is much less overlap between pairs with different presynaptic neurons. This becomes even more obvious when looking at the 95% credibility intervals drawn underneath the curves which show more or less overlap depending on the nature of the presynaptic neuron: the credibility intervals for pairs with a presynaptic D1 neuron share an overlapping interval roughly covering probabilities of 0.05 to 0.15, while the overlapping interval for connections with a presynaptic D2 SPN ranges between probabilities of about 0.20 to 0.28. A similar pattern repeats itself in the data of Planert et al. (2010) (Figure 1D) although the exact values of the overlapping regions are shifted towards 0 compared to Taverna et al. (2008), an effect which is potentially due to maximum distance of sampling as explained later. This opens the possibility that there is indeed an asymmetry in terms of probability of connection that is dependent on the subtype of the presynaptic neuron, with no or little effect of the postsynaptic target subtype, something we will explore more thoroughly later.

One of the main advantages of Bayesian inference is that it forces researchers to be explicit about their priors and gives them the opportunity to choose appropriate ones. In order to illustrate this, we applied three further priors to the experimental data. Firstly, the so-called non-informative Jeffreys prior sets $a = b = 1/2$. An intuitive way of understanding this prior is to picture ourselves at the very beginning of the experiment, waiting for the result of the very first paired stimulation and recording. This test will

either be successful or not, meaning that the shape of the prior should give most and equal weight to these two outcomes (inset of Figure 1E). Figures 1E and F show the posteriors that result from using this prior and we can see how they are practically identical to the posteriors obtained with a uniform prior. This was also the case when using the Haldane prior for which a and b equal 0 (not shown).

Our third prior is based on prior data, for Bayesian inference also provides us with a principled way of integrating previous knowledge into the prior. Earlier work (Taverna et al., 2004; Czubayko and Plenz, 2002; Koos et al., 2004) quantified the rate of lateral connections between SPNs without distinguishing SPN subtypes and concluded that lateral connections occurred at a rate of about 0.12. Using this information, we can design a beta distribution with a mean of 0.12 and an arbitrary variance of 0.005 (see Methods) which serves as our third and final prior shown in the inset of Figure 1G. Although the posteriors are more clearly different from the ones obtained with the uniform and Jeffreys prior they still look very similar. In fact, we find that the MAP values for a given type of connection (Table 2) are very close whatever the choice of prior, and the previous observation that the curves seem to segregate according to the subtype of the presynaptic neuron is valid in every case. On the other hand, because this prior is more informative than the two previous ones, uncertainty is reduced, as witnessed by the smaller credibility intervals. Given this robustness to the different priors, we shall henceforth exclusively use the prior based on previous literature when analysing connections between SPNs.

D1 neurons make fewer connections than D2 neurons

We previously observed that D1 neurons seem to make fewer connections than D2 neurons without necessarily targeting one subtype over the other, based on how the posterior distributions appear to segregate by presynaptic subtype in Figure 1. We can go beyond this qualitative analysis by calculating a density function f_Δ for the difference between two probabilities of connection. For instance, if we're interested in the difference in the probability of connection between D1 $\rightarrow$ D1 and D1 $\rightarrow$ D2 pairs (Figure 2A), using the posterior distributions $f_{D1\rightarrow D1}(p)$ and $f_{D1\rightarrow D2}(p)$, we can find the density function for $\Delta_{(D1\rightarrow D1)-(D1\rightarrow D2)}$ in the following manner:

$$f_{\Delta_{(D1\rightarrow D1)-(D1\rightarrow D2)}}(\Delta = k) = \int_{max(0,-k)}^{min(1,1-k)} f_{D1\rightarrow D1}(p = x+k) f_{D1\rightarrow D2}(p = x)dx \tag{5}$$

with the bounds of the integral such that $x+k$ lies between 0 and 1. We can then calculate the probability that $\Delta_{(D1\rightarrow D1)-(D1\rightarrow D2)}$ is smaller than 0 by integrating this distribution between -1 and 0 (or calculate if it is greater than 0 by integrating the distribution between 0 and 1). By contrast, the frequentist strategy would be to compute a p-value giving the probability of getting an experimental result at least as extreme as the one observed assuming the null hypothesis of no difference in connection probabilities (i.e. $\Delta = 0$). By contrast, the Bayesian approach allows us to calculate the probability that Δ is less than (or greater than) 0 given experimental results. Thus whereas the p-value tells us how surprising the actual data is if we accept the null hypothesis, the Bayesian approach can quantify precisely how unlikely the null hypothesis actually is.

A first application of this method is to answer the question: do SPNs of either kind preferentially target SPNs of a certain subtype? In particular, based on their point estimates of connection probabilities, Taverna et al. (2008) have claimed that D1 neurons prefer to connect to other D1 neurons than to D2 neurons, a claim used in the computational study by Burke et al. (2017). As evidenced in Figures 2A and E which plot the density functions for $\Delta_{(D1\rightarrow D1)-(D1\rightarrow D2)}$ according to the data by Taverna et al. (2008) and Planert et al. (2010) respectively, this is not the case as both density functions include 0 among their most likely values. In fact, the probability that $\Delta_{(D1\rightarrow D1)-(D1\rightarrow D2)}$ is smaller than 0 is equal to 0.19 and 0.30 in Taverna et al. (2008) and Planert et al. (2010) respectively. Similarly, we do not find any difference when looking at whether D2 neurons have a preference for a particular postsynaptic neuron subtype (see Figures 2B and F for the density functions of $\Delta_{(D2\rightarrow D2)-(D2\rightarrow D1)}$): for Taverna et al. (2008), the probability that $\Delta_{(D2\rightarrow D2)-(D2\rightarrow D1)}$ is larger than 0 is 0.16, and the probability that it is smaller than 0 is 0.17 for Planert et al. (2010). We thus find no evidence that SPNs of one subtype (D1 or D2) preferentially target a certain subtype.

Having established that rates of connection are not different for a given presynaptic neuron subtype, we can collapse the data according to the subtype of the presynaptic neuron to answer another question: are D1 neurons more or less likely to make connections overall than D2 neurons? To do this, we simply add up the total number of tested pairs and connections found for the same presynaptic neuron type (e.g.

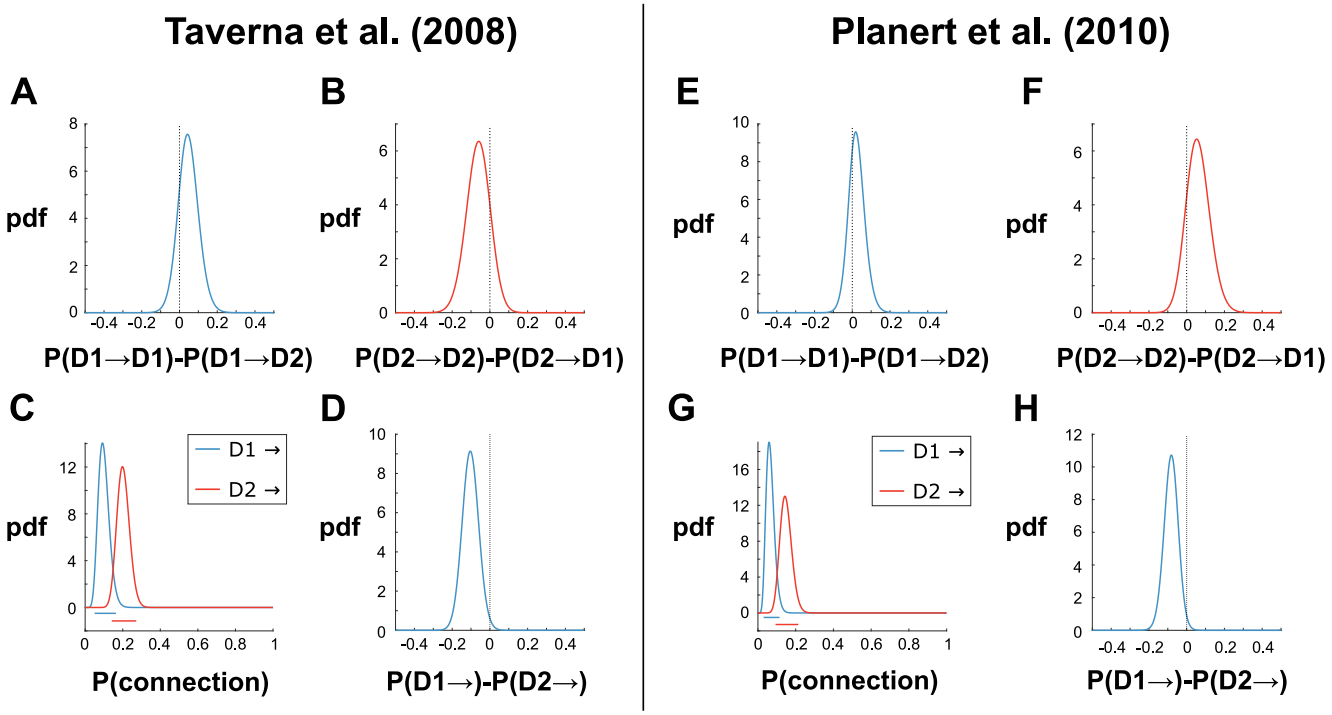

Figure 2. Comparison of the probabilities of connections between different SPN combinations. **A** Density function for the difference in the probabilities of connection in pairs with a presynaptic D1 neuron using data from Taverna et al. (2008). **B** Density function for the difference in the probabilities of connection in pairs with a presynaptic D2 neuron using data from Taverna et al. (2008). **C** Posterior density functions for the probabilities of connection collapsed according to the presynaptic neuron subtype. Bars underneath the curves correspond to the 95% credibility intervals. **D** Density function for the difference in connection probability between pairs with a presynaptic D1 neuron and pairs with a presynaptic D2 neuron. **E-H** Same as A-D, using data from Planert et al. (2010).

for D1 SPNs: $n_{D1\rightarrow} = n_{D1\rightarrow D1} + n_{D1\rightarrow D2}$ and similarly for k). In essence, this is equivalent to considering the posterior of one connection rate as the prior for connections with that same presynaptic neuron subtype, e.g. $f_{D1\rightarrow D1}(p)$ is the prior for $f_{D1\rightarrow SPN}(p)$. Figures 2C and G show the posterior distributions for the collapsed datasets, and both studies agree that D1 SPNs are less likely to make connections to other SPNs than D2 SPNs. The MAP values for connection rates from D1 neurons is 0.092 and 0.059 in Taverna et al. (2008) and Planert et al. (2010) respectively, versus 0.199 and 0.143 for connection rates from D2 neurons. If we look at the density function for the difference between the probability of connections for D1 and D2 neurons (Figure 2 D and H), the MAPs for $f_{\Delta_{(D1\rightarrow SPN)-(D2\rightarrow SPN)}}$ are -0.11 and -0.08 for Taverna et al. (2008) and Planert et al. (2010) respectively, while in both cases the probability that $\Delta_{(D1\rightarrow SPN)-(D2\rightarrow SPN)}$ is less than 0 is equal to 0.99. Thus, both studies contain very convincing evidence that D2 neurons are about twice as likely as D1 neurons to make connections to another SPN.

Probability of connection as a function of distance

So far, when considering data on SPN connections from Taverna et al. (2008) and Planert et al. (2010) we have been careful to analyse each study separately, resisting the temptation of combining the two into a more powerful dataset. We were justified in being so careful since the two experiments used different maximum intersomatic distances between neurons, namely 50 μm in the study of Taverna et al. (2008) and 100 μm in that of Planert et al. (2010). Given that probability of connection between neurons typically decreases with distance (Hellwig, 2000; Humphries et al., 2010), sampling within a larger area around a neuron will probably cause a decrease in the ratio of connected pairs. Indeed, if we compare the probability of D1 or D2 neurons making lateral connections between the two experiments, we find that these probabilities tend to be larger in Taverna et al. (2008) which has the smaller sampling area, as illustrated by the corresponding density functions shown in Figure 3A. To address this concern, instead of focusing on a raw probability of connection p that is agnostic to distance, we can try to extract information about the rate of decrease in connectivity as a function of distance in these two studies and see whether it is consistent between the two studies. To do this, we start by positing that this decrease obeys a simple exponential decay function:

$$P(connection|distance = r) = e^{-\beta r} \tag{6}$$

with β the decay parameter of unknown value, and r (for radius) the distance separating the two neurons. This is obviously a simplistic model, but its advantage for us is its dependence on a single parameter β, which, as we will demonstrate, can be linked to the binomial success rate p that we've been concerned with so far (and we note that any model with more than one parameter could not have its parameters uniquely estimated from the single estimate of p) . Ideally, to estimate this β parameter would require knowledge about the exact distance between every recorded pair of neurons, from which we could directly fit the model, but with simple assumptions on the sampling method used by experimenters, we can find an alternative way of converting values of β into p. Since the distance between each sampled pair of neurons in an experiment is in fact unknown to us, we shall consider it as a random variable. We can now express p as a function of β as

$$p(\beta) = \int_0^R f_{equi}(r) e^{-\beta r} dr, \tag{7}$$

which is the product of the probability of experimenters selecting a neuron at distance r from another $f_{equi}(r)$, and of the probability of these neurons being connected knowing r (equation 6) integrated over all possible values of r (see Figure 3 C for a visual depiction of what equation 7 means). R is the maximum distance at which the experimenters are sampling neurons, equal to 50 or 100 μm in Taverna et al. (2008) and Planert et al. (2010) respectively. We now need to find $f_{equi}(r)$.

A simple model for f_{equi} would be that, given a certain volume surrounding a central neuron, the probability of sampling any given neuron in that volume is equiprobable for all neurons (Figure 3B). With this assumption, we obtain the solution (see Methods):

$$p(\beta) = \frac{2}{R^2} \int_0^R r e^{-\beta r} dr \tag{8}$$

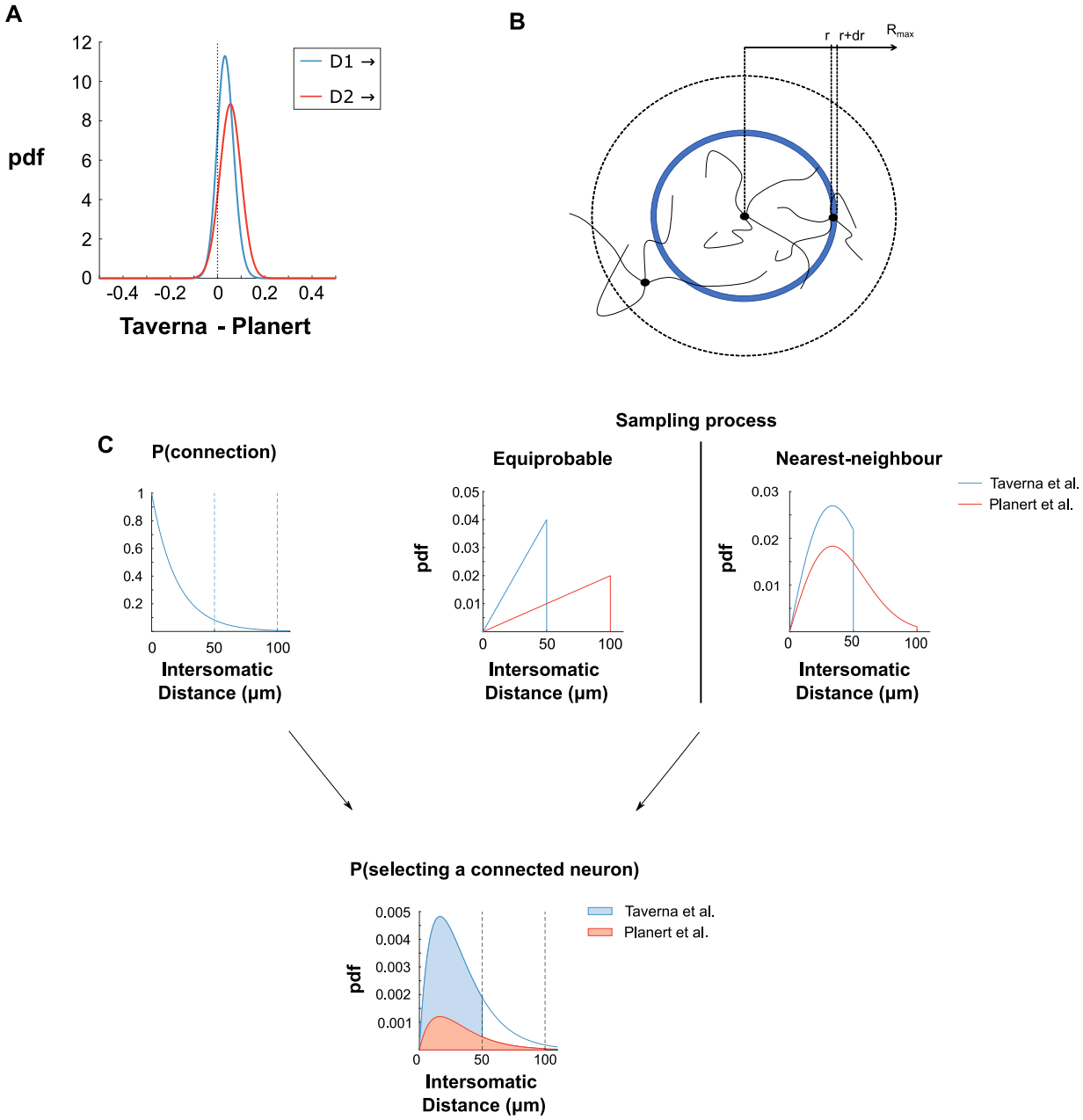

Figure 3. Estimating the probability of connection as a function of distance. **A** Density function for the difference in connection rates for a given presynaptic SPN type between the Taverna et al. and Planert et al. studies. **B** Experimenters chose their neurons within a certain maximum distance R_{max} which defined a thin cylindrical volume of interest (here we draw the top of that cylinder). In the case of equiprobable sampling, the probability of choosing neurons further away increases as the infinitesimal volume corresponding to that distance increases as a linear function of r. **C** The probability of finding a connected pair of neurons depends on two different processes. Firstly, the process of connection, modelled by the probability of connection between two neurons given the distance between them, which we postulate decays exponentially; secondly, the process of sampling neurons in the experiment, modelled as the probability of selecting another neuron at a given distance from a starting neuron. We explore here two different scenarios for the sampling process: an equiprobable scenario in which neurons within a determined volume are selected randomly, and a nearest-neighbour scenario in which the selected neuron is whichever is the closest within the maximum distance set by the experimenters. The overall rate of connection reported by the experimenters then corresponds to the integral (shaded areas) of the product of these two probability models. Hence, differences in sampling processes can cause different rates of connection, even if the probability of connection given distance is the same.

which gives us the corresponding value of p for any desired value of β. As we have posteriors $f_p(p)$ for the probability of connection between two types of neuron, we can now transform these into posteriors for β, $f_\beta(\beta)$ through parameter substitution using equation 8 (see Methods).

Probability of connection decreases faster for D1 than for D2 neurons

We apply this method to the posteriors for the probabilities of connection collapsed according to the subtype of the presynaptic SPN and obtain the posteriors for the decay rate β shown in Figure 4A and B for D1 and D2 neurons respectively. Despite not being perfect, there is a remarkable level of agreement between the two studies, which give estimates in the same ballpark. In both cases, although the posteriors do not overlap much, they do in fact lie quite close to each other providing us with a continuous, albeit broad, range of possible values . In the case of D1 neurons, the decay rate is expected to be in a region between 0.03 and 0.13 μm^{-1}. The exponential decay curves representing the probability of connection as a function of distance for a decay rate equal to the MAP of each study (which are given in Table 3) are also shown in the inset of Figure 4A, and it is evident that they are extremely close to one another. As for D2 neurons (Figure 4B), the decay rate is smaller, as expected given that we have already shown that the overall probability of connection is higher for these neurons, ranging between 0.02 and 0.07 μm^{-1}.

To get a better idea of how consistent the results are between the Taverna et al. (2008) and Planert et al. (2010) datasets, we used Monte Carlo simulations to try and recover the number of observations made by each experiment with its own value of R using the β value that seems most likely given the two posterior curves, i.e. the intersection of the two posterior curves (black dotted lines in Figure 4A,B). We ran 10000 virtual experiments by generating random distances between pairs of neurons (according to equation 20 in the Methods) and the maximum distance R used by that study, and then deciding whether they were in fact connected according to the probability of connection of equation 6. We generated the same number of pairs as tested in each study and then reported the number of times we obtained the exact same number of positive results (red bars in the histograms of Figure 4C and D). For instance, setting an intermediate decay rate of 0.075 μm^{-1} for D1 neurons, and generating 10000 replications of the experiment of Taverna et al. (2008) which recorded 85 pairs of SPNs with a D1 presynaptic neuron, we obtained more than 500 simulations where exactly 8 pairs were connected, which is the number originally reported (Figure 4C left figure). In all four cases, we managed to replicate the original results relatively often, proving that the best estimates for β are reasonable. As a sanity check, we also used the decay rate MAPs for one subtype to try and replicate the results of the other subtype and found it much harder if not impossible to replicate the results, verifying that the decay rates have to be different for the two types of SPNs (results not shown).

Study	Pair	$\hat{\beta}_{MAP}$ (μm^{-1})	95% Credibility Interval
Taverna et al. (2008)	D1 SPN → SPN	0.084	[0.064, 0.125]
	D2 SPN → SPN	0.054	[0.043, 0.070]
Planert et al. (2010)	D1 SPN → SPN	0.053	[0.040, 0.082]
	D2 SPN → SPN	0.033	[0.026, 0.045]
	FSI → D1 SPN	0.002	[0.0004, 0.009]
	FSI → D2 SPN	0.006	[0.002, 0.017]
Gittis et al. (2010)	FSI → D1 SPN	0.004	[0.003, 0.005]
	FSI → D2 SPN	0.007	[0.005, 0.009]
	FSI → FSI	0.003	[0.001, 0.008]
Ibáñez-Sandoval et al. (2011)	NGF → SPN	0.002	[0.001, 0.006]

Table 3. MAPs and 95% credibility intervals (in μm^{-1}) of the posterior curves for β.

If we compare the estimates between the two datasets more critically, there is a clear bias for the posterior curves extracted from Taverna et al. (2008) to be shifted to the right compared to the posterior curves from Planert et al. (2010) (Figure 4A and B). Indeed, if we refer to the insets in Figure 4A and B, the best estimate of β according to the study by Taverna et al. (2008) would predict a 50% drop in probability of connection every 8 μm for D1 neurons versus every 13 μm according to the study by Planert et al. (2010). In the case of D2 neurons, the difference between the two exponential curves is even greater, with a half distance of 13μm versus 21 μm according to Taverna et al. (2008) and Planert et al. (2010) respectively. Furthermore, the density functions for the difference in posteriors between the two studies both lie predominantly in the positive domain (Figure 4E and F) indicating that the decay

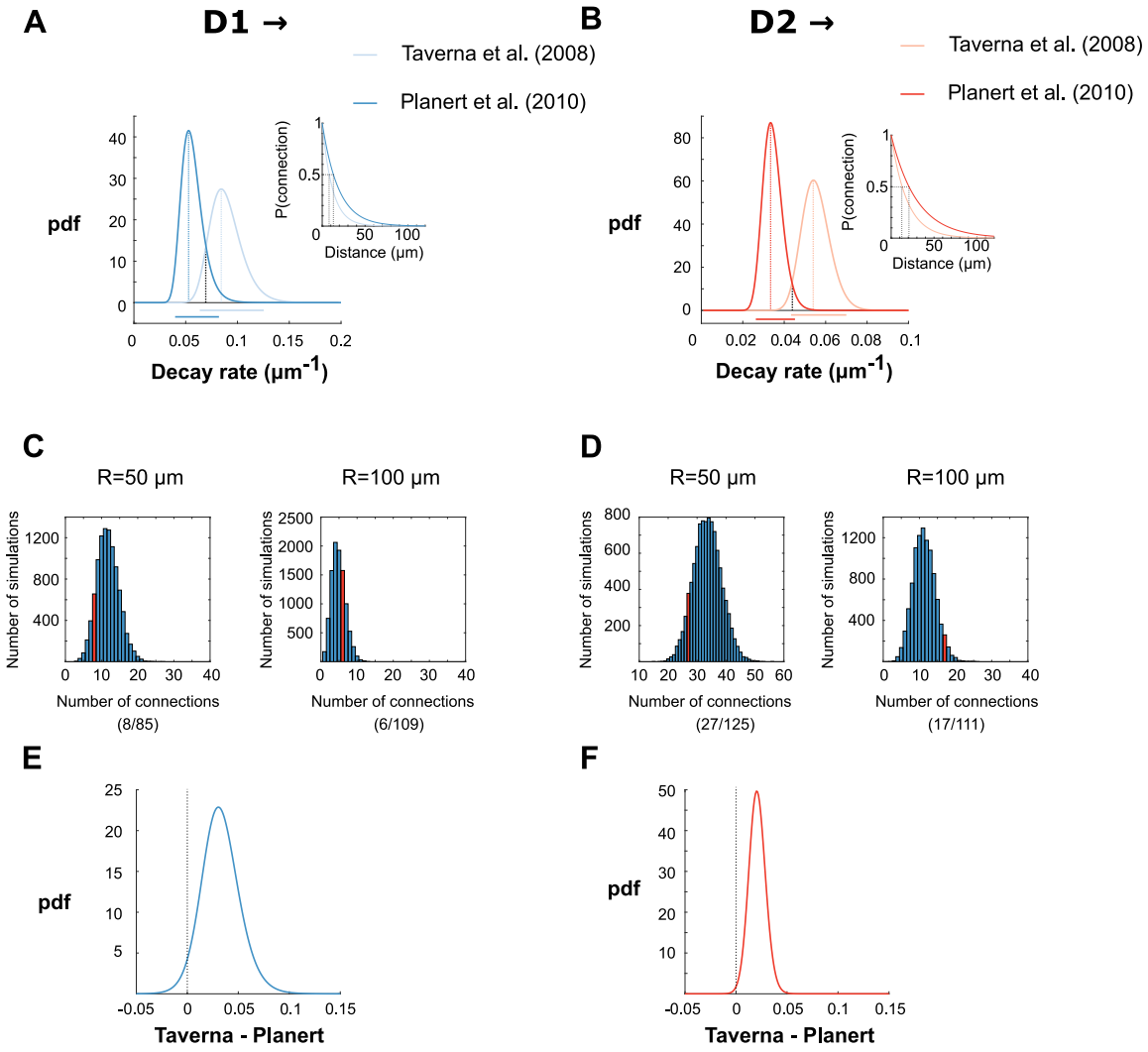

Figure 4. Estimates for the distance-dependence of connection probability between SPNs **A & B** Posterior density functions for the decay parameter of an exponential function representing the probability that a D1 or D2 neuron connects to a neighbouring neuron. Bars underneath represent 95% credibility intervals. Vertical black dashed line indicates the value of β at the maximum intersection of the two posteriors. Inset: Probabilities of connection given distance using the MAP values from the decay rate posteriors. **C & D** Monte Carlo simulations in which the best intersection estimate of β from **A & B** is used to try and replicate the exact experimental results of Taverna et al. (2008) (left) and Planert et al. (2010) (right) concerning pairs with a D1 presynaptic neuron (**C**) or a presynaptic D2 neuron (**D**). The exact results obtained by the experimenters correspond to the red bars and given between brackets underneath the bar graphs. **E & F** Density functions for the difference in decay rates between the two studies.

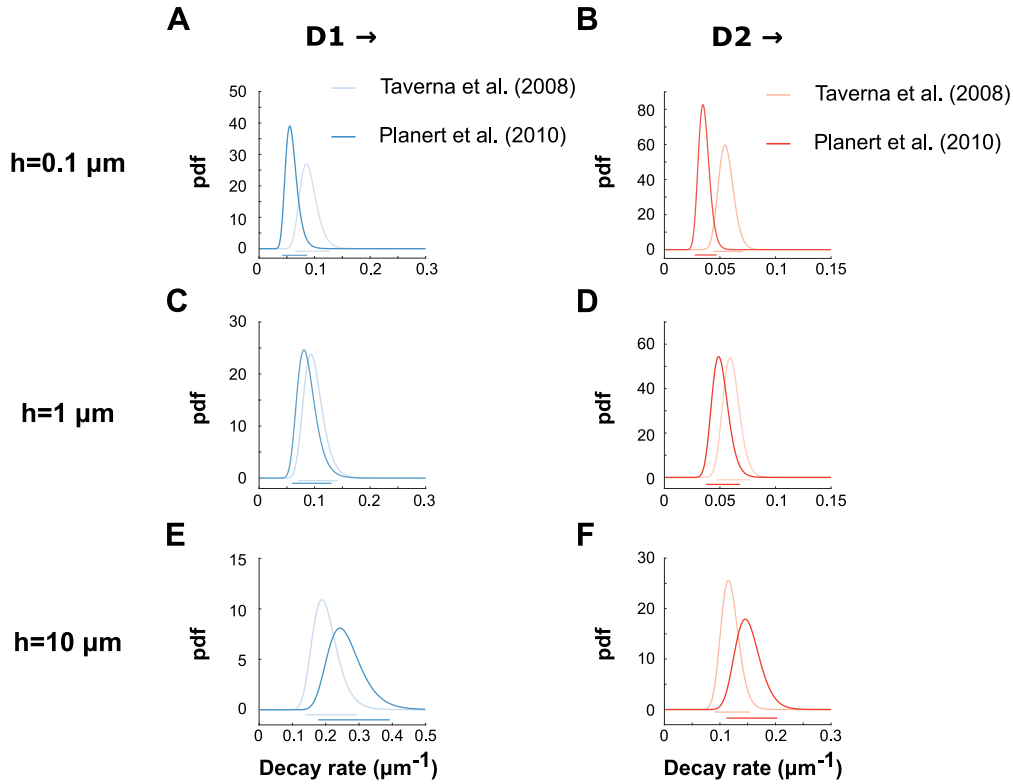

Figure 5. Density functions for the decay parameter assuming a nearest-neighbour model of neuron selection for different values of the depth h of the sampling region. **A** Density function for β for D1 neurons when $h = 0.1\mu m$. **B** Density function for β for D2 neurons when $h = 0.1\mu m$. **C-D** Same as **A-B** for $h = 1\mu m$. **E-F** Same as **A-B** for $h = 10\mu m$.

parameter in the study of Taverna et al. (2008) is probably greater than in the study of Planert et al. (2010). We next consider explanations for this.

Biased neuron sampling could explain differences between datasets

One potential explanation is that the sampling of neuron pairs was more complex than the equiprobable sampling model we first assumed. In this section, we explore this possibility by considering a model where, rather than choose neurons equiprobably within a visible area surrounding a neuron, experimenters preferentially tested neurons which were closer to one another to maximise the probability of detecting connections.

To explore this model, we used the same probability of connection given distance in combination with a new density function for the probability of the distance of the tested neuron to derive new posteriors for the decay parameter. To simplify matters, we will consider a case where experimenters only patched pairs of neurons, ignoring the fact that Planert et al. (2010) would also patch triplets or quadruplets, and we will assume that they always patched the closest neuron within the maximum distance they set themselves. This means that we are looking for the density function for the nearest neighbour, $f_{NN}(r)$, which has a solution of the form:

$$f_{NN}(r) = kre^{-\pi r^2 hN} \tag{9}$$

with k a normalising constant (see Methods for derivation). Normally, k is defined so that integrating f_{NN} between 0 and infinity is equal to 1, i.e. the nearest neighbour must be somewhere in that interval. In our particular case however, experimenters set themselves a maximum distance of either 50 or $100\mu m$, meaning that the closest neuron must be closer than this distance (if there was no neuron closer than this, experimenters would simply look for another pair). In other words, k is such that:

$$k\int_0^R re^{-\pi r^2 hN}dr = 1 \tag{10}$$

which ultimately gives us:

$$k = \frac{2\pi hN}{1-e^{-\pi R^2 hN}} \tag{11}$$

Contrary to the previous equiprobable sampling hypothesis where these parameters cancelled out between numerator and denominator, this density function depends on N, the density of SPNs in the striatum, and h, the height of the cylinder in which sampling takes place. Given that we have collapsed probabilities of connection based on the nature of the presynaptic neuron, we can simply use an overall SPN density in the mouse brain of 80500 per mm^3 following the convention of Hjorth et al. (2020) who chose this number based on the work of Rosen and Williams (2001) (and which is close to the estimated density of 84900 per mm^3 in the rat brain Oorschot (1996)). As for h, the experimenters tell us that neurons were sampled in the same field of focus which would correspond to a height with an order of magnitude of a tenth or even a hundredth of micrometer. However, given that for a neuron to be in the same field of focus as another, it suffices that some part of its soma, whose diameter is between 10 and 20 μm in mice according to Gagnon et al. (2017), intersects this very small volume, we can expect h to be much larger in practice. Because of this uncertainty, we used three different values of h to get three different nearest-neighbour distributions: 0.1 μm, 1 μm and 10 μm. As previously, we can now combine these probabilities of sampling a neuron at a given distance with the probability of connection given distance to find the overall probability of connection:

$$p(\beta) = \int_0^R P(connection|distance = r).f_{NN}(r)dr = k\int_0^R re^{-r(\pi rhN+\beta)}dr \tag{12}$$

which gives us a new mapping between p and β. We then transform the posteriors for p into posteriors for β as before (Methods). The first column of Figure 5 shows these posteriors for D1 neurons. The picture for $h = 0.1\mu m$ is not so different from that obtained under the equiprobable sampling hypothesis, but for greater values of h, the posteriors overlap far more. In particular, for h equal to 1 μm the posterior curves practically coincide. Similarly for D2 neurons, $h = 0.1\mu m$ does not much improve the agreement between the two studies, but greater values of h do (Figure 5, second column). Leaving aside the correct value of h and the fact that experimenters routinely patched more than two neurons at once, this approach successfully illustrates how a tendency to select neurons closer together might account for the discrepancy observed in estimates of the decay parameter using the simpler equiprobable sampling model. Moreover, this analysis shows how the details of data sampling matter when estimating connectivity statistics from intracellular recording data.

Fast Spiking interneurons preferentially connect to D1 SPNs

We now turn to completing our Bayesian map of the striatal microcircuit by evaluating the connections of different species of interneurons to the SPNs and to each other – we present the full map in the Discussion (Figure 10). Although they make up less than 5% of the total neuron population of the striatum (Kreitzer, 2009; Oorschot, 2012; Burke et al., 2017), interneurons are a crucial component. Three main types of interneurons are commonly documented (Kreitzer, 2009): Fast Spiking (FS), Persistent Low Threshold Spiking (PLTS), and cholinergic (Ach) interneurons. We will also include in this list tyrosine-hydroxylase (TH) and NPY-NGF interneurons because of their relationship to Ach interneurons, which is crucial to the function of the cholinergic component of the circuitry (Ibáñez-Sandoval et al., 2011; English et al., 2012; Dorst et al., 2020). We took data on pairwise intracellular recordings of these interneuron types from the range of studies listed in Table 1, and determined Bayesian posteriors for p as previously. Unlike for the SPNs we did not have prior studies of the interneuron connections to help us design a prior, so we relied on a uniform prior instead.

We focus first on GABAergic interneurons that project to the SPNs. FS interneurons in particular are known for making a great number of powerful inhibitory connections onto spiny projection neurons (Tepper et al., 2004, 2008). To examine these connections, we used intracellular recording data from Planert et al. (2010) and Gittis et al. (2010). For the connections from FS interneurons to SPNs, the posteriors we obtain for Planert et al. (2010) (Figure 6B) are consistent with quite high connection

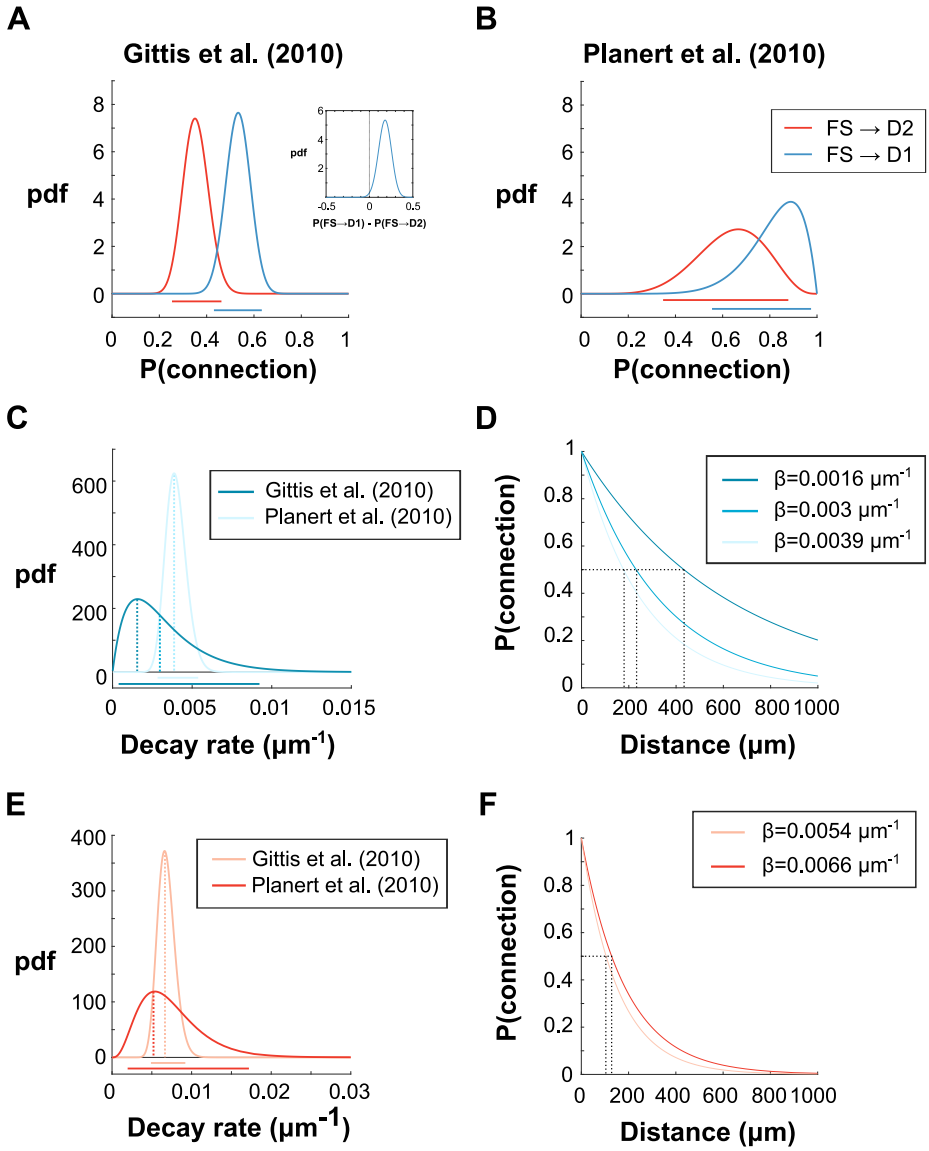

Figure 6. Bayesian analysis of connection probabilities of FS interneurons onto D1 and D2 SPNs. **A** Connection probabilities of fast spiking (FS) interneurons connecting to D1 and D2 SPNs according to the data of Gittis et al. (2010) who set a maximum distance of 250 μm between neurons. Inset: density function for the difference in probability of connection. **B** As for panel A, according to the data of Planert et al. (2010) who used a maximum distance of 100 μm instead. **C** Posterior density functions for the decay rate of probability of connection for FS→ D1 pairs assuming equiprobable sampling of neurons. **D** Probabilities of connection given distance for three different values of β corresponding to the MAP estimates of each study and the intersection of the two posterior curves. **E-F** Same for FS→D2 pairs as C and D respectively. Because the MAP estimate for Planert et al. (2010) coincides with the intersection of the two posteriors, only two exponential decays are tested in F.

probabilities: the MAPs are 0.67 and 0.89 for connections to D2 and D1 neurons respectively. However, the small size of the samples means that the range of possible values is also quite broad: the 95% credibility interval extends from about 0.35 to 0.88 in the case of connections to D2 neurons, and 0.56 to 0.98 for D1 neurons. Fortunately, the study of Gittis et al. (2010) is based on a much larger sample resulting in narrower posteriors shown in Figure 6A. The MAP for D2 neurons is 0.35 with a 95% credibility interval ranging from 0.25 to 0.46, versus 0.53 with a 95% credibility interval going from 0.43 to 0.63 for D1 neurons. Thanks to these narrower posterior curves, it is possible to conclude that FS interneurons preferentially target D1 neurons. In fact, when we integrate $\Delta_{FS\rightarrow D1-FS\rightarrow D2}$ (Figure 6A inset), we find the probability that FS interneurons prefer to connect to D1 neurons is greater than 0.99.

The two studies seem to disagree as Planert et al. (2010) gives much higher estimates of p, but this is resolved by taking into account the maximum distance used by the two studies, 100 μm for Planert et al. (2010) and 250 μm for Gittis et al. (2010). Indeed, if we convert the posteriors for p into posteriors for the exponential decay rate β of the probability of connection given distance, with the assumption of equiprobable sampling as previously explained, we obtain posteriors which very largely overlap (Figure 6C and E; Table 3) thus reconciling the two studies. In line with the already discussed overall smaller rate of connection to D2 neurons, the probability of connection drops much faster as distance increases for connections to D2 neurons (dropping to 50% after about 100 μm, see Figure 6F) than for connections to D1 neurons (50% connection rates occurring at a distance of at least 200 μm, see Figure 6D). Consequently, there are at least two length-scales in the striatal microcircuit, with connections between SPNs falling to 50% probability within a few tens of micrometers (Figure 4A,B), but connections to SPNs from FS interneurons falling to 50% probability at a hundred micrometers or more (Figure 6D,F).

PLTS interneurons make few local connections in striatum

We turn now to the connections that FS interneurons make on other interneurons of the striatum. To assess these, we analysed data from Gittis et al. (2010) on connections FS interneurons make to PLTS, cholinergic, and other FS interneurons (Figure 7A). Their data on connections between FS and PLTS interneurons were pooled with the data on the same connections from Szydlowski et al. (2013): We checked that the data from the two studies were in agreement by calculating posteriors separately for each study and found the density functions for the difference between the posteriors (Δ) for both directions (FS$\rightarrow$PLTS and PLTS$\rightarrow$FS) included 0 among their most likely values (results not shown).

We see from these data that the probability of connection from FS interneurons to PLTS interneurons is low ($\hat{p}_{MAP} = 0.10$) but uncertainty regarding these connections is quite large (95% credibility interval = [0.03, 0.29]), while the probability of connection to cholinergic interneurons is even more uncertain with a credibility interval ranging from 0 to 0.60 and therefore requires more investigation. Connections between FS interneurons are relatively common ($\hat{p}_{MAP} = 0.58$) but uncertainty is also very high for these pairs (95% credibility interval = [0.32, 0.81]). While this broad uncertainty in p translates into a broad uncertainty for the decay rate β of the probability of connection given the distance between a pair of FS interneurons (Figure 7B), we see that the distance-dependence for pairs of FS interneurons is similar to that for connections of FS interneurons to SPNs with a half-distance for connection probability as a function of distance of about 200 μm for $\hat{\beta}_{MAP} = 0.003\mu m^{-1}$. Unfortunately, Szydlowski et al. (2013) do not provide a maximum distance which prevents us from transforming $f_p(p)$ for FS $\rightarrow$ PLTS pairs into the corresponding $f_\beta(\beta)$.

The combined data of Gittis et al. (2010) and Szydlowski et al. (2013) (Figure 7C) shows no evidence that PLTS interneurons connect locally to other interneurons ($\hat{p}_{MAP} = 0$ for all pairs), but connections to MSNs, although sparse, are clearly established within a maximum distance of 250 μm ($\hat{p}_{MAP} = 0.033$, 95% credibility interval = [0.010, 0.114]). Compared to FS interneurons, there is less uncertainty concerning the rates of connection for these PLTS interneurons, as evidenced by the smaller credibility intervals shown in Figure 7C. It remains to be seen whether there is an asymmetry in connection probability to D1 and D2 SPNs as the experimenters did not make this distinction when testing PLTS $\rightarrow$ SPN connections. Furthermore, it has been reported that PLTS interneurons prefer to make connections to faraway SPNs (Straub et al., 2016) meaning that this low rate of connection locally is not the full extent of PLTS interneuron connectivity, and also prevents us from using an exponential decay model of probability to extract a decay rate as we did for the SPNs and FS interneurons.

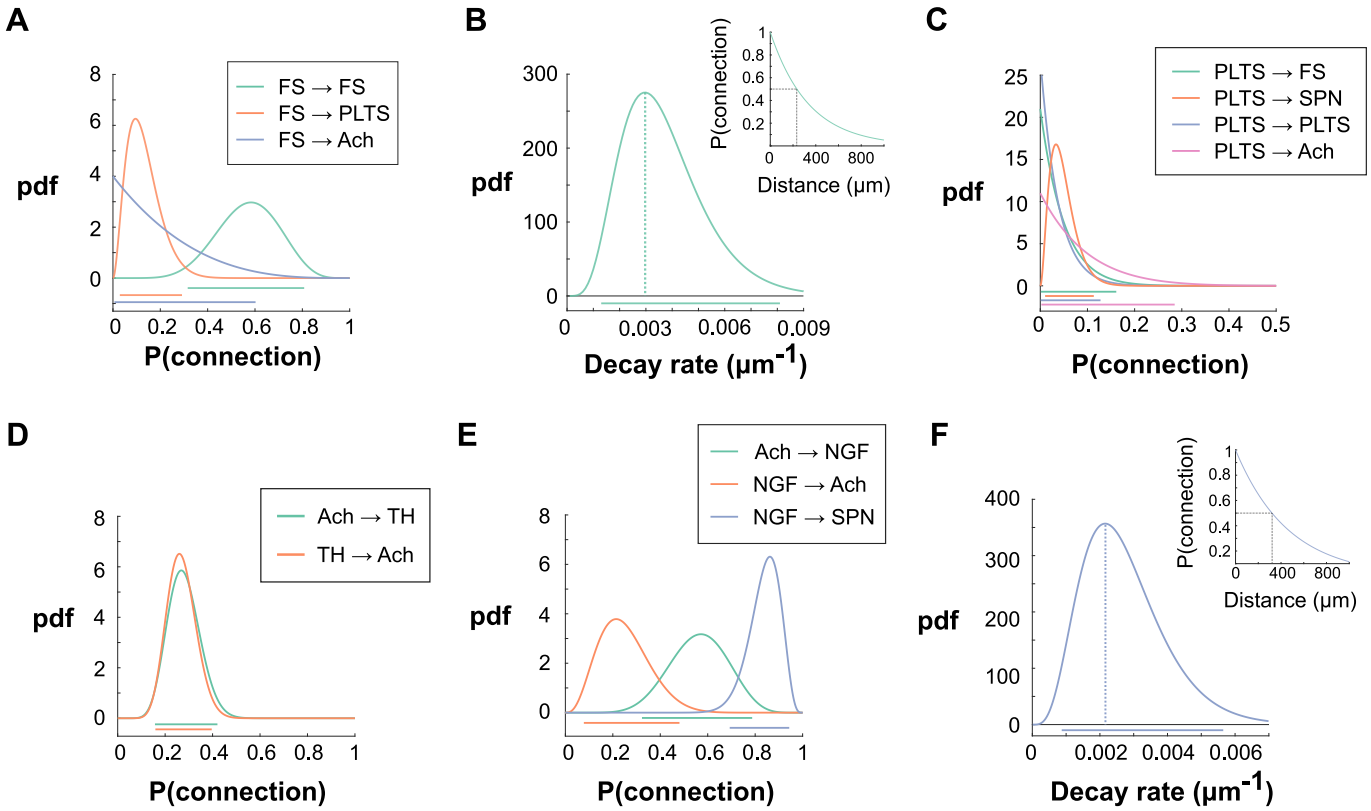

Figure 7. Bayesian analysis of connection probabilities between striatal interneurons using a uniform prior. **A** Posterior density functions for FS interneurons connections onto other interneurons according to the data of Gittis et al. (2010). **B** Posterior density functions for the decay rate of probability of connection for FS → FS pairs assuming equiprobable sampling of neurons. Inset: Exponential decay function for the probability of connection between pairs of FS interneurons corresponding to the MAP estimate of the decay rate. **C** Posterior density functions for PLTS interneuron connections onto other interneurons according to the data of Gittis et al. (2010). **D** Posterior density functions for connections between cholinergic and TH interneurons according to data from Dorst et al. (2020). **E** Posterior density functions for connections between cholinergic interneurons, NGF interneurons and SPNs according to the data of English et al. (2012) and Ibáñez-Sandoval et al. (2011). **F** Posterior density functions for the decay rate of the probability of connection for NGF → SPN pairs assuming equiprobable sampling of neurons. Inset: Exponential decay function for the probability of connection between NGF → SPN pairs corresponding to the MAP estimate of the decay rate in E.

Further evidence that the effect of cholinergic interneurons onto SPNs is mediated by GABA interneurons

Finally, a very recent study (Dorst et al., 2020) pursued further the investigation of the connectivity of cholinergic interneurons to other interneurons. They relied mostly on optogenetic approaches to rule out connections between cholinergic interneurons and various interneurons, which, given that it is based on optogenetic stimulation of an entire subpopulation of neurons, is not suited to our Bayesian method which depends on a precise count of tested pairs. However, having obtained a response when testing connections between cholinergic interneurons and a subtype of GABAergic interneurons that express tyrosyne-hydroxylase (TH interneurons), they then went on to quantify the probabilities of these connections using similar techniques as the ones we have dealt with in this paper (see Table 1). When we apply our Bayesian method to this dataset (Figure 7D), we find that TH and cholinergic interneurons connect reciprocally to one another quite frequently and with practically equal probabilities ($\hat{p}_{Ach \to TH} = 0.268$, $\hat{p}_{TH \to Ach} = 0.260$). Our Bayesian method supplements the original finding by Dorst et al. (2020) with uncertainty estimates which are relatively good compared to other interneuron connections (95% credibility interval = [0.159, 0.396] for TH $\to$ Ach connections, and [0.157, 0.420] for Ach $\to$ TH connections).

The activity of cholinergic interneurons indirectly affects SPNs via at least one type of GABAergic interneuron (English et al., 2012). To examine this route, we combine pairwise intracellular recording data from English et al. (2012) on connections between cholinergic interneurons and NPY-NGF interneurons, with data from Ibáñez-Sandoval et al. (2011) on connections from NPY-NGF interneurons to SPNs. Our analysis (Figure 7E), reveals that cholinergic neurons connect frequently to NPY-NGF interneurons, which in turn connect very frequently to SPNs, making them an effective relay of cholinergic signals; this relay may also be regulated by the NPY-NGF interneurons frequent feedback connections on cholinergic interneurons. Furthermore, given that Ibáñez-Sandoval et al. (2011) used a maximum distance between neurons of 100 μm, it is possible to transform the posteriors for the probability of NGF $\to$ SPN into posteriors for the exponential decay rate β of probability of connection given distance. The MAP of β is the lowest we have found at about 0.002 μm^{-1} (Figure 7F; Table 3). This means that even at a distance of 100 μm from an NGF interneuron, an SPN still has a 0.8 probability of receiving a connection from this interneuron, which partly explains the effectiveness of the cholinergic system in regulating SPN activity.

Evidence used to compare SPN sub-type connection rates in wild type and Huntington's disease mice is insufficient

To this point we have used our Bayesian approach to evaluate the probability of connection, the evidence for it, and (where possible) its dependence on the distance between neurons for every unique connection within the striatal microcircuit (for which there is extant data). We turn now to showing how our Bayesian approach lets us not just construct a map of the microcircuit, but also quantitatively test evidence for changes in the microcircuit. To do so, in this section we evaluate evidence that connections between SPNs change in a mouse model of Huntington's disease (Cepeda et al., 2013); in the following section we evaluate evidence for how connections between SPNs change over development.

The study of Cepeda et al. (2013) used smaller samples of identified SPN pairs than the (Taverna et al., 2008) and (Planert et al., 2010) studies of SPN connectivity (see Table 4) and consequently the resulting posteriors are notably impacted by the choice of the prior (Figure 8A and D). Indeed, the posterior curves obtained from the data of Cepeda et al. (2013) with a uniform prior or the prior based on previous literature look very different, contrary to those of Taverna et al. (2008) (compare Figures 1C and G) and Planert et al. (2010) (compare Figures 1G and H). In particular, the posteriors for wild-type mice obtained from the prior based on the previous literature look very similar to the initial prior (Figure 8D) simply because of the small number of samples. Independently of the choice of prior, the posteriors overlap so much that it is not possible to confirm the rates of connection differ between any pairs of SPNs in the wild type mice (Figures 8A and D). Given how broad the posteriors are, we infer that this lack of difference is due to insufficient data rather than a true absence of difference.

Crucially, this lack of data is also true when comparing connection rates between the wild type and Huntington's model mice (Figure 8A-B or D-E). We find no evidence to support one of the conclusions reached by the authors that D1 $\to$ D1 connections are more likely in the Huntington's model than wild-type mice. Indeed, if we plot the density functions for the difference in probabilities of connection for D1 $\to$ D1 pairs between the two animal groups, we can see that it is very probable for this difference to be 0, but also any other value ranging from roughly -0.3 to 0.3 if we use a uniform prior (Figure 8C) or a

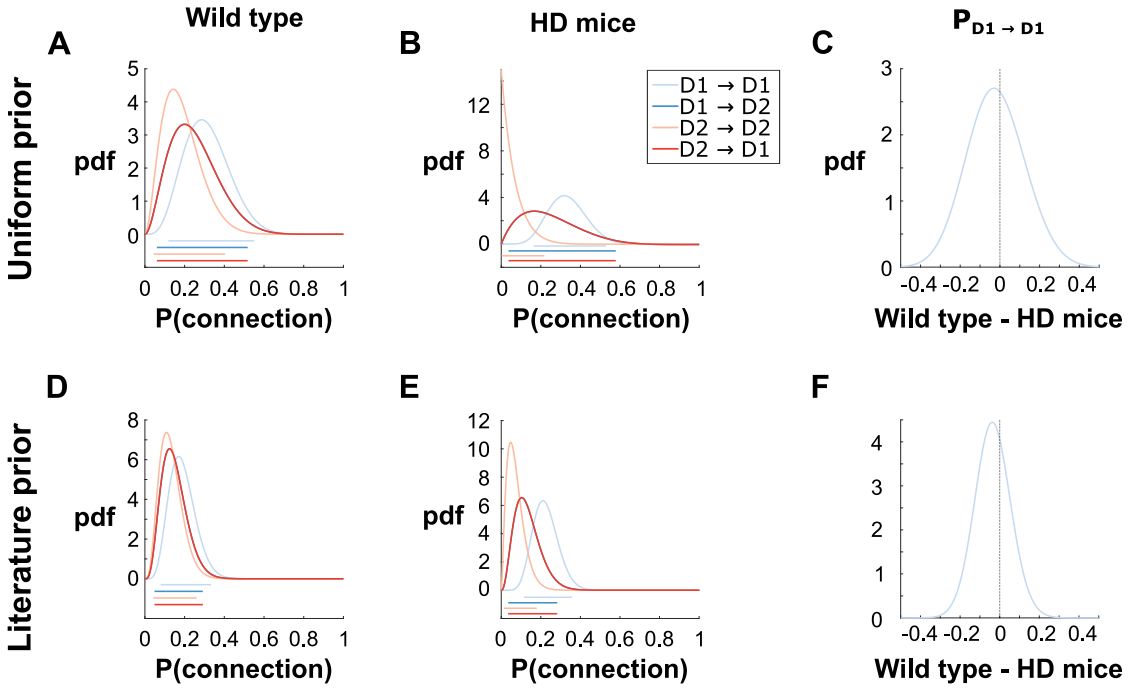

Figure 8. Bayesian analysis of the study by Cepeda et al. (2013), comparing the probability of lateral SPN connections in wild type mice and a model of Huntington's disease. **A** Posterior density functions for the probabilities of connection in the wild type mice using a uniform prior. Bars underneath represent the 95% credibility intervals. The curves for D2 → D1 and D1 → D2 coincide exactly. **B** Posterior density functions for Huntington's disease animals using a uniform prior. The curves for D2 → D1 and D1 → D2 also coincide exactly. **C** Probability density function for the difference in probabilities of connection for D1 → D1 pairs between the two animal groups. **D-F** Same as **A-C** using the prior based on the past literature.

Data	Pair	k	n	Prior	$\hat{p}_{MAP}$	95% credibility interval
Wild type	$D1SPN \to D1SPN$	4	14	Uniform	0.286	[0.118, 0.551]
				Literature	0.170	[0.079, 0.333]
	$D1SPN \to D2SPN$	2	10	Uniform	0.200	[0.060, 0.518]
				Literature	0.124	[0.048, 0.292]
	$D2SPN \to D1SPN$	2	10	Uniform	0.200	[0.060, 0.518]
				Literature	0.124	[0.048, 0.292]
	$D2SPN \to D2SPN$	4	14	Uniform	0.143	[0.043, 0.405]
				Literature	0.109	[0.042, 0.260]
HD model	$D1SPN \to D1SPN$	7	22	Uniform	0.318	[0.164, 0.529]
				Literature	0.210	[0.14, 0.359]
	$D1SPN \to D2SPN$	1	6	Uniform	0.167	[0.037, 0.579]
				Literature	0.104	[0.035, 0.283]
	$D2SPN \to D1SPN$	1	6	Uniform	0.167	[0.037, 0.579]
				Literature	0.104	[0.035, 0.283]
	$D2SPN \to D2SPN$	0	14	Uniform	0	[0, 0.218]
				Literature	0.048	[0.013, 0.180]

Table 4. Experimental data from Cepeda et al. (2013) alongside results of the Bayesian analysis using either a uniform or literature prior. For each type of connection, k is the number of connections that were found and n the total number of tested connections. HD: Huntington's disease.

slightly more conservative -0.2 to 0.15 using the prior based on previous literature (Figure 8F).

D2 SPN connection asymmetries appear during development

The development of connections between SPNs and their asymmetry can be tracked through postnatal development thanks to a recent study by Krajeski et al. (2019) who measured the probability of connection for different SPN pairs at three different stages of post-natal mouse development. The researchers reported that D1 neurons established lateral connections earlier than D2 neurons (a reproduction of these results with our added Wilson confidence intervals is shown in Figure 9A-C). We used our Bayesian approach to check this conclusion against the uncertainty in the experimental data, and provide further details of the development of the striatal microcircuit.

In order to apply our Bayesian method to this data we resort to the uniform prior, since we have no particular expectation about these connection rates at these stages of development, and obtain the posteriors for each combination of neurons shown in Figure 9D-F. We have also tested the Jeffreys prior and obtained practically identical results (not shown). We can see that D1 neurons have already made some connections in the first 3-6 days of postnatal development (P3-6; Figure 9D), but it is hard to say from inspecting the posteriors at each subsequent developmental stage (P9-12 and P21-28) whether the connections made by D1 SPNs continue to develop or have already finished by P3-6 (Figure 9E and F). If we instead look at the difference in posteriors (f_Δ) between consecutive developmental stages (Figure 9 G and H), we see some evidence that connections made by D1 neurons continue to develop, with respective probabilities of 0.81 (P3-6 to P9-12) and 0.77 (P9-12 to P21-28) that the connection density of D1 neurons increases (probabilities again found by integrating f_Δ between 0 and 1). Comparing the earliest (P3-6) and latest (P21-28) stages gave a similar probability of 0.85 that D1 connections increased (not shown).

For presynaptic D2 neurons on the other hand, it is clear that no or very few connections are present at P3-6, and that they gradually appear later (Figure 9D-F). Computing the difference in posteriors (f_Δ) for $P(D2 \to D2)$ and $P(D2 \to D1)$ between consecutive stages (Figure 9 G and H), we find the probability that D2 SPNs increase their connection density is 0.95 between P3-6 and P9-12 and 0.92 between P9-12 and P21-28, indicating a gradual development of these connections up to postnatal days 21-28. By also calculating $f_{\Delta_{((D2\to D1)-(D2\to D2))}}$ at each of these three stages (plotted as insets in Figure 9 D-F), we find good evidence that D2 neurons first connect to other D2 neurons before connecting to D1 neurons, supporting the claim by the authors of the original article (Krajeski et al., 2019). Notably, our analyses in this paper have thus shown that while there is no evidence for a difference in the probability of pre-synaptic D2 SPNs connecting to D1 or D2 SPNs in the adult striatum (Figure 2B,F), there is evidence that the D2 $\to$ D1 and D2 $\to$ D2 connections develop at different rates.

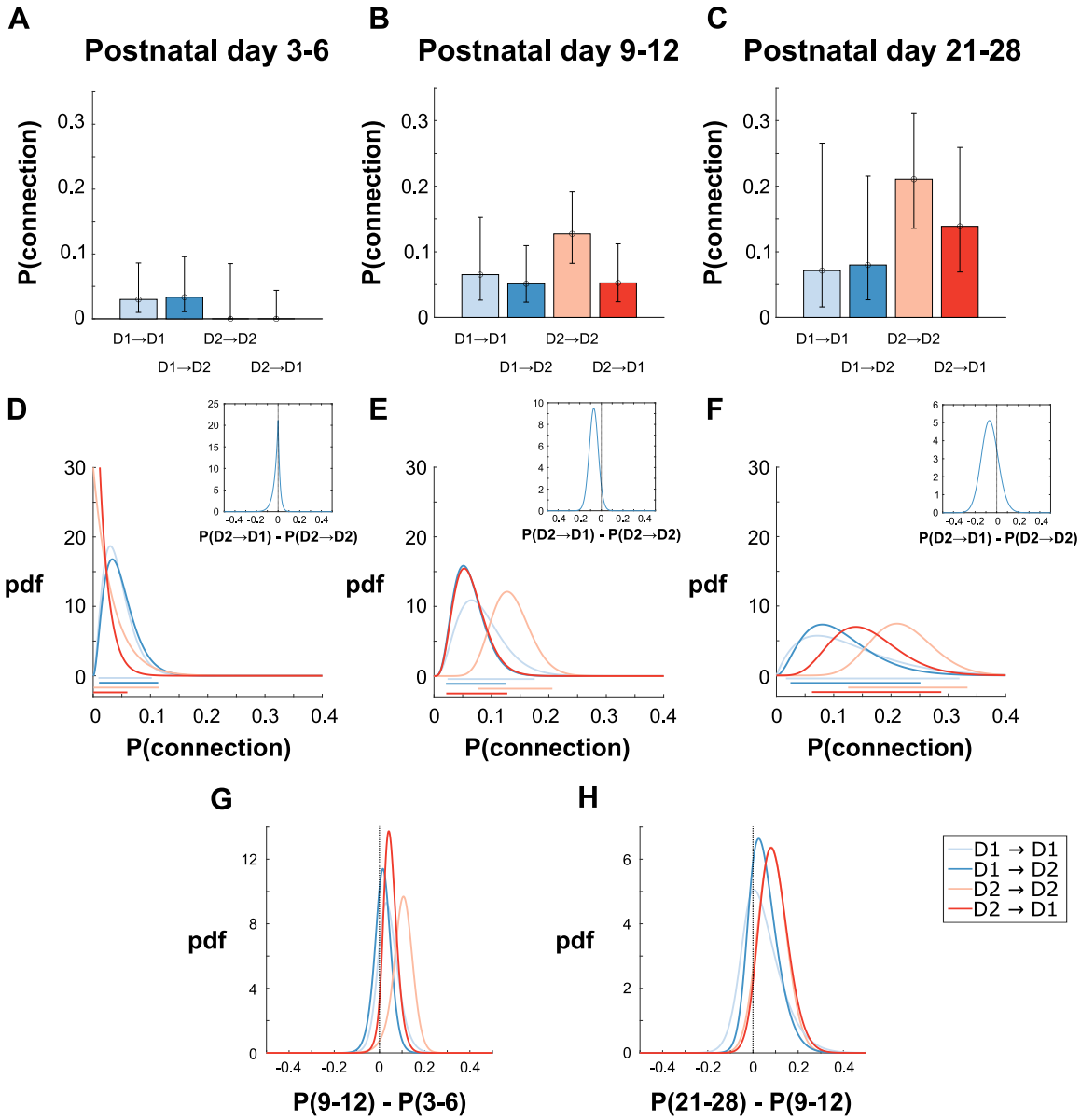

Figure 9. Post-natal development of the lateral connections of SPNs using data from Krajeski et al. (2019). **A-C** Point estimates of the probabilities of connection at different developmental stages from Krajeski et al. (2019). We add here the 95% Wilson confidence intervals. **D-F** Posterior probability density functions for the probability of connections between SPNs at each developmental stage. Coloured bars underneath the plot represent the 95% credibility intervals. A uniform prior as in Figure 1C is used. Inset: Density function for the difference in probability of connection for pairs with a D2 presynaptic neuron. **G-H** Density functions for the difference in connection probabilities for each pair of neuron types between consecutive stages of postnatal development.

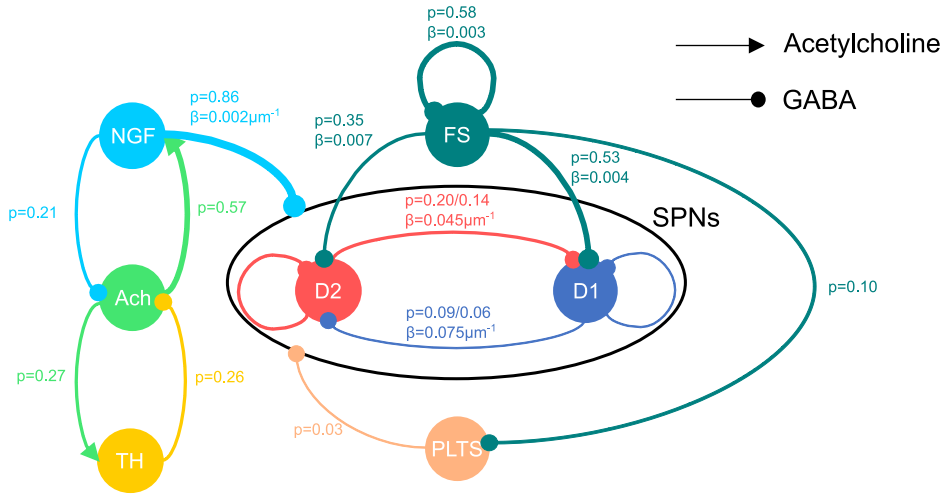

Figure 10. Map of the striatum microcircuitry based on the MAP estimates for p and, when a maximum intersomatic distance was available, the decay rate β assuming equiprobable sampling. Line thickness is indicative of the relative probability of these connections. Connections between and within SPN subtypes are assumed to be the same for a given presynaptic subtype, as established in the main text, and the two different estimates for p correspond to the two different maximum distances used in Taverna et al. (2008) and Planert et al. (2010). Modellers wishing to use this map should beware of the relative population size of these different neurons. For instance, although the probability of connection between SPNs is relatively small compared to connections from FS interneurons, this is potentially counterbalanced by the much greater number of SPNs within a given volume (Humphries et al., 2010).

DISCUSSION

In this paper, we presented novel methods based on Bayesian inference for analysing connectivity using intracellular recording data, and applied them to reconstruct the microcircuit of the striatum. We drew on a range of data from pairwise intracellular recordings, some of which have been previously used to draw conclusions about the existence and probability of connections between striatal neurons, and in turn used to construct models and theories of the striatum and basal ganglia. However, we emphasise that none of these data have had any assessment of the uncertainty in their connection estimates or of the strength of evidence they provide. Our new approach allows us to now draw rigorous conclusions about the strength of evidence for claims about the microcircuit, and in turn synthesise these data into as complete a map as the data allow. Thus, it is reassuring for researchers of the striatum that a number of conclusions we have reached here agree with previous inferences.

A Bayesian map of the striatal microcircuit in rodents

Figure 10 synthesises the complete map we obtained of the striatal microcircuit in rodents. It emphasises our key results: first, there is strong evidence of a connection asymmetry that depends on the presynaptic type of SPN – namely that D2 SPNs are roughly twice as likely to contact another SPN as D1 SPNs – but no evidence of an asymmetry that depends on the postsynaptic type of SPNs; second, that there is strong evidence for FS interneurons preferentially connecting to D1 SPNs; third that there is strong evidence of dense projections from NPY-NGF interneurons to SPNs, likely as dense or denser than those from FS interneurons; and, finally, that connections between SPNs occur on much shorter length-scales then the connections made by interneurons.

For ease of interpretation, Figure 10 summarises each connection probability as the best point estimate we can obtain (namely, the MAP of the corresponding posterior distribution). In this form, the map is not so different in spirit from what a frequentist approach would achieve, such as the one used in Figure 1C of Hjorth et al. (2020). But of course we now have the full posterior distributions underlying these estimates, and some of these are broad – for example, the connection probability from NGF to Ach interneurons has a 95% credible interval twice as wide as its best (MAP) estimate (Table 1). Our posterior distributions

thus reveal that a wide range of potential striatal wirings are consistent with current data.

From this is follows that any model of the striatum should sample its connection probabilities from these posteriors to understand the robustness of the results. After all, it is now well established that parameters of neural models fall into two classes: those whose variation over a wide range does not meaningfully alter the dynamics of the model, and those whose variation the model's dynamics are exquisitely sensitive to – and thus whose precise values are critical to the resulting predictions of the model (Panas et al., 2015; Ponce-Alvarez et al., 2020). And it is likely that striatal dynamics are indeed sensitive to variations in the probabilities and distances of connections (Humphries et al., 2010; Spreizer et al., 2017). Thus, we propose computational researchers change their usual practice of setting a single value for connection parameters, and instead sample from the posterior distribution on each run of their model – to this end, we give the complete form of all our posteriors for p in Table 1 and our confidence intervals for the decay parameter β in Table 3.

Improving the microcircuit map

The construction of our map also revealed or reemphasised further research questions that need addressing to complete it. First, because we are completely reliant on the experimental data at our disposal, the map is silent on such anatomical issues as whether connection probabilities differ between the patch and matrix compartments of the striatum or between different regions of the striatum, especially along the dorsal-ventral axis. For example, for the data on SPN collaterals Taverna et al. (2008) obtained their recordings in slices from both the ventral and dorsal striatum without reporting the exact origin of each test, while Planert et al. (2010) are unclear about the origin of their recordings (the caption of their Figure 5 refers to a network of neurons from the ventral striatum). Thus it remains possible that, rather than being a single microcircuit, there are variations across the striatum.

Second, we lack data on the connectivity of some types of striatal interneuron thought distinct to those examined here, including the calretinin-expressing interneurons and rare subtypes of 5HT3a-expressing interneurons (Tepper et al., 2018). An advantage of our approach is that any new data on pairwise recording data in the striatum can build directly on our analyses, by either updating the posteriors we arrived at, or by estimating new posteriors for connections that lack data at present.

Third, this is a map of connectivity: full knowledge of the influence of one neuron type on another requires data on the strength of the different connections, which may in turn depend on where on the target neuron they fall (Oorschot et al., 2013; Du et al., 2017). Taverna et al. (2008) noted for instance that the connections made by D1 SPNs tended to be weaker than those of D2 SPNs. In future work, using an appropriate Gaussian or log-normal prior, one could extract analogous posteriors for the distributions of conductances of various connection types.

Fourth, we emphasise that this is a map of the local microcircuit, for the connections from a source neuron of one type to other types of neurons in its neighbourhood. These neighbourhoods could be quite large on the scale of the rodent striatum. Where we could estimate the distance-dependence of the probability of connection, we found connections between SPNs fell to negligible within a distance between neuron bodies of 400 μm, but the connections from FS interneurons to D1 SPNs could be detectable out to 1 mm or more. However, some subtypes of interneurons may send axons long distances (Tepper et al., 2018), such as reports of PLTS interneurons with an axon that spans a distance of over 1mm, and which make infrequent bouquets of terminals (Kawaguchi, 1993). More detailed knowledge of these long-distance connections would allow for a more complete map of striatal connectivity.

Finally, we introduced a new analytical method to determine the distance-dependence of the probability of connection between two neuron types, which depended on simple models for the distribution of neurons and how the recorded neurons were sampled in an experiment. Either of these models could be improved with further data. Our estimates of the distance-dependence of connection probability between SPNs varied between two different models for how the experimenters chose which neurons to record. Ideally, knowing the exact distance between each pair of recorded neurons would allow us to dispense with the model assumptions and directly estimate the distance-dependent probability, though measuring each distance would be challenging experimentally.

When comparing connection rates of our two main studies of SPN connectivity (Taverna et al., 2008; Planert et al., 2010), we hypothesised that one reason for their disagreement is the maximum intersomatic distance they used, and its effect on sampling of recorded neurons. Another potential explanation lies in the orientation of the brain slices in which the neurons were tested, as Taverna et al. used parasagittal

slices and Planert et al. both parasagittal and coronal slices, raising the intriguing possibility that SPNs branch out and connect with other SPNs along preferential directions. Penny et al. (1988) noted for instance that while most dendritic trees of SPNs are spherical, some neurons presented more unusual shapes, even planar. Figure 2 of Krajeski et al. (2019) mentions the tendency in SPNs to extend their dendrites along a lateral-ventral to medial-dorsal axis in coronal slices of young mice striatum. More data on any direction preference for the dendritic or axonal projections in the striatum would allow further refinement of the distance-dependent probabilities established here.

Implications for striatal development

Our finding of strong evidence that D1 neurons are less likely to receive connections from SPNs than D2 neurons, both in adults (Figure 2) and at later stages of development (Figure 9), implies a role for active wiring processes in the developing striatum. Consider a simple model of a passive wiring process in which the contact of an axon from a first neuron onto the dendrite of a second neuron is determined only by the probability that an axon segment and a dendritic segment simultaneously occupy the same location (Liley and Wright, 1994; Kalisman et al., 2003; Humphries et al., 2010). For SPNs we have the repeated observation that D1 SPNs have denser dendritic trees for the same volume as D2 SPNs (Gertler et al., 2008; Fujiyama et al., 2011; Gagnon et al., 2017). The passive wiring model would thus predict that D1 SPN dendrites would receive more axonal contacts then D2 SPN dendrites from the same pre-synaptic type of SPN.

If this were true, we would expect to find in our analyses here that the asymmetry of connection rates would depend on the type of the postsynaptic neuron, when we instead find it depends on the type of presynaptic neuron; and we would expect D1 $\rightarrow$ D1 connections to be quite numerous when in fact these are quite rare. Together with our confirmation that the data of Krajeski et al. (2019) show D1 and D2 neurons develop their connections at a different rate, our analyses thus suggest that there is an active wiring process in striatal development that causes either an underexpression of connections to D1 SPNs, or overexpression of connections to D2 SPNs.

Implications for theories of the striatum

Ever since the paper by Jaeger et al. (1994) which reported finding no lateral connections among SPNs, computational modelling of the striatum and wider basal ganglia system have moved away from earlier lateral inhibition circuit models implementing a winner-takes-all strategy where these connections took centre stage (Groves, 1983) towards a predominantly feedforward view (see Plenz (2003) and Tepper et al. (2004) for a more detailed historical account). According to these feedforward models (for example Mink, 1996; Gurney et al., 2001b; Frank, Michael, 2005; Leblois et al., 2006; Humphries et al., 2006), the output of the striatum is entirely determined by the pattern of its cortical inputs modulated by the strength of the different cortico-striatal synapses.

Our Bayesian map of the striatal microcircuit provides further evidence that this feedforward model is limited because the striatum's internal circuit is crucial in shaping its outputs. In the classic direct-indirect pathway model of the basal ganglia (Alexander and Crutcher, 1990), the D1 and D2 SPNs respectively form the populations originating the direct and indirect pathways to the basal ganglia output nuclei, D1 SPNs sending direct axonal projections, and the D2 SPNs projecting first to the globus pallidus, which relays to the output nuclei. We have shown that data on connections between D1 and D2 SPNs (Taverna et al., 2008; Planert et al., 2010) provides strong evidence for D2 SPNs making more connections to other SPNs than D1 SPNs. Some models and theories have interpreted these data as showing that the indirect pathway will dominate the direct pathway output (Bahuguna et al., 2015; Burke et al., 2017). However, that the same data provide no evidence of an asymmetry in the preferred targets of D1 or D2 SPNs further suggests there is not a selective inhibition of the D1 SPNs by D2 SPNs, but that the D2 SPNs are equally likely to inhibit themselves as D1 SPNs. As such, the nature of the interaction of the direct and indirect pathways, and hence their response to cortical and thalamic inputs, remains unclear.

Beyond confirming there is strong evidence for the well-known strong feedforward inhibition of SPNs by FS interneurons (Koos and Tepper, 1999; Tepper et al., 2004), now included in numerous recent theories (e.g. Damodaran et al., 2014; Lindahl and Hellgren Kotaleski, 2016; Fountas and Shanahan, 2017; Girard et al., 2021), our analysis shows strong evidence for a more nuanced picture with FS interneurons preferentially projecting to D1 SPNs, NPY-NGF interneurons projecting to SPNs more densely than do FS interneurons, and both interneurons connecting to SPNs over a much further distance than the SPNs connect to each other. While recent models of the striatum include NPY-NGF interneurons in their

construction (Hjorth et al., 2020), and we know differences in the distance of connectivity within the striatum can radically alter its dynamics (Humphries et al., 2010; Spreizer et al., 2017), we have as yet no theory that accounts for any of these interneuron features, and so their functional role is unclear.

A Bayesian approach to microcircuit mapping can be used anywhere in the brain

We showed a range of advantages that our Bayesian approach has over more traditional frequentist approaches. A first qualitative advantage of Bayesian inference is that it replaces a single point estimate surrounded by a flat confidence interval by a posterior distribution covering all the possible values. One immediate consequence is more informative visualisations of the uncertainty in the data, as the potential overlap between connection rates becomes immediately apparent and estimated. Second, as argued in Dienes (2014), when differences between connection rates are non-significant, Bayesian methods allow us to distinguish between cases where the data is insufficient to draw a conclusion from cases where there really is no difference. For instance, the posterior curves for connection rates in the study of Cepeda et al. (2013) (Figure 8) overlap very much as though there were no difference. However, because these curves are so broad, we know this is due to insufficient data evidence. On the contrary, when we failed to find a difference in connection rates for different postsynaptic targets of the D1 and D2 SPNs, the posterior curves are sufficiently narrow for us to confidently conclude that such a difference is indeed either absent or quite small (Figure 2). Third, when making comparisons between probabilities of connection, we can use the full posteriors to compute an explicit probability that the difference is less than or greater than zero. We used this to calculate, for example, the explicit probabilities of whether D1 SPNs preferentially targeted D1 or D2 SPN subtypes (and concluded that they did not). A final advantage of Bayesian inference is its easy and explicit way of specifying our prior assumptions and incorporating past results. In our particular case, earlier literature had concluded that the average connection rate between SPNs of all subtypes was about 0.12, so we could incorporate this estimate as a data-driven prior. Better still, for the connections between SPN subtypes we could show that the estimates of connection probability were robust to our choice of prior, indicating we have sufficient experimental data for a robust estimate.

The methods we present here can easily be applied to any other brain regions where paired recording experiments have taken place, such as the recent study of Ellender et al. (2019) concerning connection rates of cortical neurons depending on their embryonic origin. Indeed, obtaining the posterior distributions given k positive test of connection and $n-k$ negative tests requires a single line of MATLAB or Python thanks to the built-in functions (for example, in MATLAB this is `posterior = betapdf(p, a + k, b + n-k)` with p a vector of probabilities of connection for which we want the corresponding probability density value, and a and b the shape parameters of the initial prior). Consequently, not only are these Bayesian methods easily applicable to intracellular recording data from any brain region, but also may be a rare case where it easier to be Bayesian than frequentist.

METHODS AND MATERIALS

Data processing

Because of the way Taverna et al. (2008) gave their results, namely reporting the number of connected pairs and specifying if any were bidirectional instead of reporting the number of connections, the number of tests for non-mixed pairs we use in this paper is doubled compared to the original study. For instance, when Taverna et al. (2008) say they found 5 connected pairs out of 19 pairs of D1 neurons, we interpret this as 5 connections out of 38 tests, to be consistent with the mixed D1 and D2 pairs which by necessity are unidirectional (a D1 $\rightarrow$ D2 connection can only be tested in one direction or it becomes a D2 $\rightarrow$ D1 connection). This was also the case for the data of Cepeda et al. (2013) on SPN connections in wild-type and Huntington's disease model mice.

Design of the prior based on previous literature

Knowing its mean μ and variance v, the shape parameters of a beta distribution are:

$$a = \mu\left(\frac{\mu(1-\mu)}{v} - 1\right) \tag{13}$$

$$b = (1-\mu)\left(\frac{\mu(1-\mu)}{v} - 1\right) \tag{14}$$

Previous studies have shown that SPNs connect to one another at a mean rate of 0.12, leaving us with a decision to make about the desired variance of the prior. Despite their thoroughness (325 tested pairs in Koos et al. (2004)), we could not directly use a beta distribution based on the number of pairs in the initial studies, as the resulting variance, which would reflect uncertainty attached to the measurement of the average connection rate between all types of pairs, would be so small that the new evidence with SPN subtype distinction would be unable to significantly affect the posterior. Indeed, the desired variance should reflect the fact that the average connection rate of 0.12 masks the potential existence of four distinct connection rates for each pair. We were unable to find a principled way of deriving this desired variance and, for this reason, different values of variance were tested before settling for 0.005 which gives the corresponding beta distribution a shape that makes such a prior both sufficiently informative as to be interesting without being completely insensitive to the addition of new data. Setting $\mu = 0.12$ and $v = 0.005$, we find $a = 2.56$ and $b = 18.12$.

Conversion of p into β

By definition of a probability density function, $f_{equi}(r)$ must be such that the probability that r is found between two arbitrary values r_1 and $r_1 + \Delta r$ is:

$$P(r_1 < r < r_1 + \Delta r) = \int_{r_1}^{r_1+\Delta r} f_{equi}(r)dr \tag{15}$$

This probability distribution for distance depends on how the experimenters sampled their pairs. A simple assumption would be that given a certain volume surrounding a central neuron, the probability of testing any given neuron is equiprobable for all neurons, hence why we chose to call the density function f_{equi}. So for a given distance r_1 from a central neuron, the probability of selecting a neuron within the volume bounded by r_1 and $r_1 + \Delta r$ is also equal to the ratio of the expected number of neurons found within it over the expected number of neurons in the total volume:

$$P(r_1 < r < r_1 + \Delta r) = \frac{N.V(r_1)}{N.V_{tot}} \tag{16}$$

with $V(r_1)$ the subvolume bounded by r_1 and $r_1 + \Delta r$, V_{tot} the total volume and N the density of SPNs of whichever given type experimenters are currently trying to sample. Note that N cancels out in the fraction, which implies that the probability distribution for distance is, counter-intuitively perhaps, independent of post-synaptic SPN subtype, as long as the density is constant everywhere. According to the reported methods of these two studies, experimenters selected neurons within the same field of focus at a maximum distance R of either $50\mu m$ (Taverna et al., 2008) or $100\mu m$ (Planert et al., 2010) which means that the total volume of interest is a cylinder of height h, corresponding to the depth of the field of focus, and the subvolume is a hollow cylinder, as depicted in Figure 3B:

$$V_{tot} = h\pi R^2 \tag{17}$$

$$V(r_1) = h\pi((r_1 + \Delta r)^2 - r_1^2) = 2\pi r_1 h\Delta r + \pi h\Delta r^2 \tag{18}$$

If we now combine the general definition of a probability density function (Eq 15) with this particular equiprobable sampling assumption, we now have:

$$\int_{r_1}^{r_1+\Delta r} f_{equi}(r)dr = \frac{2r_1\Delta r + \Delta r^2}{R^2} \tag{19}$$

which we can solve to find f_{equi} by differentiating the right hand side of the equation to obtain:

$$f_{equi}(r) = \frac{2r}{R^2} \tag{20}$$

So that finally, by plugging equation 20 into equation 7 we obtain:

$$p(\beta) = \frac{2}{R^2}\int_0^R re^{-\beta r}dr \tag{21}$$

which can be used to create a mapping from β to p, $p(\beta)$.

Transformation of posterior distributions

As explained in the main text, our objective is to make sure the two experimental studies are consistent with one another, despite the posteriors $f_p(p)$ of Taverna et al. (2008) being shifted towards generally higher values that the posteriors of Planert et al. (2010). To achieve this, we need to transform the posteriors we have previously obtained for p, $f_p(p)$, into posteriors for β, $f_\beta(\beta)$, to check that the decay rate at which probability of connection decreases with distance is the same. By definition of a density function, for any possible values a and b of β, we have:

$$\int_a^b f_\beta(\beta)d\beta = P(a < \beta < b) \tag{22}$$

Thanks to the mapping from β to p (which is monotonically decreasing), we can also write:

$$P(a < \beta < b) = P(p(b) < p < p(a)) = -\int_{p(a)}^{p(b)} f_p(p)dp \tag{23}$$

Finally, integration by substitution tells us:

$$-\int_{p(a)}^{p(b)} f_p(p)dp = -\int_a^b f_p(p(\beta))\frac{dp(\beta)}{d\beta}d\beta \tag{24}$$

Hence, by identification with equation 22:

$$f_\beta(\beta) = -f_p(p(\beta))\frac{dp(\beta)}{d\beta} \tag{25}$$

In order to draw f_β, we converted regularly interpolated values of β into the corresponding values of p using equation 21. Thanks to the functional definition of f_p obtained by Bayesian updating of the shape parameters of the beta distribution, we could precisely retrieve the corresponding values of the posterior for p. As for the derivative of p with respect to β in equation 25, this depends on the sampling process. In the case of equiprobable sampling, after an integration by parts of equation 21, we arrive at the following expression of $p(\beta)$:

$$p(\beta) = \frac{-2}{R^2\beta}(Re^{-\beta R} + \frac{1}{\beta}e^{-\beta R} - \frac{1}{\beta}) \tag{26}$$

which can be differentiated with respect to β:

$$\frac{dp(\beta)}{d\beta} = \frac{2}{R^2\beta^3}(e^{-\beta R}((R\beta+1)^2+1)-2) \tag{27}$$

However, in the case of the nearest-neighbour distribution, we were unable to find a closed-form for $\frac{dp(\beta)}{d\beta}$, and resorted to a numerical approximation based on the regularly interpolated values of β and the corresponding values of p given by equation 12.

Nearest-neighbour distribution

Because information about the nearest neighbour distribution was hard to find, we reproduce its derivation here, basing ourselves on Krider and Kehoe (2004), in case such a derivation would be of interest to others. Such a density function must satisfy:

$$f_{NN}(r) = (1 - \int_0^r f_{NN}(x)dx)2\pi rhN \tag{28}$$

which states that the probability density that the nearest neuron is found at distance r is the product of the probability that the first neuron is not in fact found at a shorter distance from the central neuron (the first element in brackets on the right hand side) and of the probability that there is a neuron between r and $r+dr$. This latter probability is itself the product of the infinitesimal cylindrical volume found between r and $r+dr$, i.e. $2\pi rhdr$ with h (μm) the height of the cylinder, and of N (μm^{-3}) the average density of neurons. If we now differentiate $f_{NN}(r)$, we get:

$$\frac{df_{NN}}{dr} = -\frac{d(\int_0^r f_{NN}(x)dx)}{dr}2\pi rhN + (1 - \int_0^r f_{NN}(x)dx)2\pi hN \tag{29}$$

According to Leibniz' rule, we get on the one hand:

$$\frac{d(\int_0^r f_{NN}(x)dx)}{dr} = f_{NN}(r) \tag{30}$$

We can also substitute the term in brackets in the second half of equation 29, thanks to the following rewriting of equation 28:

$$1 - \int_0^r f_{NN}(x)dx = \frac{f_{NN}(r)}{2\pi rhN} \tag{31}$$

After these substitutions and factoring by f_{NN}, we can now rewrite equation 29:

$$\frac{df_{NN}}{dr} = f_{NN}(r)(\frac{1}{r} - 2\pi rhN) \tag{32}$$

$$\frac{df_{NN}}{f_{NN}} = (\frac{1}{r} - 2\pi rhN)dr \tag{33}$$

$$ln(f_{NN}(r)) = ln(r) - \pi hNr^2 + constant \tag{34}$$

$$ln(\frac{f_{NN}(r)}{r}) = -\pi hNr^2 + constant \tag{35}$$

$$f_{NN}(r) = kre^{-\pi hNr^2} \tag{36}$$

with k, the normalisation constant.

Code availability

All code used for this work was written with MATLAB. The code necessary for the Bayesian analysis of p and the transformation of $f_p(p)$ to $f_\beta(\beta)$ using either an equiprobable or nearest-neighbour hypothesis is available on the Github account of the Humphries lab: `https://github.com/Humphries-Lab/Bayesian-map-of-striatum-circuitry`.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council [grant numbers MR/J008648/1, MR/P005659/1 and MR/S025944/1]. We thank Robert Schmidt and Benoît Girard for their helpful comments on a draft of this paper.

Author Contributions

F.C. and M.D.H designed the analyses. F.C. analysed the data. F.C. and M.D.H wrote the manuscript.

Declaration of Interest

The authors declare no conflicts of interest.

REFERENCES

Alexander, G. E. and Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends in Neuroscience*, 13:266–272.

Alexander, M. E. and Wickens, J. R. (1993). Analysis of striatal dynamics: the existence of two modes of behaviour. *J Theor Biol*, 163(4):413–438.

Bahuguna, J., Aertsen, A., and Kumar, A. (2015). Existence and Control of Go/No-Go Decision Transition Threshold in the Striatum. *PLoS Computational Biology*, 11(4):1–35.

Bogacz, R. and Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput*, 19(2):442–477.

Bornstein, A. M. and Daw, N. D. (2011). Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr Opin Neurobiol*, 21:374–380.

Brown, L. D., Cai, T. T., and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16(2):101–133.

Burke, D. A., Rotstein, H. G., and Alvarez, V. A. (2017). Striatal Local Circuitry: A New Framework for Lateral Inhibition. *Neuron*, 96(2):267–284.

Cepeda, C., Galvan, L., Holley, S. M., Rao, S. P., André, V. M., Botelho, E. P., Chen, J. Y., Watson, J. B., Deisseroth, K., and Levine, M. S. (2013). Multiple sources of striatal inhibition are differentially affected in Huntington's disease mouse models. *Journal of Neuroscience*, 33(17):7393–7406.

Czubayko, U. and Plenz, D. (2002). Fast synaptic transmission between striatal spiny projection neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24):15764–15769.

Damodaran, S., Evans, R. C., and Blackwell, K. T. (2014). Synchronized firing of fast-spiking interneurons is critical to maintain balanced firing between direct and indirect pathway neurons of the striatum. *J Neurophysiol*, 111(4):836–848.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5(July):1–17.

Ding, L. and Gold, J. I. (2010). Caudate encodes multiple computations for perceptual decisions. *J Neurosci*, 30(47):15747–15759.

Ding, L. and Gold, J. I. (2012). Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron*, 75(5):865–874.

Dorst, M. C., Tokarska, A., Zhou, M., Lee, K., Stagkourakis, S., Broberger, C., Masmanidis, S., and Silberberg, G. (2020). Polysynaptic inhibition between striatal cholinergic interneurons shapes their network activity patterns in a dopamine-dependent manner. *Nature Communications*, 11(1).

Du, K., Wu, Y.-W., Lindroos, R., Liu, Y., Rózsa, B., Katona, G., Ding, J. B., and Kotaleski, J. H. (2017). Cell-type-specific inhibition of the dendritic plateau potential in striatal spiny projection neurons. *Proc Natl Acad Sci USA*, 114:E7612–E7621.

Ellender, T. J., Avery, S. V., Mahfooz, K., Scaber, J., von Klemperer, A., Nixon, S. L., Buchan, M. J., van Rheede, J. J., Gatti, A., Waites, C., Pavlou, H. J., Sims, D., Newey, S. E., and Akerman, C. J. (2019). Embryonic progenitor pools generate diversity in fine-scale excitatory cortical subnetworks. *Nature Communications*, 10(1).

English, D. F., Ibanez-Sandoval, O., Stark, E., Tecuapetla, F., Buzsáki, G., Deisseroth, K., Tepper, J. M., and Koos, T. (2012). GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. *Nature Neuroscience*, 15(1):123–130.

Fountas, Z. and Shanahan, M. (2017). The role of cortical oscillations in a spiking neural network model of the basal ganglia. *PloS one*, 12:e0189109.

Frank, Michael, J. (2005). Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, 17(1):51–72.

Franklin, N. T. and Frank, M. J. (2015). A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. *eLife*, 4.

Fujiyama, F., Sohn, J., Nakano, T., Furuta, T., Nakamura, K. C., Matsuda, W., and Kaneko, T. (2011). Exclusive and common targets of neostriatofugal projections of rat striosome neurons: A single neuron-tracing study using a viral vector. *European Journal of Neuroscience*, 33(4):668–677.

Gagnon, D., Petryszyn, S., Sanchez, M. G., Bories, C., Beaulieu, J. M., De Koninck, Y., Parent, A., and Parent, M. (2017). Striatal Neurons Expressing D1 and D2 Receptors are Morphologically Distinct and Differently Affected by Dopamine Denervation in Mice. *Scientific Reports*, 7(December 2016):9–17.

Gerfen, C., Engber, T., Mahan, L., Susel, Z., Chase, T., Monsma, F., and Sibley, D. (1990). D1 and d2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science*, 250:1429–1432.

Gerfen, C. R. and Surmeier, D. J. (2011). Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci*, 34:441–466.

Gertler, T. S., Chan, C. S., and Surmeier, D. J. (2008). Dichotomous Anatomical Properties of Adult Striatal Medium Spiny Neurons. *Journal of Neuroscience*, 28(43):10814–10824.

Girard, B., Lienard, J., Gutierrez, C. E., Delord, B., and Doya, K. (2021). A biologically constrained spiking neural network model of the primate basal ganglia with overlapping pathways exhibits action selection. *European Journal of Neuroscience*, 53:2254–2277.

Gittis, A. H., Nelson, A. B., Thwin, M. T., Palop, J. J., and Kreitzer, A. C. (2010). Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. *Journal of Neuroscience*, 30(6):2223–2234.

Groves, P. M. (1983). A theory of the functional organization of the neostriatum and the neostriatal control of voluntary movement. *Brain Research Reviews*, 5(2):109–132.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, 84(6):401–410.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, 84(6):411–423.

Gurney, K. N., Humphries, M. D., and Redgrave, P. (2015). A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. *PLoS Biology*, 13(1).

Hellwig, B. (2000). A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. *Biol Cybern*, 82(2):111–121.

Hjorth, J., Blackwell, K. T., and Kotaleski, J. H. (2009). Gap junctions between striatal fast-spiking interneurons regulate spiking activity and synchronization as a function of cortical activity. *J Neurosci*, 29(16):5276–5286.

Hjorth, J. J., Kozlov, A., Carannante, I., Nylén, J. F., Lindroos, R., Johansson, Y., Tokarska, A., Dorst, M. C., Suryanarayana, S. M., Silberberg, G., Kotaleski, J. H., and Grillner, S. (2020). The microcircuits of striatum in silico. *Proceedings of the National Academy of Sciences of the United States of America*, 117(17):9554–9565.

Humphries, M. D., Stewart, R. D., and Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *Journal of Neuroscience*, 26(50):12921–12942.

Humphries, M. D., Wood, R., and Gurney, K. (2009). Dopamine-modulated dynamic cell assemblies generated by the GABAergic striatal microcircuit. *Neural Networks*, 22(8):1174–1188.

Humphries, M. D., Wood, R., and Gurney, K. (2010). Reconstructing the three-dimensional gabaergic microcircuit of the striatum. *PLoS Computational Biology*, 6(11).

Ibáñez-Sandoval, O., Tecuapetla, F., Unal, B., Shah, F., Koós, T., and Tepper, J. M. (2011). A novel functionally distinct subtype of striatal neuropeptide Y interneuron. *Journal of Neuroscience*, 31(46):16757–16769.

Jaeger, D., Kita, H., and Wilson, C. J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. *Journal of Neurophysiology*, 72(5):2555–2558.

Kalisman, N., Silberberg, G., and Markram, H. (2003). Deriving physical connectivity from neuronal morphology. *Biol Cybern*, 88(3):210–218.

Kawaguchi, Y. (1993). Physiological, morphological, and histochemical characterization of three classes of interneurons in rat neostriatum. *Journal of Neuroscience*, 13(11):4908–4923.

Khamassi, M. and Humphries, M. D. (2012). Integrating cortico-limbic-basal ganglia architectures for learning model-based and model-free navigation strategies. *Front Behav Neurosci*, 6:79.

Klaus, A., Planert, H., Hjorth, J. J. J., Berke, J. D., Silberberg, G., and Kotaleski, J. H. (2011). Striatal fast-spiking interneurons: from firing patterns to postsynaptic impact. *Front Syst Neurosci*, 5:57.

Koos, T. and Tepper, J. M. (1999). Inhibitory control of neostriatal projection neurons by gabaergic interneurons. *Nature Neuroscience*, 2(5):467–472.

Koos, T., Tepper, J. M., and Wilson, C. J. (2004). Comparison of IPSCs evoked by spiny and fast-spiking neurons in the neostriatum. *Journal of Neuroscience*, 24(36):7916–7922.

Krajeski, R. N., Macey-Dare, A., Heusden, F., Ebrahimjee, F., and Ellender, T. J. (2019). Dynamic postnatal development of the cellular and circuit properties of striatal D1 and D2 spiny projection neurons. *The Journal of Physiology*, 21:5265–5293.

Kreitzer, A. C. (2009). Physiology and Pharmacology of Striatal Neurons. *Annual Review of Neuroscience*, 32(1):127–147.

Krider, E. P. and Kehoe, K. E. (2004). On Quantifying the Exposure to Cloud-to-Ground Lightning. *27th International Conference on Lightning Protection (ICLP)*, pages 1–6.

Lau, T., Gage, G. J., Berke, J. D., and Zochowski, M. (2010). Local dynamics of gap-junction-coupled interneuron networks. *Phys Biol*, 7:16015.

Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *J Neurosci*, 26(13):3567–3583.

Liénard, J. and Girard, B. (2014). A biologically constrained model of the whole basal ganglia addressing the paradoxes of connections and selection. *Journal of Computational Neuroscience*, 36(3):445–468.

Liley, D. T. J. and Wright, J. J. (1994). Intracortical connectivity of pyramidal and stellate cells: estimates

of synaptic densities and coupling symmetry. *Network*, 5:175–189.

Lindahl, M. and Hellgren Kotaleski, J. (2016). Untangling basal ganglia network dynamics and function: Role of dopamine depletion and inhibition investigated in a spiking network model. *eNeuro*, 3:ENEURO.0156–16.2016.

Makin, T. R. and De Xivry, J. J. O. (2019). Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *eLife*, 8:1–13.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, 50:381–425.

Oorschot, D. E. (1996). Total number of neurons in the neostriatal, pallidal, subthalamic, and substantia nigral nuclei of the rat basal ganglia: a stereological study using the cavalieri and optical disector methods. *Journal of Comparative Neurology*, 366(4):580–99.

Oorschot, D. E. (2012). The percentage of interneurons in the dorsal striatum of the rat, cat, monkey and human: A critique of the evidence. *Basal Ganglia*, 3:19–24.

Oorschot, D. E., Lin, N., Cooper, B. H., Reynolds, J. N., Sun, H., and Wickens, J. R. (2013). Synaptic connectivity between rat striatal spiny projection neurons in vivo: Unexpected multiple somatic innervation in the context of overall sparse proximal connectivity. *Basal Ganglia*, 3:93–108.

Panas, D., Amin, H., Maccione, A., Muthmann, O., van Rossum, M., Berdondini, L., and Hennig, M. H. (2015). Sloppiness in spontaneously active neuronal networks. *J Neurosci*, 35(22):8480–8492.

Penny, G. R., Wilson, C. J., and Kitai, S. T. (1988). Relationship of the axonal and dendritic geometry of spiny projection neurons to the compartmental organization of the neostriatum. *Journal of Comparative Neurology*, 269:275–289.

Planert, H., Szydlowski, S. N., Hjorth, J. J. J., Grillner, S., and Silberberg, G. (2010). Dynamics of Synaptic Transmission between Fast-Spiking Interneurons and Striatal Projection Neurons of the Direct and Indirect Pathways. *Journal of Neuroscience*, 30(9):3499–3507.

Plenz, D. (2003). When inhibition goes incognito: Feedback interaction between spiny projection neurons in striatal function. *Trends in Neurosciences*, 26(8):436–443.

Ponce-Alvarez, A., Mochol, G., Hermoso-Mendizabal, A., de la Rocha, J., and Deco, G. (2020). Cortical state transitions and stimulus response evolve along stiff and sloppy parameter dimensions, respectively. *eLife*, 9.

Ponzi, A. and Wickens, J. (2010). Sequentially switching cell assemblies in random inhibitory networks of spiking neurons in the striatum. *J Neurosci*, 30(17):5894–5911.

Ponzi, A. and Wickens, J. R. (2013). Optimal balance of the striatal medium spiny neuron network. *PLoS Comput Biol*, 9(4):e1002954.

Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem?

Reynolds, J. N., Hyland, B. I., and Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, 413:67–70.

Rosen, G. D. and Williams, R. W. (2001). Complex trait analysis of the mouse striatum: Independent QTLs modulate volume and neuron number. *BMC Neuroscience*, 2.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, 310(25):1337–1340.

Spreizer, S., Angelhuber, M., Bahuguna, J., Aertsen, A., and Kumar, A. (2017). Activity dynamics and signal representation in a striatal network model with distance-dependent connectivity. *eNeuro*, 4(4):1–15.

Stocco, A. and Lebiere, C. (2014). Inhibitory synapses between striatal projection neurons support efficient enhancement of cortical signals: a computational model. *J Comput Neurosci*, 37(1):65–80.

Straub, C., Saulnier, J. L., Bègue, A., Feng, D. D., Huang, K. W., and Sabatini, B. L. (2016). Principles of Synaptic Organization of GABAergic Interneurons in the Striatum. *Neuron*, 92(1):84–92.

Szydlowski, S. N., Pollak Dorocic, I., Planert, H., Carlén, M., Meletis, K., and Silberberg, G. (2013). Target selectivity of feedforward inhibition by striatal fast-spiking interneurons. *Journal of Neuroscience*, 33(4):1678–1683.

Taverna, S., Ilijic, E., and Surmeier, D. J. (2008). Recurrent Collateral Connections of Striatal Medium Spiny Neurons Are Disrupted in Models of Parkinson's Disease. *Journal of Neuroscience*, 28(21):5504–5512.

Taverna, S., Van Dongen, Y. C., Groenewegen, H. J., and Pennartz, C. M. (2004). Direct Physiological

Evidence for Synaptic Connectivity between Medium-Sized Spiny Neurons in Rat Nucleus Accumbens In Situ. *Journal of Neurophysiology*, 91(3):1111–1121.

Tepper, J. M., Koós, T., and Wilson, C. J. (2004). GABAergic microcircuits in the neostriatum. *Trends in Neurosciences*, 27(11):662–669.

Tepper, J. M., Koós, T., Ibanez-Sandoval, O., Tecuapetla, F., Faust, T. W., and Assous, M. (2018). Heterogeneity and diversity of striatal gabaergic interneurons: Update 2018. *Frontiers in neuroanatomy*, 12:91.

Tepper, J. M., Wilson, C. J., and Koós, T. (2008). Feedforward and feedback inhibition in neostriatal GABAergic spiny neurons. *Brain Research Reviews*, 58(2):272–281.

Tomkins, A., Vasilaki, E., Beste, C., Gurney, K., and Humphries, M. D. (2014). Transient and steady-state selection in the striatal microcircuit. *Frontiers in Computational Neuroscience*, 7(January):1–16.

Wickens, J. R., Alexander, M. E., and Miller, R. (1991). Two dynamic modes of striatal function under dopaminergic-cholinergic control: simulation and analysis of a model. *Synapse*, 8(1):1–12.

Yartsev, M. M., Hanks, T. D., Yoon, A. M., and Brody, C. D. (2018). Causal contribution and dynamical encoding in the striatum during evidence accumulation. *eLife*, 7:34929.